From Small Molecules to Infinite Crystals: Probing Noncovalent Interactions through Vibrational Spectroscopy

Sadisha Nanayakkara
Southern Methodist University, snanayakkara@smu.edu

Follow this and additional works at: https://scholar.smu.edu/hum_scichemistry_etds

Part of the Materials Chemistry Commons, Other Chemistry Commons, and the Physical Chemistry Commons

Recommended Citation
https://scholar.smu.edu/hum_scichemistry_etds/23

This Dissertation is brought to you for free and open access by the Chemistry at SMU Scholar. It has been accepted for inclusion in Chemistry Theses and Dissertations by an authorized administrator of SMU Scholar. For more information, please visit http://digitalrepository.smu.edu.
FROM SMALL MOLECULES TO INFINITE CRYSTALS:
PROBING NONCOVALENT INTERACTIONS THROUGH
VIBRATIONAL SPECTROSCOPY

Approved by:

Dr. Elfi Kraka
Professor of Chemistry

Dr. Tomče Runčevski
Assistant Professor of Chemistry

Dr. Oliviero Andreussi
Assistant Professor of Physics
FROM SMALL MOLECULES TO INFINITE CRYSTALS:
PROBING NONCOVALENT INTERACTIONS THROUGH
VIBRATIONAL SPECTROSCOPY

A Dissertation Presented to the Graduate Faculty of the
Dedman College
Southern Methodist University
in
Partial Fulfillment of the Requirements
for the degree of
Doctor of Philosophy
with a
Major in Theoretical and Computational Chemistry
by
Sadisha Manendri Nanayakkara
B. Sc., Chemistry, University of Colombo, Sri Lanka

August 4, 2021
ACKNOWLEDGMENTS

I want to express my profound and sincere gratitude to my research advisor Prof. Elfi Kraka, for her continuous support and guidance from the Day 1 of my Ph. D. study and research. Her bursting energy, constant drive for research, and passion for science as a whole was contagious and it motivated me to be a hard worker and a better researcher myself. I am extremely grateful to her for showing me the ropes first and then altogether handing me the reins, giving me full freedom to try out bold ideas and be independent. Without her genuine help and understanding, both in my professional and personal life, I wouldn’t have made it this far.

I am also grateful to my incredible Ph. D. committee members, Dr. Tomče Runčevski and Dr. Oliviero Andreussi for investing time in me, their invaluable comments regarding my work, and supporting me in many other ways.

I want to thank profusely Dr. Marek Friendorf for being a leading figure among the many to shape up my scientific journey in the CATCO group. Besides his creative and innovative ideas that lead to some of the projects I worked on, he imparted knowledge in numerous other ways too by giving lectures and workshops, for all of which I feel highly appreciative. Another person whom I have drawn so much inspiration and insights from is Dr. Yunwen Tao and I want to thank him for his constant guidance, stimulating conversations, and never hesitating to offer me a helping hand. Also, I want to acknowledge Dr. Velmurugan Gunasekaran, Dr. Daniel Sethio, Dr. Pratanphorn Nakliang, and Dr. Wenli Zou, for nourishing various aspects of my research and sharing their expertise. I also want to kindly remember late Dr. Dieter Cremer who laid the ground works for CATCO and was a source of great inspiration for many freshly starting researches like myself.

My heartfelt thanks also goes to the current and former members of the Computational and Theoretical Chemistry Group (CATCO) and all my friends here at SMU, for being
amazing colleagues and making my graduate life memorable. Specially, Małgorzata Makoś for apart from being my teammate at CATCO, also being the best roommate one can ask for; Niraj Verma for readily sharing his talents, particularly his knack for coding; and numerous other people whom I might have failed to mention here but regardlessly have vastly enriched my experience at SMU.

I want to sincerely acknowledge the entire faculty and staff in the SMU Chemistry Department. A special tribute to Dr. Peng Tao who taught me Advance Computational Chemistry during my first year and with whom I had the greatest pleasure of working as a teaching assistant. I want to thank for his thought provoking lectures and easygoing nature as TA mentor. My kind regards also goes to my other TA mentors during my first year, Ms. Andrea Adams, Ms. Helen Babbili, and Dr. Jennifer O’Brien. This list would not be complete without thanking heartily Dr. Michael Lattman, our graduate program director and Mandy Graham, our department coordinator for their support in multiple occasions.

I want to thank all the dear friends I have made during my time in Dallas, specially Nipunika Somatilaka and Amila Nanayakkara who had literally been there in my health and sickness. I cherish their friendship, kind heartedness and thank them for providing me a home away from home. Even if it is not really a person, I want to fondly thank their furry friend Leo the bunny who did a great job in cheering me up at times. I am also forever grateful to Darshan Karunarathne who in the first place encouraged me to apply for SMU and ever since been a supportive and a great friend.

I want to immensely thank all those who are closest to my heart. This includes my father and mother, who had been my utmost strength all the time, my siblings, Dushan, Lahiru, and Shashini, along with my sisters-in-law, Mahesha and Sanda for being a really caring bunch of people. I also want to thank Dumindu for considering my struggles as his own and riding out the storm with me. His empathy has always revived me and helped me bounce back from hard periods in life. My thanks also goes to Sandya aunty for her kindness and care. I also want to thank my best friends Erandi, Nisitha, Dulashani, Malsha, Chathuni, Ishani, Ayeshini, Vindi, Wathsala for being around for a cozy chat and cheering me with
uplifting words.

Last but not least, I want to thank all the people who have supported and encouraged me along the way; mentors and friends from my undergraduate studies; all my teachers since pre-school; my former workplace friends, and many other people whom I wish I had enough space to spare. A big thank you to all for being a part of my journey.
During my four years in the CATCO group supervised by Dr. Kraka, I have been exposed to various topics targeted to address a multitude of chemical problems, broadening our knowledge of modern chemistry research as we know it. In undertaking the age-old problem of understanding the nature of chemical bonding, I could modestly contribute with my work, using in-house tools based on vibrational spectroscopy.

First part of my dissertation is focused on inventing new methods and tools to efficiently investigate chemical bonding, followed by the study of some non-covalent interactions, imperative in catalysis and solid-state chemistry. This includes analysis of hydrogen bonds in trivial systems like small dimers which is all the way extended to complex hydrogen-bond networks in ice crystals. Also, as dihydrogen-bonded complexes are key intermediates in catalytically driven hydrogen generation from water, their role is probed for a series of B- and Al-based hydride catalyzed reactions. The nature of much debated H···H interactions in sterically congested systems like naphthalene derivatives are also examined. The insights offered by these combined efforts will guide the future experiments and help in new material and catalyst design.
# TABLE OF CONTENTS

LIST OF FIGURES.............................................................................................................. viii

CHAPTER

1. INTRODUCTION ........................................................................................................... 1
    1.1. Summary of Publications .................................................................................. 1
    1.2. Overview ........................................................................................................... 2
    1.3. General Introduction ....................................................................................... 2

2. COMPUTATIONAL METHODS ...................................................................................... 6
    2.1. Local Vibrational Mode Theory and Extension to Crystals ..................... 6
    2.2. Unified Reaction Valley Approach ............................................................... 8

3. DEVELOPING NEW METHODS & TOOLS TO STUDY CHEMICAL BONDING AND VIBRATIONS ............................................................................................. 10
    3.1. A Novel Approach to Study Bond Formation and Cleavage ................. 10
    3.2. PyVibMS for Visualizing Vibrations in Molecules and Solids ............... 13
    3.3. A Revised Generalized Subsystem Vibrational Analysis ....................... 14

4. HYDROGEN BONDING, DIHYDROGEN BONDING & H···H INTERACTIONS 15
    4.1. Hydrogen Bonds in Small Dimers ................................................................. 15
    4.2. Hydrogen Bonds in Ice Crystals .................................................................. 17
    4.3. Dihydrogen Bonds in Catalysis .................................................................. 26
    4.4. H···H Interactions in Naphthols .................................................................. 29

5. HALOGEN BONDING .................................................................................................. 32
    5.1. Halogen Bonds in Crystals .......................................................................... 32

BIBLIOGRAPHY ................................................................................................................. 33

APPENDIX

vii
A. Paper on URVA and QTAIM .................................................. 48
B. Paper on hydrogen-bonded dimers ............................................ 61
C. Paper on hydrogen bonds in ice ................................................ 68
D. Paper on hydrogen release from water ........................................ 107
E. Paper on PyVibMS ............................................................... 124
F. Paper on GSVA ................................................................. 137
G. Paper on H· · ·H interactions .................................................. 143
H. Paper on halogen bonds in crystals ............................................ 154
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>A summary of some known methods which monitor the development of either PES or ED related properties along the reaction path to identify bond formation/cleavage events. Figure adopted from Ref. 1.</td>
</tr>
<tr>
<td>4.1</td>
<td>H-bond networks of ten ices investigated in this work. Primitive unit cells are indicated in black. For order–disorder pairs, their (dis)order counterparts are also given in black text.</td>
</tr>
<tr>
<td>4.2</td>
<td>Comparison between H-bond strength distribution; (a) uniform in ice Ih and (b) inhomogeneous in ice VI/XV. H-bonds with same/different strengths are indicated by same/different color(s).</td>
</tr>
<tr>
<td>4.3</td>
<td>The statistical distribution of H-bond strengths in 16 isomers via a box–and–whisker diagram. The blue lines connecting mean (dotted) and median (solid) H-bond strengths are there only to guide the eye.</td>
</tr>
<tr>
<td>4.4</td>
<td>Relationship between local stretching force constants $k_n^a$ and bond lengths $r$ for 7 predicted ices. A power function in the form $k_n^a = a \cdot r^b$ was used to fit data.</td>
</tr>
<tr>
<td>4.5</td>
<td>The statistical distribution of H-bond strengths in 7 predicted ices via a box–and–whisker diagram. Median is indicated.</td>
</tr>
<tr>
<td>4.6</td>
<td>Schematic representation of all starting reaction complexes investigated in this work. Hydrogens participating in a dihydrogen bond later in the reaction mechanism are labeled.</td>
</tr>
<tr>
<td>4.7</td>
<td>CNM analysis of cis-1-naphthol. O–H local stretching mode is indicated with red arrows and its % Contribution to O–H normal mode is shown in yellow. Numbering of local modes are as in the scheme.</td>
</tr>
</tbody>
</table>
To Niketh my nephew and Wilson the cat
who taught me sometimes it helps
to see the world upside down and knock a few things around
1.1. Summary of Publications

The following provides a summary of all my first- and second-authored publications during my four year period in CATCO. The details of these work are given in Chapter 3–5.

Table 1.1. Overview of publication list during 2017–2021.

<table>
<thead>
<tr>
<th>Project name</th>
<th>Journal (Status) $^a$</th>
<th>Authorship $^b$</th>
<th>Chapter/Appendix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bond form/break</td>
<td><em>Phys. Chem. Chem. Phys.</em> (P)</td>
<td>1$^{st}$</td>
<td>3.1/A</td>
</tr>
<tr>
<td>H-bonds in dimers</td>
<td><em>J. Comput. Chem.</em> (P)</td>
<td>1*$^{st}$</td>
<td>4.1/B</td>
</tr>
<tr>
<td>H-bonds in ice</td>
<td><em>J. Chem. Theory Comput.</em> (UR)</td>
<td>1*$^{st}$</td>
<td>4.2/C</td>
</tr>
<tr>
<td>H$_2$ from water</td>
<td><em>J. Phys. Chem. A</em> (P)</td>
<td>1$^{st}$</td>
<td>4.3/D</td>
</tr>
<tr>
<td>H···H interactions</td>
<td><em>Phys. Chem. Chem. Phys.</em> (UR)</td>
<td>2$^{nd}$</td>
<td>4.4/G</td>
</tr>
<tr>
<td>Halogen bonds</td>
<td><em>Molecules</em> (P)</td>
<td>2$^{nd}$</td>
<td>5/H</td>
</tr>
<tr>
<td>PyVibMS</td>
<td><em>J. Mol. Model.</em> (P)</td>
<td>2$^{nd}$</td>
<td>3.2/E</td>
</tr>
<tr>
<td>GSVA</td>
<td><em>Theor. Chem. Acc.</em> (P)</td>
<td>2$^{nd}$</td>
<td>3.3/F</td>
</tr>
</tbody>
</table>

$^a$ P: published; UR: under review

$^b$ 1$^{st}$: first-author; 1*$^{st}$: equal contribution; 2$^{nd}$: second-author
1.2. Overview

This dissertation offers a window into the gist of my work carried out during my four year journey in the Computational and Theoretical Chemistry Group here at SMU. From the very outset, my research focus was on understanding nature of chemical bonding, in particular non-covalent interactions like hydrogen bonding, utilizing in-house tools and methods derived from vibrational analysis described in Chapter 2.

On the path to realize my research goals, I have been motivated to develop new methods and tools that would provide efficient and more accurate means to probe chemical bonding and analyze vibrations in molecules. Towards the later years of my research, I grew more bold and extended my interest from non-covalent interactions in small molecular systems to periodic systems, leading to my interesting work on studying hydrogen bonds in ice crystals, which I personally took so much pleasure working on.

1.3. General Introduction

An improved understanding of chemical bonding ensues the use of new methods and tools and different perspectives to reveal insights unknown hitherto. In my work, in addition to bringing new approaches to study chemical bonding in general, I focused on investigating a set of highly important non-covalent interactions and their potential utilities. Entirety of my research work envelops, probing hydrogen bonds in dimers and ice crystals, evaluating reaction mechanisms of dihydrogen assisted catalysis, and examining controversial H···H interactions.

In tackling the age old problem of describing the nature of chemical bonding, chemists have come a long way from simple theories like VSEPR\textsuperscript{1} theory to more elegant theories based on quantum chemistry. These theories/methods aimed at a quantum chemical description of the chemical bond often exploit either features of the potential energy surface (PES), like energy and forces, or on the other hand features of electron density (ED), like topo-

\textsuperscript{1}Valence Shell Electron Pair Repulsion
logical properties; specially when they concern formation and breaking of chemical bonds. In this context, Hohenberg-Kohn theorem [3] is remarkable in manifesting the link between energy and electron density. Even though, we do not know the explicit form of this link, establishing the latter has been the breaking grounds for modern day quantum chemical theories like density functional theory. Driven by these forerunning theories/methods and a chemist’s fundamental desire to understand the nature of chemical bonding better, I thought of constructing a new combined approach in my work, which draws information from both PES- and ED-based approaches and uses this to efficiently describe a formation/cleavage of a chemical bond. Chapter 3.1 is dedicated to discussing details of this work. I further battered around this topic and this time the focus was on utilizing vibrational analysis to extract information about chemical bonding. This lead to my involvement in two other work, which introduces innovative ways to employ force constant matrix, i.e., Hessian matrix, to arrive at an in-depth analysis of chemical bonding. Chapter 3.2–3.3 is dedicated to discussing details of these work.

Within the framework of studying chemical bonding and vibrations, I pinpointed my focus on non-covalent interactions. One such important subset of interactions being those which involve hydrogen as a participating element. This can lead to a whole spectrum of possible interactions and in my work I pursued some of the most important non-covalent interactions; hydrogen bonding, dihydrogen bonding, and H···H interactions. Hydrogen bonds play an imperative role in a myriad of systems which range from governing structure and dynamics of liquid water [4–6] to more complex functions like rendering DNA environment [7, 8] and assisting in protein folding. [9, 10] In my work, I first started with a critical analysis of hydrogen-bonded dimers, addressing some crucial factors that can be overlooked and result in misinterpretation of hydrogen bond strengths, even within these simple dimer systems. Chapter 4.1 is dedicated to discussing details of this work. Having tackled these trivial systems, I expanded my interest into hydrogen bonds in more convoluted settings as in ice crystals. In my work, I characterized the hydrogen bond networks in several known ices, from which valuable insights about new ice forms could be derived. Chapter 4.2 is dedicated
to discussing details of this work.

Another type of important non-covalent interactions that stem from H atoms are dihydrogen bonds which are often referred to as an unconventional type of hydrogen bonds since they have comparable interaction energies. [11] Dihydrogen bonds are formed between chemically distinguishable hydridic (H$^{\delta-}$) and protonic (H$^{\delta+}$) hydrogen atoms and have useful applications in crystal engineering, solid state chemistry, and catalysis. In my work, I focused on the latter, where dihydrogen bonded complexes act as key intermediates in catalytic reactions such as in hydrogen evolution. Here, considering their potential utility as chemical hydrogen storage compounds, [12, 13] I targeted small main group element hydrides as boranes, alanes and their derivatives. Specifically, I investigated the mechanism of hydrogen release by these compounds hydrolytically, i.e., by reacting with water, using our in-house tools. Chapter 4.3 is dedicated to discussing details of this work. By gaining a profound understanding of dihydrogen bonds, I then moved onto analyze H···H interactions which differ from the dihydrogen bonds in the sense that they occur in sterically congested environments such as in bay region of phenanthrene [14] and have unfavorable polarity (e.g. C–H$^{\delta+}$$\cdots$$\delta^+$$H$). The nature of these H···H contacts has been the subject of much debate, as to whether they qualify as bonding interactions, [15] but a number of compelling studies have proven to the contrary. [16, 17] I could get involved in one such study where they investigated the nature of interaction between H atoms in close proximity in 1-naphthol. For the latter, AIM$^2$ analysis suggested a bonding interaction, but careful cross examination by our in-house tools confirmed the traditional view of a steric repulsion. Chapter 4.4 is dedicated to discussing details of this work.

Overlapped with my interest to study non-covalent interactions in periodic systems, I got to participate in the exploration of another type of important non-covalent interactions; halogen bonds in crystals. The primary authors were interested in characterizing X–I$\cdots$OA-type (X= I or Cl; OA: carbonyl, ether, and N-oxide groups) halogen bonding in molecular

$^2$Atoms-in-Molecule
crystals, considering their high bond strength tunability that lead to potential applications in crystal engineering [18] and material science. [19] This work could demonstrate how bond strengths of I–I and I–Cl bonds in crystals can be fine tuned by using different oxygen-containing acceptor molecules and modifying substituents. Chapter 5 briefly reflects details of this work.
2.1. Local Vibrational Mode Theory and Extension to Crystals

Vibrational spectroscopy has become an indispensable tool in extracting intricate details about electronic structure and chemical bonding. In a diatomic molecule, stretching force constant obtained via normal mode analysis serves as a quantitative measure of intrinsic bond strength of its chemical bond. However, this kind of powerful analysis is dampened in a polyatomic molecule by the fact that normal vibrational modes are no longer localized in a specific chemical bond, but delocalized all over the molecule as a result of electronic and mass mode-mode coupling. [20–23] Thus, Konkoli and Cremer drafted a successful solution to this problem via the development of local vibrational mode theory. [24–28] The fundamental idea was to derive local stretching force constants from local vibrational modes, free from any mode-mode coupling, which will then serve as a direct measure of intrinsic strength of a chemical bond. In dealing with the two different coupling mechanisms, mass coupling and electronic coupling, the latter was already eliminated by solving the Wilson equation of spectroscopy. [20,29] Then, by solving the mass-decoupled analogue [24] of Wilson’s equation, local vibrational modes (LVMs) could be derived which were associated with a given internal coordinate ($q_n$) such as bond length, bond angle, dihedral, or puckering coordinate. [30]

After the surfacing of the LVM theory, Zou and Cremer [31] justified the use of the local stretching force constant $k_n^a$ as a unique and universal measure of intrinsic strength of a chemical bond/interaction. They showed that $k_n^a$ yields the curvature of the potential energy surface (PES) in the direction of the bond stretching defined by $q_n$. Thus, $k_n^a$ not only captures pure electronic effects being independent from masses but also characterizes intrinsic bond strength, thereby serving as an ideal bond strength descriptor. Also, Zou and
co-workers [32] proved that there is a one-to-one correspondence between local and normal vibrational modes as verified via an adiabatic connection scheme (ACS), which allows a smooth transition from local to normal modes. Additionally, this led to the characterization of normal modes (CNM) procedure, [26,33,34] where each normal vibrational mode could be decomposed into their respective local mode counterparts, stretching the analytic potential of LVM theory even beyond its general role as a bond strength descriptor. Thus, LVM analysis has advanced as a powerful analytical tool and ever since has been extensively applied to a broad range of chemical systems covering both covalent bonds [31,35–46] and non-covalent interactions including hydrogen bond, [4,7,33,47–52] halogen bond, [53–58] pnictogen bond, [59–61] chalcogen bond, [45,62,63] tetrel bond, [64] and atom-π interactions. [65,66]

One of the recent and important milestones in LVM theory was its extension to periodic one-dimensional (1D) through three-dimensional (3D) systems, i.e. crystals. [67] A local vibrational mode in a periodic system is defined as a vibration driven by a specific internal coordinate in all primitive cells while relaxing all other parts of the periodic system. This approach allows for a head-to-head comparison of the intrinsic bond strength in periodic and molecular systems. Also, it does not suffer from certain restrictions imposed by periodicity; e.g. lattice structure prevents the calculation of bond dissociation energies, or the reduced functionality of some standard analytical tools; e.g. periodic NBO [68] does not provide second order perturbation energies. Due to these superiorities, periodic LVM analysis has already made its debut in to the crystal domain and is being used to gain a deeper understanding of several types of crystal bonding. [67,69,70]

In my work, LVM analysis and its extension to periodic systems were successfully applied to assess bond strengths of hydrogen bonded dimers (Chapter 4.1), hydrogen bond networks in ice crystals (Chapter 4.2), dihydrogen bonded intermediates of catalytically driven hydrogen release reactions (Chapter 4.3), and H···H interactions in naphthols (Chapter 4.4). Among the work which I contributed to as a second author, one study made use of periodic LVM analysis for studying halogen bonds in crystals (Chapter 5) and two other work used LVM theory as a precursor to derive new tools/methods related to vibration analysis.
2.2. Unified Reaction Valley Approach

Reaction path hamiltonian (RPH) approach [71] by Miller, Handy, and Adams enables to describe a chemical reaction within a specific region of the potential energy surface (PES), i.e. reaction valley, which encompasses the reaction path connecting reactants, TS, and products. URVA exploits the RPH approach exhaustively to retrieve in-depth mechanistic information about a chemical reaction by focusing on the curving of the reaction path. [72–76] As a reaction proceeds along the reaction path, electronic structure changes of the reacting complex is reflected in the changes of the normal vibrational modes. Consequently, mode-mode coupling between the vibrational and translational motion along the reaction path is changed, which produces the curvature of the reaction path. In practice, scalar curvature $\kappa(s)$ is monitored where $s$ is the arc length of the reaction path. Each chemical reaction is characterized by a unique curvature profile, with curvature minima related to minimal chemical change which are termed as hidden intermediates, [42,77–79] and curvature maxima representing locations of significant chemical changes as bond breaking/forming, charge polarization and transfer, rehybridization, etc. Accordingly, the region flanked by two curvature minima enclosing a curvature peak is defined as a reaction phase. [74,77] The decomposition of $\kappa(s)$ into internal coordinate components $q_n$ describing the reacting complex, [80] unravels their individual contributions to the curvature, leading to a comprehensive analysis of the reaction mechanism. A positive contribution to the $\kappa(s)$ supports the chemical event while a negative contribution indicates resistance. URVA has been used as an advanced mechanistic tool to study a multitude of reaction mechanisms leading to the strategical design of catalysts as highlighted in some interesting examples: Claisen rearrangement of chorismate in the gas phase, water, and enzyme environment; [51] cycloaddition of ethylene or acetylene to 1,3–dipoles; [81,82] $\beta$–hydride elimination in Au(II) and Au(III) complexes; [83] and Au(I) assisted [3,3]-sigmatropic rearrangement of allyl acetate. [84]
In my work, URVA analysis was successfully employed to elucidate reaction mechanisms of a diverse set of reactions studied utilizing a novel combined PES-ED approach (Chapter 3.1) and B/Al-based hydride facilitated hydrogen evolution from water (Chapter 4.3).
3.1. A Novel Approach to Study Bond Formation and Cleavage

3.1.1. Background

The results of this work are published in Ref. 1 (Appendix A), therefore here a short overview is given.

Monitoring the formation and cleavage of a chemical bond are important to unravel the mechanism of a chemical reaction. The evolution of a chemical reaction—starting from reactants, through transition state (TS), and all the way into products—can be represented by a reaction path, often by the intrinsic reaction coordinate (IRC) of Fukui. [85] The mainstream approaches aimed at capturing bond formation/cleavage processes along the reaction path can be formally classified into two categories: those which focus on features of the potential energy surface [86–89] (PES), and those which follow properties of electron density [90–94] (ED), $\rho(r)$. A summarization of some known methods can be found in Figure 3.1.

In 1964, Hohenberg and Kohn [3] established there is a one-to-one relationship between energy and electron density, although the exact functional form of this mapping is unknown. The underlying motivation in our work was to explore this connection during a bond formation/cleavage event using a combination of PES– and ED–based methods, thereby obtaining a qualitative picture of energy–electron density relation and more importantly, developing a novel concert method to study reaction mechanisms.
By employing URVA which is based on the curvature of the reaction path on the PES, hence providing powerful insights about changes in the PES, we could pinpoint bond formation/cleavage events along the reaction path. On the other hand, a description of chemical bonding could be obtained based on ED, within the framework of Bader’s quantum theory of atoms–in–molecule (QTAIM) analysis. [95, 96] The latter exploits topological properties of ED and associated Laplacian $\nabla^2 \rho(r)$ [97–101] to identify specific changes in the ED, that could be closely associated with formation/cleavage of a bond. Furthermore, our work demonstrated an efficient way of utilizing the information extracted by $\nabla^2 \rho(r)$ to visualize bond formation/cleavage processes; based on analyzing overlapping of valence shell charge concentration (VSCC). [102,103]

To showcase the connection between energy and electron density, we studied bond formation/cleavage mechanisms in a diverse set of 10 representative chemical reactions [80,83,104,105] ranging from hydrogen migration, [77] cycloaddition reactions [82] to gold(I) catalysis. In particular, we wanted to answer the following questions: First, do the significant changes in the features of PES and ED during a bond formation/cleavage event occur simultaneously along the reaction path? Second, how do we come up with an efficient way of navigating the reaction path in order to locate bond formation/cleavage processes?
Figure 3.1. A summary of some known methods which monitor the development of either PES or ED related properties along the reaction path to identify bond formation/cleavage events. Figure adopted from Ref. 1.

3.1.2. Results and Discussion

The systematic comparison between URVA and QTAIM approaches to study a bond formation/cleavage event led to a few important conclusions. A bond formation event characterized by a curvature peak is accompanied by the changes of the ED, which occur shortly before or at the corresponding curvature peak. The corollary is also true where the changes of the ED in a bond cleavage event take place shortly at or after the corresponding curvature peak. In a nutshell, in any case of bond formation/cleavage event, the significant ED changes occur in the vicinity of the curvature peaks which are related to significant changes in the features of the PES. Importantly, we could provide a rational basis for this relation based on a decomposition of energy density $H(r)$ along the reaction path. This parallel behavior nicely agrees with the Hohenberg–Kohn theorem. Furthermore, this implies if one were to scan the changes of the ED along the reaction path in the wake of a bond formation/cleavage event, the search could only be confined to the curvature peak region, thereby saving the cumbersome exercise of searching the entire reaction path.
3.1.3. Summary

In summary, we could vindicate energy and electron density linkage proposed in the Hohenberg–Kohn theorem by means of a combined PES–ED approach, offering valuable guidelines to identify bond formation/cleavage events along the reaction path.

This work is published in Physical Chemistry Chemical Physics (2019) 21:15007–15018 titled:
*A new way of studying chemical reactions: A hand–in–hand URVA and QTAIM approach* (see Appendix A).

3.2. PyVibMS for Visualizing Vibrations in Molecules and Solids

I was a second author in this software report, published in Ref. 106 (Appendix E). The program was written by the first– and co–authors and I was involved in the editing of the manuscript and running test calculations. A brief summarization of the work is as following.

Development of post–processing tools are important in visualizing vibrational modes calculated with different mainstream *ab initio* packages. Here, a PyMOL plugin called PyVibMS was developed which can be used to visualize vibrational modes for both molecular and solid systems calculated with widely used quantum chemical programs including Gaussian, Q-Chem, VASP and CRYSTAL. The PyVibMS is user–friendly and has a set of very useful functionalities. It is an open source program written with Python and therefore more flexible. This work demonstrates for the first time, the Konkoli–Cremer LVMs.

This work is published in Journal of Molecular Modeling (2020) 26:290 titled:
*PyVibMS: a PyMOL plugin for visualizing vibrations in molecules and solids* (see Appendix E).

PyVibMS is freely available at: https://github.com/smutao/PyVibMS
3.3. A Revised Generalized Subsystem Vibrational Analysis

I was a second author in this work, published in Ref. 107 (Appendix F). A brief summarization of the work is as following.

This work presents a revised formulation for the previously implemented generalized subsystem vibrational analysis [108] (GSVA). In the new formulation, the explicit definition and construction of a non-redundant set of internal coordinates for the subsystem is no longer required. Instead, an alternative formulation is proposed using massless Eckart conditions [109,110] to the subsystem fragment paired with a Gram–Schmidt orthogonalization to span the same internal vibration space indirectly. This circumvents the effort to construct a non-redundant parameter set for the subsystem. The revised GSVA is applicable to equilibrium and transition state structures and implemented in the open source package UniMoVib [107,111].

This work is published in Theoretical Chemistry Accounts (2021) 140:31 titled: *A revised formulation of the generalized subsystem vibrational analysis (GSVA)* (see Appendix F).
HYDROGEN BONDING, DIHYDROGEN BONDING & H···H INTERACTIONS

4.1. Hydrogen Bonds in Small Dimers

4.1.1. Background

The results of this work were published in Ref. 112 (Appendix B), therefore here a short overview is given.

Emamian and co-workers [113] evaluated the relationship between binding energy (BE) and electron density $\rho(r)$ at the bond critical point (BCP) for a diverse set of hydrogen-bonded dimers including 28 neutral and 14 charged complexes. The authors observed a significantly strong correlation between BE and $\rho(r)$ and advocated the use of this correlation for a quick assessment of BE when the latter is unattainable.

Even though this study deals with a set of carefully handpicked simple dimers at very accurate computational levels, so that the evaluation of BE is reliable and not hampered by other external factors, in our opinion it still overlooks some important facts and raises certain caveats; Does BE only capture the "pure" hydrogen bond strength? Are the systems explored here have no other additional (de)stabilization interactions beyond the hydrogen bond? How rational is it to present a simplified and unified picture of hydrogen bonding using BE and $\rho(r)$ as bond strength descriptors?

In our work, employing LVM stretching force constant $k_\alpha^n$ as a direct measure of hydrogen bond strength, we aimed to clarify above concerns and provide explanations to some of the more elusive aspects in their original work.
4.1.2. Results and Discussion

We calculated $k_{n}^{a}$ for 28 hydrogen bonds and analyzed the correlation between BE-$k_{n}^{a}$ and $k_{n}^{a}$-$\rho(r)$. We could observe a strong correlation between $k_{n}^{a}$ and $\rho(r)$ of the hydrogen bonds which is a consequence of the fact that both $k_{n}^{a}$ and $\rho(r)$ are local descriptors of the hydrogen bond. There was also a considerably high correlation between BE-$k_{n}^{a}$ which was marginally low than the correlation observed for BE-$\rho(r)$ in the original work by Emamian and co-workers.

This comparatively low correlation was the result of five significantly deviated systems as identified by a standard statistical method using deviation criterion of $\sigma$. Interestingly, four of these systems were flagged as significantly deviated cases also in the BE-$\rho(r)$ correlation. We could explain this by comprehending the underlying physical nature of these complexes as well as comparing the extent of physical interactions covered by BE and $k_{n}^{a}$ bond strength descriptors. We could associate these significantly deviated systems with two types of (de)stabilizing interactions within monomers involved in the hydrogen bond. They were classified into two groups based on the interaction type. Group 1 have the undesirable situation where the F–H monomer unit is destabilized by the repulsive interactions between F and H atoms, whereas Group 2 have the favorable situation of attractive interactions between the ionic forms of monomers which add extra stabilization. When measuring the bond strength, BE captures the effect of all these additional (de)stabilization interactions while $k_{n}^{a}$ strictly quantifies only the hydrogen bond strength. Thereby we could explain the discrepancy between these two bond strength descriptors and assess the nature of some systems that deviate from the simple hydrogen bond situation which was overlooked in the original work. Furthermore, as a caveat we pointed out it is not possible to obtain an absolutely simplified and unified picture of hydrogen bonding as was the attempt of the original work.
4.1.3. Summary

This work revealed the importance of accurately capturing hydrogen bond strengths and shed light on some serious considerations that ought to be taken into account when using different bond strength descriptors, like local stretching force constant and binding energy. Even within these trivial H-bonded systems a number of decisive factors come into play, influencing H-bond strengths.

This work is published in Journal of Computational Chemistry (2020) 42:516-521 titled: Comment on "Exploring nature and predicting strength of hydrogen bonds: a correlation analysis between atoms–in–molecules descriptors, binding energies, and energy components of symmetry–adapted perturbation theory" (see Appendix B).

4.2. Hydrogen Bonds in Ice Crystals

4.2.1. Background

The complete results of this study are presented in Ref. 114 (Appendix C), therefore here a short overview is given.

Ice has some intriguing properties and rich in polymorphism. The latter is induced by the tetrahedral coordination around each water molecule, where two hydrogen bonds (H-bonds) are donated and two are accepted. Currently, there are 19 experimentally resolved different forms of ice, [115–118]; some existing in order–disorder pairs (i.e., ices Ih–XI, III–IX, V–XIII, VI–XV, VII–VIII, and XII–XIV). Ice XIX is the most recently discovered ice polymorph [119–121] which occurs as a result of pressure–induced structural distortions in ice VI.

Since the H-bond network of ice is pivotal to determining various structural, physical, and chemical properties of these different ice forms, measuring accurate lattice energies [122,123] prompts to understand the cumulative strength of the H-bond network. However, quantifying individual H-bond strengths of ice structures will lead to a more detailed and minute picture of the H-bond network, but this has proven to be an exceedingly hard venture from an
experimental viewpoint. By reverting to theoretical means we can quantify intrinsic H-bond strength within ice structures thereby broadening our understanding of the nature of ice.

Thus, the major goal in our work was to characterize the H-bond network of different ice forms employing the recently extended form of LVM theory to periodic systems. We calculated for 10 different experimentally confirmed ice structures as indicated in Figure 4.1 (Ih, II, III, VI, VII, VIII, IX, XIII, XIV, XV) whose initial structures were obtained via Grimme’s ICE10 dataset. [124] In addition to providing a quantitative description of the H-bond network in different ice forms, we addressed following aspects/questions in this project: What structural parameters (H-bond length; OH-bond length; HOH-angle) can influence different H-bond strengths? How does different local topological elements, i.e., rings, that can be found within the H-bond network vary in their bond strengths? What new insights can be gained about water–ice phase diagram in connection with H-bond strengths for different ices existing at ambient to high pressures?

Figure 4.1. H-bond networks of ten ices investigated in this work. Primitive unit cells are indicated in black. For order–disorder pairs, their (dis)order counterparts are also given in black text.
4.2.2. Results and Discussion

For both covalent O–H bonds and H-bonds, we could see a strong correlation between H-bond length and $k_n^a$ for the 10 experimentally established ice forms. Also, the statistical distribution of H-bond strength as characterized by $k_n^a$ for each ice form was investigated. We could observe bond strengths vary in a narrow to broad range within different ice structures. These broad bond strength distributions could be explained in terms of local structural features, i.e. rings, in the global H-bond topology of each ice. The latter can be associated with different H-bond strengths, for example, while hexagonal rings in ice Ih can be characterized by one bond strength, octahedrons which act as primary structural elements in ice VI/XV have varying bond strengths, hence the disparity in the bond strength distributions.

![Figure 4.2](image_url)

Figure 4.2. Comparison between H-bond strength distribution; (a) uniform in ice Ih and (b) inhomogeneous in ice VI/XV. H-bonds with same/different strengths are indicated by same/different color(s).

Furthermore, we could rationalize different bond strengths observed for a particular ice based on the so called push–pull effect; [4, 50] a phenomena where a target H-bond can be strengthened by directional high polarization, when a water molecule accepts and donates H-bonds. In addition, considering the water–ice phase diagram we could see ices that occur at ambient pressure correspond to strong H-bonds in general, and those at high pressure
region have weak H-bonds in contrast.

We very recently received reviewers’ comments for our manuscript and we are in the process of incorporating some of the valuable suggestions into our work. In the new submission we will expand ICE10 set to include more recently discovered ice forms which are ultra–low density ices, ice XVI [125] along with ice XVII, [126] and another very newly discovered H-ordered pendant of ice VI, ice XIX, whose crystal structure was resolved via three independent experimental studies. [119–121] In addition, our current study considers only one form for H-disordered structures and it was brought into our attention that our analysis could be improved by addressing H-disorder better. As we found in the literature there are two well–known approaches for dealing with H-disordered ices in computer simulations; one is to construct large supercells [127–129] and the other is to construct a series of primitive unit cell structures representing different configurations. [130–132] Considering the computational cost involved with larger supercell calculations with more than 100 water molecules, particularly in the Hessian matrix evaluation step, as well as taking into account that our LModeA-nano program employs a primitive unit cell model for the computations, [67] we decided to proceed with the latter way of introducing H-disorder. There are many algorithms based on graph theory to generate all possible configurations inside a unit cell (primitive or supercell) obeying ice rules, two popular flavors being graph enumeration algorithm [133,134] and genice program. [135,136] In our revisions, we are utilizing both algorithms interchangeably to generate a reasonable number of configurations to represent each H-disorder ice, only considering primitive unit cell structures. These new results will be included in our next submission of the manuscript.

4.2.3. Summary

In summary, this work for the first time provides a quantification of individual H-bond strengths in ice crystals. The knowledge of H-bond strengths in known ices unfolded in this work can be utilized in novel ice structure prediction, as a method of screening most plausible structures out of thousands of computer generated different H-bond networks.
At the time of submitting this dissertation (July 28, 2021) the manuscript has been submitted to Journal of Chemical Theory and Computation and we are working on reviewers’ reports. The submitted draft is presented in Appendix C.

4.2.4. Some Unpublished Results

Our initial work involving characterizing hydrogen bond network of a set of known ices, persuaded us to explore other related branches as well. One such side venture was to take into account hydrogen disorder in ices, which was not covered explicitly in the scope of our main project.

Hydrogen ordering in ice is a crucial property that has largely contributed to discoveries of new ice forms. At ordinary conditions, an ice crystal is hydrogen-disordered which means while oxygen sub–lattice exhibits long–range order, no such order is retained for the H atoms. This leads to a number of hydrogen disorder–order (H-(dis)order) pairs of currently known ices. It has been observed that hexagonal ice phase Ih when doped with KOH, goes to its H-ordered phase, ice XI, [137–142] at a transition temperature of 72K. [137,138] This is evidence that unique configurations of H-ordered ice are energetically preferred, opposed to the Pauling’s early assertion that proton ordered ice structures are degenerate. [143] The H-ordered configurations of Ih (i.e. ice XI) has been characterized by several experimental [142] and quantum chemical simulation [144] methods. One such prominent theoretical study is by Hirsch and Ojamäe, [144] where they studied not one but several H-ordered structures of ice Ih, including ice XI, to evaluate how the H-bond network affects the lattice energies of the different ice polymorphs as well as to elucidate the structures of the latter. Starting from an orthorhombic unit cell with 8 water molecules, [145] using permutation of H-atoms and imposing symmetry, they derived 16 symmetry distinct configurations as H-ordered structures of Ih. They explained the difference in lattice energies within those 16 distinct configurations in terms of structural characteristics, in particular the different conformations of H-bonded water molecular pairs. It was identified that there are four different conformations [146] that can be exhibited by a water dimer pair within its H-bond network which were termed as
h-cis, h-trans, c-cis, and h-trans. [147] One of the important conclusions was that there is a correlation between lattice energy and the distribution of these H-bonded conformations. Also, they refuted the suggestion [148–152] that these different confirmations correspond to H-bond classes of different strengths based on H-bond distance.

![Image](image.png)

**Figure 4.3.** The statistical distribution of H-bond strengths in 16 isomers via a box–and–whisker diagram. [2] The blue lines connecting mean (dotted) and median (solid) H-bond strengths are there only to guide the eye.

Using periodic LVM to characterize the H-bond networks of the all 16 symmetry distinct configurations of Ih, i.e., isomers, we expected to investigate how H-bond strengths vary for these 16 different H-ordered configurations. We used structures of 16 isomers within a 8–molecule orthorombic unit cell optimized at CASTEP/PW91 level from the original work as initial guess and calculated at BLYP-D3(BJ)/6-31+G(d,2p) level in our work. Our preliminary results indicating the H-bond strength distribution of the 16 isomers, as given by LVM stretching force constant $k_n^a$, is shown in Figure 4.3. It can be seen that H-bond strengths of all 16 isomers vary in the somewhat broad range of 0.34–0.41 Å. According to
the lattice energies computed in the original work of Hirsch and Ojamäe, isomer with $Cmc_2$1
space group symmetry tends to be the most stable, in agreement with the experimental results. [153] In our analysis, structure no. 1 corresponds to $Cmc_2$1 structure and it could be seen that it has a relatively narrow H-bond strength distribution, with the lowest mean and median H-bond strength values. As all of these isomers have the same oxygen lattice and differ only in the orientation of H-bonds, it would be worthwhile to investigate how H-disorder can give rise to different H-bond strengths. As hinted by the former analysis of Hirsch and Ojamäe, the answer might lie in examining the correlations between different H-bond conformatons and H-bond strengths. We briefly checked for a linear correlation as similar to Ref. 144 between no. of selected H-bonded dimer pairs and average H-bond strengths, but having found none suggests that other important factors need to be considered, such as the magnitude of push-full effect for these different H-bonded dimer pairs.

Another aspect we were interested in was to analyze computer predicted ices, in particular to asses how their H-bond networks compare to those of known, thereby screening potential new ice candidates. For this purpose, 7 plausible ice structures predicted via Bernham-English modified basin-hopping algorithm [154] were considered. The optimization and frequency calculations for these ices were done with Crystal17 program at BLYP-D3/6-31+G(d,2p) level under ambient pressures. The periodic LVM calculations were carried out for a primitive cell model of optimized structures to obtain $k_n^a$ for H-bonds within the unit cell as well as across periodic boundaries. For both covalent O-H bonds and H-bonds, we could see a weak correlation (0.752) as indicated in Figure 4.4 in the case of 7 predicted ices, which was significantly lower than what we observed for the experimental ices (0.938) we investigated prior (see Chapter 4.2.) Also, statistical distribution of H-bond strength as characterized by $k_n^a$ for these ices was investigated as shown in Figure 4.5. A key to understanding the H-bond networks of these ices better, particularly their differences compared to experimental ices, lies in characterizing local structural elements associated with different H-bond network topologies. A similar kind of analysis as we followed for isolating local structural features in our previous work (see Ref. 112) would be of interest to pursue as a
future direction in this work.

Figure 4.4. Relationship between local stretching force constants $k_a^a$ and bond lengths $r$ for 7 predicted ices. A power function in the form $k_a^a = a \cdot r^b$ was used to fit data.
Figure 4.5. The statistical distribution of H-bond strengths in 7 predicted ices via a box-and-whisker diagram. [2] Median is indicated.
4.3. Dihydrogen Bonds in Catalysis

The results of this work were published in Ref. 155 (Appendix D), therefore here a short overview is given.

4.3.1. Background

Hydrogen has tremendous potential as a clean and renewable alternative to conventional fuels. [156–161] Chemical hydrogen storage is a widely sorted and much focused aspect in transition to a hydrogen economy. In this regard, lightweight hydrides are particularly attractive candidates for chemical hydrogen storage [162,163] due to their rich hydrogen content and being a convenient source for on–board hydrogen supply.

By using simple lightweight hydrides as boranes, alanes and their derivative as catalysts, hydrogen can be produced via a water splitting reaction. [164–167] The ability to form dihydrogen interactions of the type X–H\(^{\delta^+}\)···\(^{\delta^-}\)H–Y [168–171] with water promotes the catalytic activity, and for borane– and alane–amines it has been shown to reduce the reaction barrier. [172–175] Thermochemistry and kinetics of this process have been studied widely, an in–depth mechanistic analysis is required to fill in the missing gaps in knowledge about hydrogen release by those catalysts. Some major questions to be posed are, (a) What are the major features of the reaction mechanism along the way leading to dihydrogen bond formation? (b) What is/are driving force(s) behind this process? (c) How to account for the regeneration of the catalysts? (d) What insights can be gained to model efficient catalysts?
4.3.2. Results and Discussion

In our work, we performed a detailed mechanistic analysis and addressed the above persisting questions using URVA and LVM analysis as major tools. We studied the reaction mechanism between water and BH$_3$, AlH$_3$, B$_2$H$_6$, Al$_2$H$_6$, and BH$_3$–AlH$_3$ catalysts where starting reaction complexes are shown in Figure 4.7. For one representative reaction, the impact of the aqueous solution was also probed. Additionally, charge transfer along the reaction path was evaluated using Natural Bond Orbital (NBO) analysis [176,177] and Cremer–Pople puckering analysis [178–180] was used to comment on the conformational preferences of the intermediates.

As revealed by URVA, major chemical events during the reaction were O–H cleavage, H–H formation, B(Al)–H cleavage and O–B(Al) formation, whereas O–H cleavage preceded the reaction sequence in all cases. LVM analysis characterizing bond strengths of H–H dihydrogen bonds at the transition state of the reactions revealed that weak H–H bonds impeded the H$_2$ evolution. This criteria helped to eliminate these catalysts as suitable candidates for H$_2$ production at an early stage. On the other hand, O–B(Al) strength of the by-product probed via LVM analysis suggested some promising candidates for catalyst regeneration. The driving force behind these reactions is the charge transfer between the catalyst and water. Those which can act as charge donors during the reaction promoted H$_2$ release while the others hindered it leading to very high activation barriers. Two subsets of reactions were identified based on the correlation between activation energy and the change in the charge separation going from reactant to transition state ($\Delta q_{(R-TS)}$) for specific bonds. The activation energy for one subset could be controlled by tuning the charge separation of O–H bond while the other relied on that of B(Al)–H bond. Conformational changes of the 6–membered intermediates also contributed to activation barrier. New insights could be gleaned on how to best optimize borane and alane catalysts for efficient hydrogen evolution from water.
Figure 4.6. Schematic representation of all starting reaction complexes investigated in this work. Hydrogens participating in a dihydrogen bond later in the reaction mechanism are labeled.

4.3.3. Summary

Overall, this study offered valuable insights into elucidating reaction mechanism, elimination as well as modification of catalysts, and catalyst regeneration for hydrogen evolution reactions in the presence of B– and Al–based catalysts that are assisted via dihydrogen bonds.

This work is published in The Journal of Physical Chemistry A (2020) 124:8978-8993 titled:

*Modeling hydrogen release from water with borane and alane catalysts: a unified reaction valley approach* (see Appendix D).
4.4. H⋯H Interactions in Naphthols

4.4.1. Background

The complete results of this study are presented in Ref. 181 (Appendix G), therefore here a short overview is given.

The initiative for this project was taken by Dr. Jaeger and co-workers where they studied rotational spectra of cis- and trans-1-naphthol and analyzed the electron distribution using Bader’s QTAIM approach. In the case of cis-1-naphthol, they could observe a bond path and a bond critical point between the OH hydrogen atom and the H-atom of the neighboring ring, suggesting a H⋯H interaction. They speculated that this is an interaction that is traditionally termed as "steric hindrance" and posed the question whether these are attractive or repulsive interactions. Similar type of interactions have been reported for phenanthrene and other hydrocarbons, where within the QTAIM framework it was shown that the H-atoms from the terminal rings in close proximity (also referred to as bay region) were connected by a bond path with a bond critical point. Based on these evidence, Bader and co-workers [182] claimed that these interactions are stabilizing and termed them as hydrogen-hydrogen bonds (different from dihydrogen bonding) which has later been subjected to much criticism. [14, 16, 183] Grimme and co-workers [16] used vibrational spectroscopy to analyze the H⋯H interaction in the phenanthrene bay region. They concluded that opposed to Bader’s arguments that these are stabilizing interactions, they were in fact irrelevant for the stability of phenanthrene. Dieter Cremer and co-workers [14] used LVM analysis to study phenanthrene and also concluded that H⋯H interactions in the bay region do not play a significant role in determining the stability of phenanthrene.
4.4.2. Results and Discussion

In the present work, we investigated cis- and trans-1-naphthol with LVM analysis to disclose the nature of C–H⋯H–O interactions in cis-1-naphthol. For comparison, cis- and trans- conformers of 2-naphthol were also considered. In this naphthol compound series, cis-1-naphthol emerged to have the strongest O–H bond. A bonding interaction between H atoms should result in a subsequent weakening of O–H bond, and conversely a strong O–H bond implies steric repulsion. Our initial analysis based on bond strengths suggest a repulsive interaction and this should be reflected in the IR spectrum as a blue shift. The latter is only valid when OH normal mode is not coupled with the other modes, e.g., bending modes. To ensure this we carried out a CNM \(^1\) procedure for cis- and trans-1-naphthol, and OH normal stretching mode for both conformers were observed to have a 99.9% contribution from O-H local stretching mode, indicating an almost pure O–H stretching mode. So the blue shift in IR spectrum observed by our collaborators with regard to cis- and trans-conformers, supports the traditional notion of a steric repulsion.

\(^1\)Characterization of Normal Modes
Figure 4.7. CNM analysis of cis-1-naphthol. O–H local stretching mode is indicated with red arrows and its % Contribution to O–H normal mode is shown in yellow. Numbering of local modes are as in the scheme.

4.4.3. Summary

This work was important in clarifying the nature of C–H⋯H–O interactions that take place between H atoms of close contact in a confined space. Although, AIM analyze is in favor of a weak bonding interaction, it is necessary to also consider repulsive interactions which outweigh the latter. LVM analyze followed by CNM procedure confirmed the traditional notion of a steric repulsion between the two close contact H-atoms and these findings will help in probing other similar systems with π-electron delocalization.

At the time of submitting this dissertation (July 28, 2021) the manuscript is under review in Physical Chemistry Chemical Physics. The submitted draft is presented in Appendix G.
Chapter 5

HALOGEN BONDING

5.1. Halogen Bonds in Crystals

I was a second author in this work, published in Ref. 69 (Appendix H). A brief summarization of the work is as following.

The intrinsic strength of the X−I⋯OA-type (X= I or Cl; OA: carbonyl, ether, and N-oxide groups) halogen bonding in 32 model systems derived from 20 molecular crystals were investigated at rev-vdW-DF2 density functional [184] level using periodic LVM analysis. The LVM stretching force constant $k_a$ for I⋯O halogen between donor dihalogen X-I and various acceptor OA molecules showed significant variations (0.1-0.8 mdyn/Å). Also, $k_a$ and bond length for both donor X−I bonds and acceptor I⋯O halogen bonds showed strong correlations, thereby extending generalized badger rule to crystals for the first time. The halogen atom X plays an important role in determining the intrinsic strength of I⋯O halogen bond, as it can influence the electrostatic attraction between σ-hole on I and acceptor atom O. Another factor that contributes to the I⋯O bond strength is the nature of the OA acceptor molecule where different substituents can fine tune $n \rightarrow \sigma^*$ charge transfer ability. Additionally, if X is involved in a second halogen bond, it could weaken the primary halogen bond. Overall, this study demonstrated the use of periodic LVM analysis for quantifying the halogen bond strength within crystals, which could easily be applied to study any type of non-covalent interactions in other crystalline systems.

This work is published in Molecules (2020) 25:1589–1610 titled:

*In situ assessment of intrinsic strength of X−I⋯OA-type halogen bonds in molecular crystals with periodic local vibrational mode theory* (see Appendix H).
BIBLIOGRAPHY


[75] Elfi Kraka and Dieter Cremer. Dieter cremer’s contribution to the field of theoretical chemistry. *Int. J. Quantum Chem.*, 119:e25849, 2019. 8


Appendix A

Paper on URVA and QTAIM

S. Nanayakkara, and E. Kraka, A new way of studying chemical reactions: a hand–in–hand URVA and QTAIM approach,

A new way of studying chemical reactions: a hand-in-hand URVA and QTAIM approach†‡

Sadisha Nanayakkara and Elfi Kraka*

Bond formation and bond cleavage processes are central to a chemical reaction. They can be investigated by monitoring changes in the potential energy surface (PES) or changes in the electron density (ED) distribution \( \rho (r) \) taking place during the reaction. However, it is not yet clear how the corresponding changes in the PES and ED are related, although the connection between energy and density has been postulated in the famous Hohenberg–Kohn theorem. Our unified reaction valley approach (URVA) identifies the locations of bond formation/cleavage events along the reaction path via the reaction path curvature peaks and their decomposition into the internal coordinate components associated with the bond to be formed or cleaved. One can also investigate bond formation/cleavage events using the quantum theory of atoms-in-molecule (QTAIM) analysis by monitoring changes in the topological properties of \( \rho (r) \) and the associated Laplacian \( \nabla^2 \rho (r) \). By a systematic comparison of these two approaches for a series of ten representative chemical reactions ranging from hydrogen migration to cycloaddition reactions and gold(I) catalysis, we could for the first time unravel the PES–ED relationship. In the case of a bond formation, all changes in the ED occur shortly before or at the corresponding curvature peak, and in a bond cleavage, the ED changes occur at or shortly after the curvature peak. In any case, the ED changes always occurred in the vicinity of the curvature peak in accordance with the Hohenberg–Kohn theorem. Our findings provide a comprehensive view on bond formation/cleavage processes seen through the eyes of both the PES and ED and offer valuable guidelines on where to search for significant ED changes associated with bond formation or cleavage events.

1. Introduction

Bond formation and bond cleavage are central to a chemical reaction and this provokes one to pinpoint these chemically invaluable events along a reaction path connecting reactants and products. The general approach would be to follow this reaction path and quantitatively account for the bonding status of the reaction complex (i.e. the union of reacting molecules) on the grounds of monitoring some property that is sensitive to bond formation/cleavage events. There have been a number of studies along these lines using as a reaction path the intrinsic reaction coordinate (IRC) of Fukui.† Based on the choice of bond-sensitive property, these approaches can be classified into two main categories: those that employ features of the potential energy surface (PES), i.e. energy related properties, and those that use topological properties of the electron density (ED), as summarized in Fig. 1.

With regard to features of the PES, Komorowski and co-workers‡ introduced a novel perspective to study bond formation/cleavage using the so-called atomic fragility parameter \( a^X \), which is based on the trace of the Hessian matrix. Evaluation of \( a^X \) along the IRC generates a spectrum-like profile termed the reaction fragility spectrum where regions of \( a^X \) with \( a^X < 0 \) correspond to bond dissociation and \( a^X > 0 \) corresponds to bond...
formation. Toro-Labbé et al. characterized the reaction mechanism of 1,3-intramolecular hydrogen transfer reactions by identifying different regions along the IRC, utilizing energy and reaction force profiles as the operating tools; a recurring approach in several mechanistic studies.4,5

We use the unified reaction valley approach (URVA) developed by Kraka and Cremer,6–8 which is based on the reaction path Hamiltonian of Miller, Handy and Adams.9 URVA utilizes the scalar curvature $\kappa(s)$10 of the reaction path to identify bond formation/cleavage as a reaction proceeds. The curvature maxima $Kn$ indicate locations on the reaction path where the reaction coordinate drastically steers from a subset of degrees of freedom to another subset, suggesting a chemical event (conformational, rehybridization, bond formation, bond cleavage, etc.) (see Section 2 for a detailed description of URVA). URVA has been successfully applied in our research to organic reactions,11–15 homogenous catalysis16,17 and chemical reactions taking place in an enzyme.18

Complementary to the quantum chemical study of PES related properties, a description of bond cleavage/formation processes can also be based on the physically observable ED, $\rho(r)$. Hohenberg and Kohn19 showed that the energy of a non-degenerate electronic ground state of a molecule is a unique functional of the ED, implying that other chemical and physical properties can also be related to the ED. Thus, the analysis of ED redistributions along the reaction path provides a rational approach to study bond formation and cleavage in concert with the changes of the PES-related properties, which is the underlying motivation throughout our work. The pioneering work of Bader’s quantum theory of atoms-in-molecule (QTAIM) analysis exploits the assertions in Hohenberg–Kohn theorem (HKT) by providing a rigorous protocol to define chemical bonding in terms of topological properties of the ED.20–26 Another appealing method based on the ED is the electron localization function approach of Becke and Edgecombe.27 This was further developed by Silvi and Krokidis28 in their bonding evolution theory methodology for the quantitative study of chemical bonding along the reaction path. A substantial amount of work has been done within these frameworks to provide a wealth of information about reaction mechanisms, essential for identifying bond formation/cleavage processes, as highlighted in the recent works by Abjudeufack et al.29 on the 1,3-dipolar cycloaddition reaction mechanism and also in Zahedi and co-workers30 coupled analysis using bonding evolution theory and noncovalent interaction index. In the work of Zalazar and co-workers31 aimed at the study of bond formation/cleavage events in zeolite-catalyzed reactions, various topological properties along the IRC were used as direct indicators to disclose the bonding status of selected bonds.

Despite these numerous studies, there is still a missing link bridging the features of the PES with those of the ED to provide a deeper insight into bond formation/cleavage events. In this respect, it is worth mentioning the work of Chakraborty et al.,32 where they introduced a chemical binding indicator $b(r)$, combining the Berlin function33 and the derivative of the ED. The regions of space where $b(r) < 0$ indicate a bond cleavage event, and $b(r) > 0$ indicates a bond formation event. Another study spanning the scopes of both PES and ED is that of Sun and co-workers,34 involving the conventional energy scheme of stationary points and intermediates, and more interestingly the topological analysis of the ED along the reaction path emphasizing bond formation/cleavage structures. By monitoring gradient paths of the ED, they located the points on the reaction path where gradient paths appear/disappear, signifying formation/cleavage of respective bonds and points corresponding to the maximum of the Laplacian $\nabla^2 \rho(r)$, which they termed as a structure transition state.

The preceding studies describing bond formation/cleavage along the reaction path based on the features of the PES, ED or a composite of both, provoked us to systematically explore how the changes in the PES and ED pertinent to bond formation/cleavage processes are related along the reaction path. Simply put, are the significant changes in the features of the PES and ED, which govern the formation or rupture of a bond, analogous? If we monitor the formation/cleavage of a chemical bond along the reaction path, can we identify if the PES and ED changes occur simultaneously or concomitantly? If these changes occur at different points along the reaction path, what is the order of appearance? One final question to pose is, how can we use this new-found relation between features of the PES and ED to investigate bond formation/cleavage events along the reaction path in the most efficient way?

To answer these questions, we performed a comprehensive study linking features of the PES and ED for a series of chemical reactions covering a wide range of different bond formation/cleavage mechanisms, including hydrogen migration reactions, pericyclic reactions such as the Diels–Alder reaction, and non-catalyzed catalytic sigmatropic rearrangement reactions. While we employed the URVA toolbox for investigating PES features, we used the QTAIM methodology for analyzing ED redistributions along the reaction path. Whereas the curvature maxima explicitly reveal the bond formation/cleavage processes along the reaction path, it is a more challenging task to identify the specific changes in the ED that could be closely associated with formation/cleavage of a bond. Bader’s analysis based on topological properties of the ED at bond critical points (bcps) $\rho(\mathbf{r})$, provides a rational basis to identify formation or cleavage of a bond. The appearance of a bond path indicates the formation of a new bond and conversely, the disappearance of a bond path is associated with cleavage of an existing bond. However, an ED measure solely based on the bond path could be misleading and has to be complemented with other topological properties. While the gradient vector field, $\nabla \rho(\mathbf{r})$, can be used to further characterize the bond paths to some extent, more information of chemical interest can be extracted from the Laplacian $\nabla^2 \rho(\mathbf{r})$ as it displays the regions of local concentration and depletion of charge. Additionally, as pointed out by Bader and co-workers, $\nabla^2 \rho(\mathbf{r})$ yields a density shell structure corresponding to the quantum shell structure.35,36 Thus, the outermost region of charge concentration ($-\nabla^2 \rho(\mathbf{r})$) represents the valence shell, termed the valence shell charge concentration (VSCC). The VSCC has been employed in several works to study bond formation/cleavage events.37,38 For a bonded atom pair, the covalent character increases with approaching VSCC, and vice versa. Therefore, we also used the VSCC overlap begin/end points as flag points to monitor bond formation/cleavage processes. For the characterization of the covalent nature of a chemical bond, the Cremer and Kraka
criterion draws a sufficient condition based on the total electronic energy density $H(c)$ at bond critical point $c$, which states that $H(c) < 0$ is indicative of a covalent interaction, whereas $H(c) > 0$ is indicative of an electrostatic interaction. Thus, we monitored changes in the bond critical points, bond paths, $H(c)$, $\nabla^2 \rho(c)$ and VSCC in order to identify bond formation/cleavage events, which we defined as E points. These E points include (a) catastrophe points attributed to the ED topology: appearance/disappearance of a bond path, and coalescence of bond and ring critical points, and (b) the initial point of partial overlap between the VSCCs of the two atoms of interest and conversely, the terminating point of this partial overlap separating the VSCCs to their respective atomic basins.

The objectives of this study were the following: (i) to elucidate and comment on bond formation/cleavage processes along the reaction path with the combined use of URVA and QTAIM; (ii) to probe the topology of the Laplacian at curvature peak positions; (iii) to understand the relationship between maxima in the URVA curvature and E points, and moreover to compare their relative locations along the reaction path; (iv) to evaluate QTAIM properties at the bond critical points, at the stationary points, e.g. reactants, products, and transition states, and at the reaction path curvature maxima, curvature minima and selected E points.

2. Computational methods

The theoretical basis underlying URVA has been thoroughly described in several publications and in some review articles. Therefore, we will provide only a brief overview of the major features of URVA, focusing on those used in this work. URVA follows the movement of the reaction complex (RC, the union of reacting molecules) along the reaction path on the PES from reactants to products via the transition state spanning the entire reaction path region from entrance to exit channel. As a chemical reaction proceeds, the occurring electronic structure changes of the RC are registered by its vibrational modes. The change in the vibrations results in a change of their coupling with the translational motion along the reaction path, which can be monitored by the so-called curvature coupling coefficient $B_{ij}$ at each path point $s$. Large electronic structure changes, i.e. the formation/cleavage of a chemical bond, result in large $B_{ij}$ values. The $B_{ij}$ coefficients define the scalar reaction path curvature $k(s)$, which measures the curving of the reaction path. Thus, analysis of $k(s)$ along the reaction path in the form of a curvature profile provides a direct connection to identify the electronic structure changes of the RC during the reaction. The curvature profile is an ensemble of curvature maxima $K_n$ and minima $M_n$. The $K_n$ points indicate important chemical events (e.g. bond formation/cleavage, rehybridization, or charge transfer and polarization, etc.) while at the $M_n$ locations bordering a peak, the RC falls to a neutral state or one with lower activity. The $M_n$ can be associated with hidden intermediates, which means a conversion to a real intermediate under specific reaction conditions is possible, e.g. by adding a catalyst. We coined the term “reaction phase” to mean the region stretching from one curvature minimum to the next, enveloping a curvature maximum. In this way, each curvature profile is composed of a series of reaction phases, which can be used as the fingerprint of the reaction.

A key feature of URVA related to our current work is the decomposition of the curvature into curvature components unravelling which internal coordinates (bond length, bond angle, dihedral angle, etc.) contribute to a certain curvature peak. A curvature component can be supportive (positive $k(s)$ values) or resistive (negative $k(s)$ values) to the chemical event it is describing. In this way, the locations of bond formation/cleavage events can be uniquely identified.

We included in our current study the ten reactions (R1)–(R10) shown in Fig. 2, previously investigated with URVA. The reactions were chosen to cover a wide range of different bond formation/cleavage scenarios. We focused in particular on the bond length curvature components and their associated curvature peaks describing bond formation/cleavage events. For each of the reactions (R1)–(R10), the same model chemistry was used as in the previous URVA study: UMP2/aug-cc-pVDZ for (R1) and (R2), MP2/6-31G(d,p) for (R3) and (R4), B3LYP/Def2-TZVP for (R5), B3LYP/6-31G(d) for (R6) and (R8), B3LYP/6-311+G(d,p) for (R7), B3LYP/6-31+G(d,p) for (R9) and B3LYP/6-31+G(d,p)/SDD for (R10). A sufficiently large range of points around each curvature peak was analyzed to identify all important E points including the VSCC overlap begin/end points. The ED at these points was analyzed to determine the corresponding $\rho(c)$, $\nabla^2 \rho(c)$, $\varepsilon$ and $H(c)$ values. Apart from curvature peaks, we also investigated the ED at other points of interest along the reaction pathway with the QTAIM methodology, i.e. at the location of reactants (Rs), products (Ps) and transition states (Ts), as well as at the Mn. All quantum chemical calculations were carried out with the GAUSSIAN16 program package. The ED analysis was performed with the AIMALL program and URVA was performed with the program package COLOGNE2019.

3. Results and discussion

The results of this work are presented for 4 selected reactions (see Fig. 2), and each of the subsections is dedicated to one of the reactions (R1), (R2), (R4) and (R9). The rest of the reactions are discussed in the ESI. The QTAIM parameters evaluated at curvature peak and E points of importance for the bond
formation/cleavage processes and at the stationary points of $R_s$, $T_s$, and $P_s$ are tabulated in Tables 1–4. In Fig. 3–6, the scalar curvature $κ(s)$ and its decomposition into bond length contributions relevant for the bond formation/cleavage processes are

### Table 1

<table>
<thead>
<tr>
<th>Numbering</th>
<th>Bond</th>
<th>Character</th>
<th>$s$ (amu$^{1/2}$ Bohr)</th>
<th>$ρ(c)$ (e Bohr$^{-3}$)</th>
<th>$∇^2ρ(c)$ (e Bohr$^{-5}$)</th>
<th>$κ$ ($× 10^{-6}$)</th>
<th>$H(c)$ (Hartree Bohr$^{-3}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1–H2 + H3</td>
<td>H2H3</td>
<td>R</td>
<td>–2.6799</td>
<td>0.0031</td>
<td>0.0086</td>
<td>0.0000</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E</td>
<td>–0.2999</td>
<td>0.0820</td>
<td>–0.0110</td>
<td>0.4441</td>
<td>–0.0335</td>
</tr>
<tr>
<td></td>
<td></td>
<td>K</td>
<td>–0.2000</td>
<td>0.1010</td>
<td>–0.0654</td>
<td>0.2220</td>
<td>–0.0490</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TS</td>
<td>0.0000</td>
<td>0.1557</td>
<td>–0.3345</td>
<td>1.1102</td>
<td>–0.1128</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Numbering</th>
<th>Bond</th>
<th>Character</th>
<th>$s$ (amu$^{1/2}$ Bohr)</th>
<th>$ρ(c)$ (e Bohr$^{-3}$)</th>
<th>$∇^2ρ(c)$ (e Bohr$^{-5}$)</th>
<th>$κ$ ($× 10^{-6}$)</th>
<th>$H(c)$ (Hartree Bohr$^{-3}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1H2</td>
<td></td>
<td>R</td>
<td>–2.6799</td>
<td>0.2752</td>
<td>–1.4181</td>
<td>0.0000</td>
<td>–0.3610</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E</td>
<td>–0.2999</td>
<td>0.2400</td>
<td>–1.0262</td>
<td>0.2220</td>
<td>–0.2729</td>
</tr>
<tr>
<td></td>
<td></td>
<td>K</td>
<td>–0.2000</td>
<td>0.2183</td>
<td>–0.8177</td>
<td>0.2220</td>
<td>–0.2244</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TS</td>
<td>0.0000</td>
<td>0.1557</td>
<td>–0.3345</td>
<td>1.3323</td>
<td>–0.1128</td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th>Numbering</th>
<th>Bond</th>
<th>Character</th>
<th>$s$ (amu$^{1/2}$ Bohr)</th>
<th>$ρ(c)$ (e Bohr$^{-3}$)</th>
<th>$∇^2ρ(c)$ (e Bohr$^{-5}$)</th>
<th>$κ$ ($× 10^{-6}$)</th>
<th>$H(c)$ (Hartree Bohr$^{-3}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H3–C1 ≡ N2 → C1 ≡ N2–H3</td>
<td>C1H3</td>
<td>R</td>
<td>–3.3935</td>
<td>0.2916</td>
<td>–1.2087</td>
<td>0.0000</td>
<td>–0.3404</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TS</td>
<td>0.0000</td>
<td>0.2015</td>
<td>–0.3274</td>
<td>0.8648</td>
<td>–0.1741</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E1</td>
<td>0.0199</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M1</td>
<td>0.2999</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>K1</td>
<td>0.7598</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M2</td>
<td>1.1398</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P</td>
<td>4.2527</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

| N2H3      |      | R         | –3.3935                | —                     | —                     | —                | —                         |
|           |      | TS        | 0.0000                 | —                     | —                     | —                | —                         |
|           |      | E1        | 0.0199                 | 0.2011                | –0.3148               | 0.9183           | –0.1724                   |
|           |      | M1        | 0.2999                 | 0.2038                | –0.3577               | 1.1146           | –0.1721                   |
|           |      | K1        | 0.7598                 | 0.2386                | –0.6623               | 0.4024           | –0.2716                   |
|           |      | M2        | 1.1398                 | 0.2650                | –1.0479               | 0.1686           | –0.3543                   |
|           |      | P         | 4.2527                 | 0.3384                | –1.8678               | 0.0000           | –0.5190                   |


Table 4. Local topological properties of the electron density \( \rho(c) \) along the IRC calculated at the bond critical points \( c \) for the selected bonds C16O5, C1C16, and C10O4; values in bold are for curvature peak data. \( s \) reaction parameter, \( \rho(c) \) electron density, \( \nabla^2 \rho(c) \) Laplacian, \( H(c) \) total energy density, and \( \varepsilon \) ellipticity calculated at the B3LYP/6-31++G(d,p) level of theory.

<table>
<thead>
<tr>
<th>Numbering</th>
<th>Bond</th>
<th>Character</th>
<th>( \rho(c) ) (amu(^{1/2}) Bohr(^{-3}))</th>
<th>( \nabla^2 \rho(c) ) (amu(^{1/2}) Bohr(^{-5}))</th>
<th>( \varepsilon )</th>
<th>( H(c) ) (Hartree Bohr(^{-3}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>O4C10 R</td>
<td>(-16.5259)</td>
<td>(-0.2260)</td>
<td>(-0.3863)</td>
<td>(0.0117)</td>
<td>(-0.2897)</td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>(-16.1438)</td>
<td>(-0.2261)</td>
<td>(-0.3866)</td>
<td>(0.0115)</td>
<td>(-0.2897)</td>
<td></td>
</tr>
<tr>
<td>M2</td>
<td>(-13.5044)</td>
<td>(-0.2264)</td>
<td>(-0.3907)</td>
<td>(0.0135)</td>
<td>(-0.2891)</td>
<td></td>
</tr>
<tr>
<td>M3</td>
<td>(-9.1552)</td>
<td>(-0.2248)</td>
<td>(-0.3917)</td>
<td>(0.0167)</td>
<td>(-0.2845)</td>
<td></td>
</tr>
<tr>
<td>M4</td>
<td>(-5.5260)</td>
<td>(-0.2151)</td>
<td>(-0.3698)</td>
<td>(0.0209)</td>
<td>(-0.2554)</td>
<td></td>
</tr>
<tr>
<td>K3</td>
<td>(-3.9963)</td>
<td>(0.1902)</td>
<td>(-0.2633)</td>
<td>(0.0274)</td>
<td>(-0.1798)</td>
<td></td>
</tr>
<tr>
<td>E1</td>
<td>(-3.4264)</td>
<td>(0.1673)</td>
<td>(-0.1427)</td>
<td>(0.0328)</td>
<td>(-0.1243)</td>
<td></td>
</tr>
<tr>
<td>M5</td>
<td>(-2.1366)</td>
<td>(0.1108)</td>
<td>(0.0627)</td>
<td>(0.0524)</td>
<td>(-0.0403)</td>
<td></td>
</tr>
<tr>
<td>TS</td>
<td>(0.0000)</td>
<td>(0.0354)</td>
<td>(0.1218)</td>
<td>(0.1234)</td>
<td>(-0.0051)</td>
<td></td>
</tr>
<tr>
<td>K4</td>
<td>(0.1571)</td>
<td>(0.0580)</td>
<td>(0.1241)</td>
<td>(0.1161)</td>
<td>(-0.0060)</td>
<td></td>
</tr>
<tr>
<td>M6</td>
<td>(2.1367)</td>
<td>(0.0338)</td>
<td>(0.1129)</td>
<td>(0.3731)</td>
<td>(-0.0005)</td>
<td></td>
</tr>
<tr>
<td>E2</td>
<td>(3.3364)</td>
<td>(0.0281)</td>
<td>(0.1016)</td>
<td>(0.3064)</td>
<td>(-0.0014)</td>
<td></td>
</tr>
<tr>
<td>K5</td>
<td>(3.7863)</td>
<td>(0.0263)</td>
<td>(0.0960)</td>
<td>(0.3363)</td>
<td>(-0.0015)</td>
<td></td>
</tr>
<tr>
<td>M7</td>
<td>(5.0161)</td>
<td>(0.0220)</td>
<td>(0.0801)</td>
<td>(0.4463)</td>
<td>(-0.0017)</td>
<td></td>
</tr>
<tr>
<td>M8</td>
<td>(9.8151)</td>
<td>(0.0135)</td>
<td>(0.0470)</td>
<td>(1.6929)</td>
<td>(-0.0014)</td>
<td></td>
</tr>
<tr>
<td>M9</td>
<td>(17.9135)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>(19.3478)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>

04C10 R: O4C10 bond forming reaction.

C16O5 R: C16O5 bond breaking reaction.

C1C16 R: C1C16 bond forming reaction.

C10O4 R: C10O4 bond breaking reaction.

s reaction parameter, \( \rho(c) \) electron density, \( \nabla^2 \rho(c) \) Laplacian, \( H(c) \) total energy density, and \( \varepsilon \) ellipticity calculated at the B3LYP/6-31++G(d,p) level of theory.

shown as a function of the reaction parameter \( s \) as well as \( \nabla^2 \rho(c) \) contour plots with bond paths and bond critical points at Kn, Mn, and E and stationary points listed in Tables 1–4. In addition, Cartesian coordinates of all stationary points and reaction movies of the reactions (R1)–(R10) are included in the ESL.’

3.1. \( R1: H + H_2 \) reaction

This elementary reaction has often been used as a prototype for studying bimolecular reactions from many theoretical and experimental aspects.\(^{54}\) We used this reaction as a starting

point for our combined URVA and QTAIM analysis to deduce the link between the relative locations of the significant E and curvature points on the reaction path reflecting HH breakage and HH bond formation. Because of symmetry, both forward and reverse reactions are identical. As shown in Fig. 3A, the curvature profile features two peaks, one closely before the TS at \( s = -0.20 \) units and one at \( s = +0.20 \) units, being related to the equivalent reverse reaction. The first peak is dominated by the bond formation component H2H3 (numbering the RC as H1–H2–H3), which is supportive (positive sign), whereas the
bond cleavage component H1H2 is negative, indicative of resistance to bond breaking. At the TS, with equal H1–H2 and H3–H3 distances and a 3 electron-3 center (3e-3c) bonding situation, a curvature minimum is located, denoting the starting point of the reversion of the role of the H1H2 and H2H3 components. Clearly, bond formation precedes bond breakage, leading to a non-classical 3e-3c bonding situation between s = −0.20 and +0.20 units.

The ED starts to accumulate in the region between H2 and H3 at E (−0.30 s units, see Table 1), initiating the VSCC overlap, which is also reflected by the Laplacian revealing that the ED of the incoming H atom is polarized towards the ED distribution of H2. This E point occurs fairly close to the curvature peak K, which signifies the formation of the new bond. It is interesting to note that after the E point, the supportive dominance of H2H3 declines whereas the resistance of H1H2 bond breakage becomes support and the 3e-3c bond starts to get established, showing that for this simple reaction, PES and ED features are synchronized.

3.2. R2: CH₃ + H₂ reaction
In this hydrogenation reaction of the CH₃ radical, we can recognize two distinct curvature peaks, K1 at −0.08 s units and K2 at +0.60 s units (see Fig. 4A). The first curvature peak K1 is dominated by the H2H6 bond component supporting bond breaking. In a previous URVA study, using the electron difference density distribution map Δρ(r), it has been shown that when CH₃ and H₂ approach each other, the HH bond becomes polarized towards the negatively charged C atom, which in turn leads to a draining of the ED from the HH bond region. Examining the $ν_2^2ρ(c)$ topology at the initial VSCC overlap point E1 (−0.16 s units, see Fig. 4B) confirmed this observation, where it was seen that the ED distribution in the region between H2 and H6 is slightly perturbed and after this point starts to accumulate between the C1 and H2 atomic basins. This is also reflected by $H(c)$ and $ν_2^2ρ(c)$ of C1H2 (see Table 2), which have become negative at this point. As can be seen from the curvature component C1H2 (see Fig. 4A), the resisting contribution, which is at a maximum at E1, starts to recede gradually after E1 and converts into a supporting one at the TS, which is shortly after the K1 peak. In this way, the ED flow from the H2H6 bond region to the C1H2 region causes the latter to become less resistive to bond formation. Meanwhile, the negative $H(c)$ and $ν_2^2ρ(c)$ values for H2H6 (see Table 2) are becoming more positive along the reaction path via the K1 peak implying the loss of covalent character. The other peak K2 at +0.60 s units results from finalization of the C1H2 bond formation as reflected in its supporting contribution to the peak. Here, there is no charge concentration between H2H6 atomic basins, as exhibited by the Laplacian topology (see Fig. 4B) and positive $ν_2^2ρ(c)$ value (= 0.03, see Table 2). In contrast, C1H2 has negative $ν_2^2ρ(c)$ and $H(c)$ values, evidence of the well-established covalent nature of the C1H2 bond. The charge concentration around the H6 atom starts adapting its spherical character at E2 (+0.42 s units, see Fig. 4B), when the ED starts to recede in the H2H6 region, indicating the detachment of VSCC overlap. With regard to the H2H6 component, E2 resides in a plateau of the H2H6 curvature component (in green, see Fig. 4A) after which the supporting contribution of H2H6 rapidly declines and turn into a resisting one along the path. This reveals that when the ED starts to reorganize into the CH bond situation of CH₄ and the ED distribution around H6 becomes spherical, the H2H6 bond cleavage process has reached completion and the corresponding curvature component adjusts accordingly. In other words, when the ED stops flowing into the H2H6 interatomic region, the H2H6 curvature component no longer affects the curving of the reaction path. Compared with the H + H₂ reaction, the first half of the CH₄ + H₂ reaction exhibits a similar curvature pattern with the inclusion of an E point, but the role of the HH component is reversed. The former has a resisting contribution to the first curvature peak from the HH bond component to be broken while the latter shows a supporting contribution (diminishing after E1), which can be explained by the enhanced polarization in the presence of the CH₃ radical. The E1 point associated with the CH bond formation precedes the K2 peak and is located in the vicinity of the K1 peak. On the other hand, the E2 point associated with the HH bond cleavage follows after the K1 peak and it is located near the K2 peak. The ED...
changes before a peak add up to the bond formation and the reorganizations after a peak signaled completion of bond cleavage, herein providing further insight into the link between PES and ED features.

3.3. R4: isomerization of HCN

This simple gas phase reaction of hydrogen cyanide rearranging to hydrogen isocyanide has been extensively studied using several theoretical tools\(^3\)\(^{,}\)\(^5\)\(^9\)\(^{–}\)\(^6\)\(^1\) including the URVA method.\(^1\)\(^0\) Here, we intended to revisit this reaction with our combined analogy, thereby facilitating better understanding of its bond formation/cleavage processes along the IRC. Also, this reaction can be seen in connection with the HN\(_2\)\(^+\) reaction ((R3), see ESI\(^‡\)) as characterized by a similar curvature shape. In the curvature profile, the only curvature peak K\(_1\) at +0.76 s units decomposes into major supporting contributions from C1H\(_3\) and N2H\(_3\) bonds. At this peak point, a bond path is found for N2H\(_3\) having a negative H\((c)\). It was formerly proposed that this reaction proceeds via a cyclic, non-classical structure (partial 2e–3c bonding) located after the TS.\(^1\)\(^0\) However, the ED analysis revealed the presence of a bond switching structure where the C1H3 bond path disappears and the N2H3 bond path appears at +0.02 s units (at E1, see Fig. 5B). This is rather close to the TS in which the N2H3 bond path is still absent. Since at E1, the bond path has already been switched from the C to N atom, E1 is in fact a stable structure according to structural stability requirements.\(^2\)\(^2\)\(^,\)\(^3\)\(^6\) An infinitesimal shift in the direction of the TS will result in a bond catastrophe point where the H atom will be directly linked to the CN bcp yielding a catastrophe point of the conflict type.\(^2\)\(^2\)\(^,\)\(^3\)\(^6\) (For our discussion, investigating E1 is sufficient, which is close to the catastrophe point.) At E1 for N2H3, H\((c)\) is negative, suggesting a covalent interaction and \(\varepsilon\) is significantly high (= 0.92, see Table 3) resulting from the distorted T-shaped ED distribution at that point. The negative H\((c)\) progressively increases and at +0.76 s units, it marks the finalization of the N2H3 formation as evidenced by the corresponding curvature peak K1 (see Fig. 5A). Also, an interesting observation about the topology of Laplacian charge density at K1 is that it starts depleting between the C1H3 atomic basins at this point, eventually rearranging to a more or less spherical form around C1 and H3 atoms. By the comparison of contour plots of \(\nabla^2 \rho(r)\) for M1 and its counterpart M2 (see Fig. 5B), this topological change is more distinct. In comparison to the analogous HN\(_2\)\(^+\) system, it can be inferred that changing from the symmetric HN\(_2\)\(^+\) system to one that is asymmetric results in a shift of the E point while preserving the overall curvature shape. The same cannot be said about the overall topological changes in the ED, which
have now become more complex since H has two different bonding partners. The TS is no longer a π-complex but closely resembles the E1 catastrophe structure and also as in the previous case no 3-membered ring was found. Here, an overall important observation is that the significant changes in the PES and ED, in the simultaneous process of C1H3 bond cleavage and N2H3 bond formation, do not necessarily overlap on the reaction path but are accompanied by one another. The bond catastrophe in the ED at E1 occurs well in advance of the curvature peak at K1, the latter dictating the final cleavage of the existing bond and the establishing of the new bond.

3.4. R9: [3,3]-sigmatropic rearrangement of allyl acetate

This reaction has been studied in previous studies,62 including with URVA,16 in connection with the gold-catalyzed counterpart (reaction (R10), see ESI†) to understand the mechanistic differences, particularly the influence of Au(i)–NHC catalysts.63,64 As revealed by the curvature peaks, the rupture of the C10O4 bond takes place at K3 (−4.00 s units) and the formation of the new C16O5 bond occurs at K5 (+3.79 s units) (see Fig. 6A). From R through K3, the ED starts depleting between the C10O4 atomic basins (see Fig. 6B), as represented by the decreasing negative $\nabla^2 \rho(c)$ and $H(c)$ values at C10O4 (see Table 4). At E1 (−3.43 s units), located in close vicinity of the K3 peak, the overlapping VSCC is separated into their respective atomic basins, C10 and O4. As could be anticipated, after this point, $\nabla^2 \rho(c)$ becomes positive, marking the transition of the C10O4 shared-shell interaction to that of a closed-shell interaction. Subsequently, the covalent character of the C4H10 bond deteriorates as viewed in the decrease of negative $H(c)$ values. It is interesting to notice that at E1, the supporting contribution of the O4C10 curvature component has already begun to decrease as well as the small supporting contribution from O5C16. Shortly after the TS at K4, both O4C10 and C16O5 components resist further change and as discovered by URVA, this indicates the possibility to break up the reaction into two steps with the use of a catalyst,16 which is demonstrated in the reaction (R10). At E2 (+3.55 s units, see Fig. 6B), the formation of the C16O5 bond is initiated when VSCC overlap occurs leading to the ED accumulation between the C16O5 atomic basins and bond formation is represented by the K5 peak. Both O5C16 and O4C10 components are supporting at this point; yet the supporting contribution from the latter starts to decline shortly after E1. When the ED starts to flow into the O5C16 region, this component positively influences the curving of the reaction path. In contrast, the ED has substantially depleted from the O4C10 region at this point and the O4C10 component affects the curving of the reaction path to a lesser degree, or in other words, it suppresses further curving. In agreement with the...
former results, the O4C10 bond cleavage is followed by E1, and E2 precedes the O5C16 bond formation.

3.5. Decomposition of $H(c)$ along the reaction path

The energy density $H(r)$ is defined as: 39,65

$$H(r) = G(r) + V(r)$$  \hspace{1cm} (1)

and its connection to the Laplacian $\nabla^2 \rho(r)$ is given by the virial theorem: 20–22

$$\frac{1}{4} \nabla^2 \rho(r) = 2G(r) + V(r)$$  \hspace{1cm} (2)

where $G(r)$ is the kinetic energy density (always positive) and $V(r)$ (always negative) is the potential energy density. The negative potential energy $V(r)$ corresponds to a stabilizing accumulation of density whereas the positive kinetic energy density $G(r)$ corresponds to a depletion of electron density. 39

Therefore, the inspection of how $G(c)$ and $V(c)$ change along the reaction path and how these changes are connected with the changes observed in $\nabla^2 \rho(c)$, in particular in bond formation/cleavage regions, will lead to a deeper understanding of the PES–ED relationship. This is discussed in the following for (R2), as an example. Fig. 7 shows how $H(c)$, $G(c)$, $V(c)$ and $\nabla^2 \rho(c)$ change for the C1H2 bond (bond to be formed) along the reaction path and Fig. 8 shows the corresponding changes for the H2H6 bond (bond to be broken). For the C1H2 bond, we observed that $H(c)$ becomes negative close to curvature minimum M1 before the K1 peak, which implies that $V(r)$ becomes dominating from this point. $G(c)$ increases up to the curvature point K2; however, the increase is not strong enough to outweigh $H(c)$. $\nabla^2 \rho(c)$ starts out as positive, slightly increasing until the E1 point is reached, where it changes sign from positive to negative (see Fig. 7e), indicating that electrons are confined to the bonding region. The link to the reaction path curvature is given by the fact that at the E1 point, the C1H2 component of the reaction path curvature changes from negative (i.e. resisting) to positive (i.e. supporting) as a consequence of electrons being more trapped in the C1H2 bond region, eventually leading to bond formation.

In the H2H6 bond cleavage event (see Fig. 8), after M1, $V(c)$ decreases and in contrast $G(c)$ increases, reaching a maximum value around M2. After this point, kinetic energy density becomes dominating as shown by the sign transition of $\nabla^2 \rho(c)$ from negative to positive (E2 point, see Fig. 8e). The link to the reaction path curvature is given by the fact that after E2, the supporting contribution from the H2H6 component rapidly declines and as kinetic energy density has already reached a maximum in the H2H6 region, it implies electrons are less likely to stay in that region or in other words the H2H6 bond has been cleaved. In summary, this analysis

Fig. 6  R9: (A) Scalar curvature as a function of the reaction parameter $s$ (solid black line). The decomposition of scalar curvature into components is given in color. The borders of the reaction phases are indicated by vertical dashed lines at curvature points M1, M2, M3, etc. The TS at $s = 0$ amu\(^{1/2}\) Bohr is also indicated by a vertical dashed line. R and P mark the first and last curvature points, corresponding to reactant and product, respectively. K1, K2, etc. denote curvature peaks. E1, E2, etc. denote special points in electron density shown by blue vertical dashed lines. Values in grey, 1, 2, 3, etc., indicate reaction phases and the highlighted region in blue shows the chemical phases. (B) Laplacian contour plots of selected curvature points R, K3, E1, TS, E2, K5, and P, plotted in the planes C1–O4–C10 for R–E1 and O5–C1–C16 for TS–P. The dashed (pink) contours denote regions of charge concentration where $\nabla^2 \rho(r) < 0$, and the solid (blue) contours denote regions of charge depletion where $\nabla^2 \rho(r) > 0$. 

PCCP Paper
shows that there is a direct connection between the interplay of $V(c)$ and $G(c)$ and the reaction path curvature.

4. Conclusions

In this work, we unravelled the relationship between the reaction path curvature peaks associated with bond formation/cleavage events and changes in the topology of the ED, vindicating the Hohenberg–Kohn theorem. For all reactions (R1)–(R10) investigated in this work, the ED changes connected with a bond formation/cleavage process occurred in the vicinity of the corresponding curvature peak. As illustrated in this work, a combined URVA-QTAIM study provides a comprehensive view on bond formation/cleavage processes seen through the eyes of both the PES and ED. It also offers valuable guidelines on where to search for significant ED changes associated with a bond formation/cleavage event. We could arrive at the following conclusions.

(1) Scanning the neighborhood of a curvature peak associated with a bond formation/cleavage event with QTAIM exposed a manifold of changes in the ED. These changes are reflected in the topology of the Laplacian, as the beginning of VSCC overlap between respective atomic basins inducing the bond formation or conversely, the ceasing of the VSCC overlap, leading to a bond cleavage. Additionally, catastrophe points of the conflict and bifurcation type could be identified in the close vicinity of a curvature peak, where bond paths appear/disappear or bond and ring critical points coalesce. These significant changes in the ED along the reaction path connected with a bond formation/cleavage event always appeared close to a curvature peak. Thus, the location of the curvature peaks can serve as starting points for an ED analysis of bond formation/cleavage processes within the QTAIM framework. This replaces the cumbersome exercise of scanning the entire reaction path to monitor topological changes in order to determine the location of bond formation/cleavage, as has been the approach of some previous studies.

(2) The locations of the reaction path curvature peaks denoting bond formation or bond cleavage events are concomitant to notable changes in the topology of the ED. For bond formation events, significant ED changes occur before or at the curvature peak, and as a corollary, for bond cleavage events, the topological changes indicated by the disappearance of a bond path or ceasing of VSCC overlap between two atomic basins occur either at or shortly after the peak. We quantified the shift of the E points relative to the corresponding bond formation/cleavage
Curvature peak and the results are included in the ESI† (see Table S7). The largest deviation in the E point is observed for (R7) bond formation, which is shifted about 26% relative to the respective curvature peak. The other pericyclic reactions (R6) and (R8) show similar deviations except for (R9), which has a very slight deviation as low as 1–2% for both bond formation and cleavage events. The delay in bond formation in (R6), (R7) and (R8) compared to (R9) can be explained by the electronegativity difference of the two bonding partners. In the latter, the ED is exchanged between the C and O atoms, which is more efficient and leads to an early bond formation as opposed to C–C bond formation in (R6), (R7) and (R8). For all the other reactions, E point deviations up to 11% are observed, which suggest these E points are in close proximity to the corresponding curvature peaks. This information can be used as a useful guide in an ED analysis, particularly to narrow down the region scanned around the curvature peak. 

(3) Decomposition of $H(c)$ into kinetic and potential energy density contributions along the reaction path revealed that the relationship between PES and ED is mainly governed by the interplay of $V(c)$ and $G(c)$, where the latter reaches a maximum near Kn and Mn. As a consequence, significant changes in the ED occur close to a curvature peak.

(4) This combined URVA-QTAIM study leads us to further solidify the fact that the TS is mechanistically often not relevant, in accordance with previous URVA studies.12 We observed for all reactions studied in this work, with the exception of (R3) [see ESI†] where TS/E/K points coincide because of symmetry, that the curvature peaks associated with bond formation/cleavage events are located before or after the TS. We also could not find any specific changes in the ED at the TS locations.

(5) A topological analysis going beyond the stationary points on the PES (R, TS, and P) probing the nature of bonding at the reaction path curvature maxima and minima as well provides a comprehensive new understanding of bond formation/cleavage events.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank Yunwen Tao, Daniel Sethio and Velmurugan Gunasekaran for their valuable insights and suggestions, and Niraj Verma for his computational support with regard to the automation of the electron density calculations. This work was financially supported by the National Science Foundation [CHE 1464906]. We thank the SMU for providing computational resources.

References

Appendix B

Paper on hydrogen-bonded dimers

LETTER TO THE EDITOR

Comment on “Exploring nature and predicting strength of hydrogen bonds: A correlation analysis between atoms-in-molecules descriptors, binding energies, and energy components of symmetry-adapted perturbation theory”

Sadisha Nanayakkara | Yunwen Tao | Elfi Kraka

Department of Chemistry, Southern Methodist University, Dallas, Texas, USA

Correspondence
Elfi Kraka, Department of Chemistry, Southern Methodist University, 3215 Daniel Avenue, Dallas, TX 75275-0314.
Email: ekraka@gmail.com

Abstract
We evaluate the correlation between binding energy (BE) and electron density \( \rho(r) \) at the bond critical point for 28 neutral hydrogen bonds, recently reported by Emamian and co-workers (J. Comput. Chem., 2019, 40, 2868). As an efficient tool, we use local stretching force constant \( k_{HB}^{a} \) derived from the local vibrational mode theory of Konkoli and Cremer. We compare the physical nature of BE versus \( k_{HB}^{a} \), and provide an important explanation for cases with significant deviation in the BE–\( k_{HB}^{a} \) relation as well as in the BE–\( \rho(r) \) correlation. We also show that care has to be taken when different hydrogen bond strength measures are compared. The BE is a cumulative hydrogen bond strength measure while \( k_{HB}^{a} \) is a local measure of hydrogen bond strength covering different aspects of bonding. A simplified and unified description of hydrogen bonding is not always possible and needs an in-depth understanding of the systems involved.

KEYWORDS
binding energy, electron density, hydrogen bond, local hydrogen bond energy, local vibrational mode theory

In a recent article, Emamian and co-workers judiciously selected a diverse collection of 42 hydrogen-bonded dimers including 28 neutral and 14 charged complexes, in order to study the nature of hydrogen bonds with quantum chemistry. The geometry of these 42 dimer complexes was optimized at B3LYP-D3(BJ)/ma-TZVPP level and the binding energies (BEs) characterizing the strength of H-bonding were calculated at CCSD(T)/jul-cc-pVTZ level employing half of basis set superposition error correction. The authors utilized symmetry-adapted perturbation theory (SAPT) at the SAPT2 + (3)sMP2/aug-cc-pVTZ level to decompose the BEs into energy components and proposed a novel classification of hydrogen bonding into four categories: very weak, weak to medium, medium, and strong. Emamian and co-workers observed fairly strong linear correlation between the CCSD(T) BEs and the electron density \( \rho(r) \) at the (3,−1) hydrogen bond critical point (BCP) for both the neutral \((R^2 = 0.9732)\) and charged \((R^2 = 0.9644)\) complexes. They advocated the use of this correlation between BE and \( \rho(r) \) as a convenient way to predict in the molecular systems the BE of individual hydrogen bonds when these BEs are unattainable. We reassessed the correlation between BE and \( \rho(r) \) for those hydrogen bonds in the 28 neutral complexes by utilizing the local vibrational mode theory and obtained interesting insights which provide a different perspective complementing the work of Emamian and co-workers.

As shown in Figure 1(A), the correlation between BE and \( \rho(r) \) at the BCP of 28 neutral hydrogen bonds contains four cases with significant deviation identified with a deviation criterion of 1.5 \( \sigma \) \( \sigma \) is the standard deviation of the residuals after fitting). However, Emamian and co-workers did not provide any remarks on possible reasons.

In order to better understand and assess the above correlation, we employed the local vibrational mode theory originally developed...
by Konkoli and Cremer to calculate the local stretching force constant $k_{HB}$ of these 28 hydrogen bonds (collected in Table 1) based on the Hessian (second-order derivatives of energy) matrices of the equilibrium geometries optimized at the B3LYP-D3(BJ)/ma-TZVPP level. Konkoli and Cremer derived in 1998 the local vibrational modes associated with individual internal coordinates $q_n$ (e.g., H–H–C–O hydrogen bond length in D–H–C–O hydrogen bonding) directly from the normal vibrational modes (within the harmonic approximation) by solving the mass-decoupled Euler-Lagrange equations. The local vibrational mode associated with a bond stretching can be considered as the motion being obtained via an infinitesimal change in the bond length followed by the relaxation of all other atoms in the molecular system. Each local mode has its corresponding local mode frequency $\omega_n$ and local mode force constant $k_{HB}$ associated with a chemical bond for a hydrogen bond $k_{HB}$ independent of atomic masses and characterizes pure electronic structure effects. The underlying physical nature of local mode force constant associated with a chemical bond is the curvature of the Born–Oppenheimer potential energy surface (PES) in the direction of the bond (i.e., diatomic) stretching. Therefore, the local stretching force constant has been extensively used to quantify the

**FIGURE 1**  Correlation between BE and $\rho(r)$ at the BCP of the, (A) 28 neutral hydrogen bonds in the original dataset (reproduced from ref. [1]); (B) 22 neutral hydrogen bonds in the standard model or Group 0; (C) 35 neutral hydrogen bonds in the extended dataset. Group 1+ and Group 2+ dimers are represented by red triangles and squares, respectively. For a definition of Group 0, Group 1+, and Group 2+, see main text. $m_i$ ($i = 1–7$) denotes the new dimers. Coefficient of determination $R^2$, standard deviation of the residuals RMSE, and the sample size $n$ are indicated in blue for each linear regression line.

<table>
<thead>
<tr>
<th>Complex</th>
<th>$k_{HB}$</th>
<th>Complex</th>
<th>$k_{HB}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.018</td>
<td>15</td>
<td>0.095</td>
</tr>
<tr>
<td>2</td>
<td>0.025</td>
<td>16</td>
<td>0.098</td>
</tr>
<tr>
<td>3</td>
<td>0.026</td>
<td>17</td>
<td>0.099</td>
</tr>
<tr>
<td>4</td>
<td>0.028</td>
<td>18</td>
<td>0.099</td>
</tr>
<tr>
<td>5</td>
<td>0.028</td>
<td>19</td>
<td>0.127</td>
</tr>
<tr>
<td>6</td>
<td>0.027</td>
<td>20</td>
<td>0.177</td>
</tr>
<tr>
<td>7</td>
<td>0.042</td>
<td>21</td>
<td>0.167</td>
</tr>
<tr>
<td>8</td>
<td>0.054</td>
<td>22</td>
<td>0.176</td>
</tr>
<tr>
<td>9</td>
<td>0.049</td>
<td>23</td>
<td>0.164</td>
</tr>
<tr>
<td>10</td>
<td>0.066</td>
<td>24</td>
<td>0.195</td>
</tr>
<tr>
<td>11</td>
<td>0.082</td>
<td>25</td>
<td>0.251</td>
</tr>
<tr>
<td>12</td>
<td>0.094</td>
<td>26</td>
<td>0.194</td>
</tr>
<tr>
<td>13</td>
<td>0.088</td>
<td>27</td>
<td>0.315</td>
</tr>
<tr>
<td>14</td>
<td>0.092</td>
<td>28</td>
<td>0.372</td>
</tr>
</tbody>
</table>

Note: The ordering of these 28 complexes stays the same as the original work.
intrinsic strength of both covalent bonds\textsuperscript{14–16} and noncovalent interactions including hydrogen,\textsuperscript{17–20} halogen,\textsuperscript{21–24} chalcogen,\textsuperscript{25–27} pnictogen,\textsuperscript{28–30} and tetrel bonding.\textsuperscript{31} Concerning the detailed mathematical derivation of local vibrational mode theory, interested readers are referred to a recent review article.\textsuperscript{13}

We observed a rather strong correlation between the local stretching force constant $k_{HB}$ and the electron density $\rho(r)$ at the BCP of the, (A) 28 neutral hydrogen bonds in the original dataset; (B) 22 neutral hydrogen bonds in the standard model or Group 0; (C) 35 neutral hydrogen bonds in the extended dataset. Group 1+ and Group 2+ dimers are represented by red triangles and squares, respectively. For a definition of Group 0, Group 1+, and Group 2+, see main text. $\text{mi} (i = 1–7)$ denotes the new dimers. Coefficient of determination $R^2$, standard deviation of the residuals RMSE, and the sample size $n$ are indicated in blue for each linear regression line.

FIGURE 2 Correlation between local stretching force constant $k_{HB}$ and $\rho(r)$ at the BCP of the, (A) 28 neutral hydrogen bonds in the original dataset; (B) 22 neutral hydrogen bonds in the standard model or Group 0; (C) 35 neutral hydrogen bonds in the extended dataset. Group 1+ and Group 2+ dimers are represented by red triangles and squares, respectively. For a definition of Group 0, Group 1+, and Group 2+, see main text. $\text{mi} (i = 1–7)$ denotes the new dimers. Coefficient of determination $R^2$, standard deviation of the residuals RMSE, and the sample size $n$ are indicated in blue for each linear regression line.

The linear correlation between BE and local stretching force constant $k_{HB}$ (see Figure 3(A)) is in the same range as the correlation between BE and $\rho(r)$. On one hand, Emamian and co-workers carefully picked small-sized molecules for their set of 28 neutral hydrogen bonded dimers with marginal secondary interactions, i.e. additional interactions between the two monomers besides the target hydrogen bonding, so that in this particular case the BE predominantly reflects the pure hydrogen bonding strength. On the other hand, the local stretching force constant $k_{HB}$ with its physical nature as the curvature of PES in the direction of hydrogen bond stretching has been widely recognized as an intrinsic bond strength descriptor derived from vibrational spectroscopy.\textsuperscript{13} Therefore, it is of interest to check in more detail to what extent BE and $k_{HB}$ as two different hydrogen bond strength descriptors are compatible with each other.

We got particularly interested in the five cases with significant deviation in the BE versus $k_{HB}$ correlation identified with the deviation criterion of 1.5 $\sigma$ (see Figure 3(A)) because four of them overlap with the significantly deviated points in the BE versus $\rho(r)$ correlation. Before we delve into this, however, it is necessary to review the underlying physical nature of BEs when characterizing hydrogen bond strength, especially on what physical interactions are covered in
comparison with local stretching force constants. Emamian and co-workers claimed that the BE "is best to exhibit intrinsic binding strength of the monomers without considering the structural relaxation of the monomers." In other words, the BE includes all energy contributions collectively upon complexation of the monomers. As we sketch in Figure 4, the local stretching force constant $k_{HB}$ for H–A

bond covers the pure hydrogen bond interaction connecting the two monomers, it does not cover additional interactions within the monomers caused by hydrogen bonding. Based on this analysis, we can explain the cases with significant deviation in the correlation between BE and $k_{HB}$ shown in Figure 3(A).

The five points with significant deviation can be grouped into two parts including three above the regression line as Group 1+ and two below the regression line as Group 2+. Additionally, Group 0 refers to the standard model obtained after excluding Group 1+ and Group 2+. For a definition of Group 0, Group 1+, and Group 2+, see main text. $m_i (i = 1–7)$ denotes the new dimers. Coefficient of determination $R^2$, standard deviation of the residuals RMSE, and the sample size $n$ are indicated in blue for each linear regression line. Linear regression equations are shown for original and standard models.
electronegative fluorine atom. This provides the hydrogen atom in
$\text{F} \cdots \text{H}$ with excess electron density leading to repulsive interaction with
the electron density of the fluorine atom. Such a repulsion counting
as a destabilizing factor toward the total BE is not reflected in the
local stretching force constant $k_{\text{HB}}$ explaining why these four data
points are located above the regression line.

For Group 2 complexes, the donor monomers are hydrogen azide
$(\text{HN}_3)$ and hydrogen fluoride (HF), both of which are weak acids while
the acceptor monomer is ammonia (NH$_3$) as weak base. The association
of ammonia and weak acid molecules reminds us of the formation
of ammonium salt consisting of an ammonium cation (NH$_4^+$) and
an anion. Therefore, in the case of hydrogen bonding between ammonia
(NH$_3$) and a weak acid (D–H), the partially ionic character of ammonium
salt arising from the delocalization of ammonia lone pair into the
n$^\sigma$(D–H) antibonding orbital (covalent character of H–N) leads to an
electrostatic attraction between the anion (D$^-$) and the ammonium
cation (H$^+\cdot\cdot\cdot\text{NH}_3$).\textsuperscript{33} Such an ionic attraction counts as a stabilizing fac-
tor toward the total BE but is not reflected by the local stretching
force constant $k_{\text{HB}}$ of H–N explaining why these two cases with sig-
nificant deviation are found below the regression line.

Furthermore, to demonstrate that the observed deviations to the
BE and $k_{\text{HB}}$ model are not artifacts of the statistical method used here,
but instead have a sound physical basis, we calculated BE, $\rho(r)$ (at the
same levels of theory as in the ref. 1), and $k_{\text{HB}}$ for a set of newly
designed dimers (see Table 2). These were designed to have similar
dehydrogen bonding environment as in Group 1 or Group 2 dimers,
which deviate from the simple hydrogen bond situation due to the
presence of additional (de)stabilization. The new dimers contain as
their acceptor or donor at least one of the monomers incorporated in
the original 28 dimers. For a systematic comparison of how the pres-
ence of complexes that deviate from simple hydrogen bond situation
can affect the quality of the BE and $k_{\text{HB}}$ model, we did the following.
First, we removed all the systems belonging to and as characterized
by Group 1 and Group 2 dimer classes (i.e., Group 1: (20), (22), (25),
(27), and Group 2: (26), (28), respectively) and only considered the
rest of the 22 dimers (denoted as Group 0). These dimer systems
have the simple situation in their hydrogen bonding nature and the
the corresponding linear regression line is considered as a standard/
sumple model between BE and $k_{\text{HB}}^0$, which has $R^2 = 0.9871$ with
RMSE = 0.1937 (see Figure 3(B)). Then, with the inclusion of the
seven new dimers (denoted by Group 1+ and Group 2+) we did linear
regression for the extended data set with 35 dimers, and the
the corresponding linear regression line between BE and $k_{\text{HB}}$ has
$R^2 = 0.9413$ with RMSE = 1.0356 (see Figure 3(C)). In comparison to
the standard model between BE and $k_{\text{HB}}$, we could still see a fairly
high correlation for the extended data set, but the deviation (given by
RMSE) is much more pronounced now. This is a clear indication that
in the presence of complexes that diverge from the simple hydrogen
bond situation, the quality of the model is affected and as a caveat we
like to point out that careful consideration should be taken when pre-
senting a unified picture of hydrogen bonding under such circum-
cstances. Also, this analysis further confirmed, the identification of the
deviations defined as Group 1 and Group 2, is not merely based on the
points produced by the statistical method we used here, but rather
guided by our chemical intuition. A similar kind of evaluation was also
carried out for BE and $\rho(r)$, and $k_{\text{HB}}$ and $\rho(r)$ relationships as shown in
Figures 1(B),(C) and 2(B),(C), respectively.

The conclusions of our work can be summarized as follows.

- Those cases with significant deviation in the $\rho(r)$ versus BE correla-
tion are difficult to explain in reminiscence of the Hohenberg–
Kohn theorem which relates energy to electron density, but so far
the actual functional between the two is still not known. The local
stretching force constant $k_{\text{HB}}$ of hydrogen bond H–N character-
izing the hydrogen bond strength between the donor and acceptor
monomers is an effective tool for assessing cases with significant
development in the BE–$k_{\text{HB}}$ and BE–$\rho(r)$ correlations.

- For 28 neutral hydrogen-bonded dimer complexes investigated by
Emamian and co-workers, both BE and $k_{\text{HB}}$ are reasonable descri-
ators for quantifying the hydrogen bond strength and they provide
in general consistent ordering of bond strength for these 28 com-
plexes. We show that both BE and $k_{\text{HB}}$ provide similar measures of
hydrogen bond strength based on the judicious selection of donor/acceptor
monomers forming 28 simple dimer complexes made by
Emamian and co-workers. However, the local stretching force con-
stant $k_{\text{HB}}$ avoids the twist of borrowing additional local descriptors
to indirectly quantify the hydrogen bond strength (like the use of
$\rho(r)$ to predict BE). Therefore, the use of $k_{\text{HB}}$ is straightforward

\begin{table}[h]
\centering
\caption{Binding energy (CCSD(T)/\text{jul-cc-pVTZ} level with half counterpoise correction, indicated as BE) in kcal/Mol, $\rho(r)$ (B3LYP-D3(BJ)/ma-
tZVPP level) in a.u., local stretching force constant $k_{\text{HB}}$ in modyn/Å for the newly designed dimers.}
\begin{tabular}{|c|c|c|c|c|}
\hline
Group & Complex & Structure & BE & $\rho(r)$ & $k_{\text{HB}}$ \\
\hline
\hline
Group 1+ & m1 & FH···NCH & --7.40 & 0.0399 & 0.251 \\
 & m2 & FH···OC & --1.54 & 0.0160 & 0.083 \\
 & m3 & FH···FCCH & --1.01 & 0.0133 & 0.053 \\
Group 2+ & m4 & NH$_3$···N(Me)$_3$ & --9.47 & 0.0438 & 0.231 \\
 & m5 & FH···N(Me)$_3$ & --16.38 & 0.0763 & 0.406 \\
 & m6 & HCOOH···NH$_3$ & --12.02 & 0.0508 & 0.285 \\
 & m7 & HCOOH$^+$···NH$_3$ & --14.08 & 0.0557 & 0.309 \\
\hline
\end{tabular}
\end{table}

\textsuperscript{*}Peroxy acid.
when measuring the hydrogen bond strength and generally applicable also to complex systems like water clusters or proteins.

- Our work provides a different point of view into the nature of hydrogen bonding by emphasizing the interactions within the donor/acceptor monomer which are not reflected by local electronic structure descriptors of $H \cdots A$ such as $\rho (r)$ and $k_{\text{HB}}^\rho$ but contribute marginally to the total BE of the 28 neutral hydrogen-bonded complexes. Therefore, care has to be taken when different hydrogen bond strength measures are compared. The BE is a cumulative hydrogen bond strength measure while $k_{\text{HB}}^\rho$ is a local measure of hydrogen bond strength.

- As a caveat we want to highlight, any attempt to present a simplified and unified description of hydrogen bonding as done in the work of Emamian and co-workers via their BE and $\rho (r)$ model, can be clouded by the fact that many systems tend to diverge from the simple hydrogen bond situation due to other factors that come into play. Thus, careful consideration should be taken under such situations.

ACKNOWLEDGMENTS
This work was financially supported by National Science Foundation Grants CHE 1464906. The authors thank SMU for providing supercomputing resources. Yunwen Tao thanks Yue Qiu for helpful discussions.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID
Elfi Kraka [https://orcid.org/0000-0002-9658-5626]

REFERENCES

Appendix C

Paper on hydrogen bonds in ice

S. Nanayakkara, Yunwen Tao, and E. Kraka, Capturing individual hydrogen bond strengths in ices via periodic local vibrational mode theory: beyond the lattice energy picture, Under review in Journal of Chemical Theory and Computation
Capturing Individual Hydrogen Bond Strengths in Ices via Periodic Local Vibrational Mode Theory: Beyond the Lattice Energy Picture

Sadisha Nanayakkara,†‡ Yunwen Tao,†‡ and Elfi Kraka*†

†Computational and Theoretical Chemistry Group (CATCO), Department of Chemistry, Southern Methodist University, 3215 Daniel Avenue, Dallas, Texas 75275-0314, United States

‡These authors contributed equally to this work

E-mail: elfikraka@smu.edu

Abstract

Local stretching force constants derived from periodic local vibrational modes at the BLYP-D3(BJ) density functional level have been employed to quantify the intrinsic hydrogen bond strength of ten ice crystals that are stable under ambient to elevated pressures. Based on this characterization, relationships between local stretching force constants and structural parameters like hydrogen bond length and angle were identified. Moreover, different bond strength distributions, from uniform to inhomogeneous, were observed for the ten ices and could be explained in relation to different local structural elements within ices, i.e. rings, that are constituted of different hydrogen bond types. In addition, criteria for the classification of hydrogen bonds as strong, intermediate, and weak were introduced. The latter was used to explore a different dimension
of the water-ice phase diagram. These findings will provide important guidelines for assessing the credibility of new ice structures.

1 Introduction

Ice, the condensed state of water, is a solid presence on Earth’s surface, its interiors and moreover, can be found across the entire solar system\textsuperscript{1–4} and beyond; within comets, moons, distant planets, rings of giant planets, and even in interstellar medium\textsuperscript{5–7}. The most frequently observed form of ice, or ice polymorph, at ambient pressure on Earth is hexagonal ice, $I_h$.

Hydrogen bonding is pivotal to determining the structure of H$_2$O ices and the prospect of each water molecule participating in four hydrogen bonds, (H-bond)s, has led to the rich variety of ice polymorphs.\textsuperscript{8–13} The local tetrahedral coordination around each water molecule, also called the Walrafen pentamer,\textsuperscript{14} results in six possible orientations that can be adopted by the central and neighboring water molecules. The sole constraints that govern the arrangement of water molecules are the Bernal-Fowler ice rules\textsuperscript{15} and consequently, a large number of allowed configurations exist for one ice phase near its true ground state. This is known as hydrogen-disorder, where the oxygen sub-lattice preserves an order while the respective H-positions around the oxygen atoms are occupied arbitrarily, i.e., H-sites have equal occupancy of 0.5. Upon cooling, the H-disordered ices undergo a phase transition into their H-ordered counterparts, giving rise to a hydrogen order-disorder ice pair since both correspond to the same H-bond network topology.

Experimental and theoretical studies, that cover the manifold aspects and implications of H-bond network in ices are prolific, ranging from spectroscopical analyses\textsuperscript{16–20} to thermodynamics of proton ordering.\textsuperscript{21–23} These studies are concentrated on the actual cooperative arrangement of the water molecules\textsuperscript{24–26} and its influence on macroscopic properties of ice like structure, density, electrical, thermal, and lattice dynamics, rather than on hydrogen
bonding, an important microscopic property.

Noteworthy in this connection is the use of electron localization function (ELF) scheme by Silvi et al.\textsuperscript{27} to characterize the H-bond strength of H-bonded crystals, including ice VIII. Based on ELF profiles along the O–H···O bond line, the presence of a weak H-bond was suggested in the case of ice VIII, which was determined to be weaker than that of a water dimer and a pentamer. Moreover, on the grounds of a linear relationship between ELF values at the H-bond interaction point and pressure, they proposed ice X, which occur over 60 GPa to be constituted of very strong H-bonds. In this context, the two force constant model by Ross et al.\textsuperscript{28} is also of interest which was used to determine two different H-bond strengths in Ih as well as to interpret inelastic neutron scattering (INS) spectra of the latter, and a similar kind of analysis was applied in succession to ices Ic, XIII, XIV, XVI, and XVII\textsuperscript{29–33} to describe their H-bond strengths.

Alternatively, a delineation of the H-bonds can also be presented based on the bond strength descriptors obtained from local vibrational mode (LVM) theory introduced by Konkoli and Cremer.\textsuperscript{34–38} They derived local vibrational modes from normal vibrational modes, which generally cannot be used to directly assess the bond strength of a chemical bond/interaction in polyatomic molecules. This shortcoming stems from delocalization\textsuperscript{39,40} of normal vibrational modes and this was addressed in the process of deriving local vibrational modes.\textsuperscript{34} The reader is encouraged to refer to a comprehensive review by Kraka et al.\textsuperscript{41} on the origin and applications of LVM theory. In essence, LVM analysis has advanced as a powerful analytical tool, extensively applied to a broad range of chemical systems from simple molecular systems to proteins\textsuperscript{42} accounting for both covalent bonds\textsuperscript{43–55} and non-covalent interactions\textsuperscript{54,56–69} including H-bonds\textsuperscript{70–78} Recently, a whole new scope of chemical systems were unlocked with the extension of LVM theory to periodic systems,\textsuperscript{79} one- through three-dimensional, which has enabled to probe the nature of crystal bonding. A local vibrational mode in a periodic system is defined as a vibration driven by a specific internal coordinate in all primitive cells while relaxing all other parts of the periodic system. This approach accommodates a head-
to-head comparison of the intrinsic bond strength in periodic and molecular systems. Also, it does not suffer from certain restrictions imposed by periodicity; e.g. the lattice structure prevents the calculation of bond dissociation energies, or the reduced functionality of some standard analytical tools; e.g. periodic NBO\textsuperscript{80} does not provide second order perturbation energies. Due to these superiorities, periodic LVM analysis has already been used as an operating tool to quantify intrinsic bond strengths in materials/crystals; its debut application was focused on analyzing a special type of halogen bonding in crystals.\textsuperscript{81}

In this work, we used periodic local vibrational mode theory for a deeper understanding of H-bonds in ice crystals which entails a quantification of intrinsic H-bond strengths. The following objectives were pursued in this work.

- To quantify the intrinsic H-bond strength in ice crystals at an individual H-bond scale and discover correlations with geometrical parameters like H-bond length and angle.
- To clarify and comment on the conventional view whether there is a homogenous bond strength distribution within ice structures due to the fact that it is a three-dimensional assembly of the same tetracoordinated local H-bond environments.
- To identify and characterize distinct local structural features, i.e. rings, and different H-bond environments incorporated in those elements, responsible for disparate bond strength distributions in ices.
- To compare different H-bond strengths in different ices.
- To explore the water-ice phase diagram in connection with variation of H-bond strengths from ambient to high pressures.

The paper is structured as following: First, a description of Computational Details (Section 2) is provided. In the Results and Discussion (Section 3), key focus is on determining intrinsic H-bond strength in ices, followed by an analysis of its relation to various local struc-
tural features of ice within the global topology of H-bond frameworks. Finally, highlights of this work are emphasized in Conclusions (Section 4), along with a future outlook.

2 Computational Details

The initial structures were obtained from the ICE10\textsuperscript{82} benchmark set which includes four H-ordered and six H-disordered systems, all obeying Bernal-Fowler ice rules. The calculations were carried out with the CRYSTAL17\textsuperscript{83,84} program using the BLYP\textsuperscript{85,86} functional with D3 London dispersion correction\textsuperscript{87} applied in the Becke-Johnson damping variant\textsuperscript{88} (BLYP-D3(BJ)), utilizing Pople’s 6-31+G(d,2p) basis set.\textsuperscript{89-91} A primitive cell model was used for the calculations as required by the periodic LVM theory.\textsuperscript{79} For the sampling of the Brillouin zone, a 12×12×12 \textbf{k}-point mesh was used, generated via the Monkhorst-Pack scheme.\textsuperscript{92} A pruned XXLGRID (99,1454) integration grid was employed, and two-electron integrals were computed with an accuracy of $10^{-10}$-$10^{-12}$. Hessian matrices were obtained numerically, according to a central-difference formula\textsuperscript{93} with a step-size of 0.001 Å. All Hessian matrices were evaluated at the Γ point ($q = 0$).\textsuperscript{94} The reference systems, water monomer (H\textsubscript{2}O), dimer ((H\textsubscript{2}O)\textsubscript{2}) and hexamer ring ((H\textsubscript{2}O)\textsubscript{6}) were calculated using the Gaussian 16 package\textsuperscript{95} with an UltraFine (99,590) integration grid, and a model water 2D layer adapted from Ref.\textsuperscript{79} was calculated using CRYSTAL17, where all reference systems were evaluated at the BLYP-D3(BJ)/6-31+G(d,2p) level of theory.

For the statistical analysis presented in Section 3.2.2, a box-and-whisker diagram\textsuperscript{96} was used where the distribution of data is represented by a box with two whiskers. The horizontal lines correspond to the minimum and maximum. The bottom and top of the box is defined by first and third quartile of data, Q1 and Q3, respectively, while Q2 indicates the location of the median. The vertical length of the box correspond to the interquartile range $QR = Q3 - Q1$, where lower and upper whisker lengths are defined by $Q1 - 1.5\times QR$ and $Q3 + 1.5\times QR$, respectively. Data points lower than $Q1 - q\times QR$ and higher than $Q3 + q\times QR$ are
considered as outliers, where $q = 1.5$ (black dots) and $q = 3.0$ (open dots) define mild and extreme outliers, respectively.

For the local mode analysis along with local stretching force constants calculation, an in-house periodic version of the LModeA program package\textsuperscript{97} was used. The primitive cell and other graphical representations of the periodic systems were generated with the VESTA\textsuperscript{3}\textsuperscript{98} package.

3 Results and discussion

3.1 Equilibrium Volumes and Lattice Energies

Table 1 reports equilibrium volumes and lattice energies ($E_{\text{lat}}$, at 0K) calculated for the ICE10 set at the BLYP-D3(BJ)/6-31+G(d,2p) level. The back-corrected volumes suggested by Brandenburg \textit{et al.}\textsuperscript{82} are used as "experimental" reference values and accordingly, the calculated volumes give mean relative deviation (MRD) of -1.8\% and mean absolute relative deviation (MARD) of 1.9\% with respect to the references. We also compared our calculated lattice energies with the experimental lattice energies obtained by Whalley\textsuperscript{99} for seven ice systems via extrapolating measured values to 0 K, excluding zero-point vibrational energies. We could see mean deviation (MD) and mean absolute deviation (MAD) of 2.3 kcal/mol for the calculated lattice energies, which could be improved using other relatively costly density functionals as shown by Brandenburg \textit{et al.}\textsuperscript{82} in their benchmark study. However, an accurate description of H-bond strength mostly relies on the ability to produce reasonable geometries, and in this connection the selected level of theory performs satisfactorily at a relatively low computational cost.

3.2 Intrinsic Strength of H-Bonds in Ices

In the following, H-bond strengths as given by local stretching force constants $k_a$ and the relationship to geometrical parameters are probed for all the ten ices, followed by a statistical
Table 1: Calculated equilibrium primitive unit cell volumes (Å$^3$/H$_2$O) and lattice energies (kcal/mol·H$_2$O) for the 10 ices calculated at the BLYP-D3(BJ) level. Volume and energy values indicated within brackets correspond to the experimental reference values taken from the cited literature.

<table>
<thead>
<tr>
<th>Polymorph</th>
<th>#H$_2$O$^a$</th>
<th>Volume (Ref. 82)</th>
<th>E$_{lat}$ (Ref. 99$^b$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ih</td>
<td>12</td>
<td>30.14 (30.91)</td>
<td>-16.97 (-14.07)</td>
</tr>
<tr>
<td>II</td>
<td>12</td>
<td>23.68 (23.97)</td>
<td>-16.28 (-14.05)</td>
</tr>
<tr>
<td>III</td>
<td>12</td>
<td>24.45 (25.29)</td>
<td>-16.47 (-13.85)</td>
</tr>
<tr>
<td>VI</td>
<td>10</td>
<td>21.26 (21.98)</td>
<td>-15.87 (-13.68)</td>
</tr>
<tr>
<td>VII</td>
<td>16</td>
<td>19.18 (19.28)</td>
<td>-14.82 (-13.07)</td>
</tr>
<tr>
<td>VIII</td>
<td>8</td>
<td>19.19 (19.06)</td>
<td>-14.82 (-13.31)</td>
</tr>
<tr>
<td>IX</td>
<td>12</td>
<td>24.48 (24.97)</td>
<td>-16.61 (-13.97)</td>
</tr>
<tr>
<td>XIII</td>
<td>28</td>
<td>22.55 (23.03)</td>
<td>-16.27</td>
</tr>
<tr>
<td>XIV</td>
<td>12</td>
<td>21.78 (22.26)</td>
<td>-16.11</td>
</tr>
<tr>
<td>XV</td>
<td>10</td>
<td>21.38 (21.58)</td>
<td>-15.80</td>
</tr>
</tbody>
</table>

$^a$ No. of water molecules per primitive unit cell.  
$^b$ Values are available for seven ice crystal systems.

analysis of H-bond strengths and a definition of classification criteria for the same.

3.2.1 Variation of H-bond Strengths

The relationship between local stretching force constants $k_n$ and bond lengths $r$ for H-bonds and covalent OH bonds are shown in Figures 1 and 2, respectively. A strong correlation is observed between $k_n$ and bond length for OH covalent bonds ($R^2 = 0.956$) whereas that of H-bonds tend to be weaker in comparison ($R^2 = 0.938$). The bond strengths of both H-bond and OH covalent bonds in ices are revealed to vary in a broad range, 0.173-0.411 and 4.943-6.199 mdyn/Å, respectively. While H-bonds in all the ice polymorphs exhibit higher bond strengths than H-bonds in the water dimer (0.170 mdyn/Å), only few ices including ice Ih have H-bond bond strengths surpassing those of a hexamer water ring (0.340 mdyn/Å), enforced by the push-pull effect.$^{75}$ On the other hand, in comparison to a model 2D water layer (0.879 mdyn/Å) with highly directive H-bonds, $^{79}$ all the ice structures having no such unique orientation of water molecules, are bound to have weaker H-bonds.

We also investigated the relationship between the directionality and the H-bond strength, or in other words whether more linear intermolecular H-bonds in ice imply enhanced bond
Figure 1: Relationship between local stretching force constants $k_a^n$ and bond lengths $r$ for H-bonds. Circles and squares distinguish between disordered and ordered ices respectively, and different colors represent H-bonds in each ice polymorph (refer to inset); within primitive unit cell and across its boundaries. An arbitrary function in the form of $k_a^n = a \cdot c^b \cdot r^f + d \cdot r^{-g} + m \cdot \log(r)$ was used to fit all 264 data points with $R^2 = 0.938$. Marked along the second $y$-axis by stars are the $k_a^n$ values for the references, water dimer ($\text{(H}_2\text{O)}_2$) and hexamer ring ($\text{(H}_2\text{O)}_6$)). Dashed horizontal lines correspond to the strength classification in Section 3.2.3.

Figure 2: Relationship between local stretching force constants $k_a^n$ and bond lengths $r$ for OH covalent bonds. Circles and squares distinguish between disordered and ordered ices respectively, and different colors represent H-bonds in each ice polymorph (refer to inset); within primitive unit cell and across its boundaries. A power function in the form of $k_a^n = a \cdot r^b$ was used to fit all 264 data points with $R^2 = 0.956$.

strengths which is a common phenomena in proteins and other H-bonded crystals. As deduced from Figure 3, there is a weak correlation between the intermolecular OHO bond angle and $k_a^n$ as analyzed via a 95% prediction interval which reveals a tendency to show
Figure 3: Relationship between local stretching force constants $k_a$ and intermolecular OHO angles for H-bonds. The area shaded in grey bounded by dashed lines corresponds to the 95% prediction interval. Circles and squares distinguish between disordered and ordered ices respectively, and different colors represent OHO angles in each ice polymorph (refer to inset); within primitive unit cell and across its boundaries.

Stronger H-bonds with linearization; this is only considering all the ices excluding ices VII and VIII. As for the latter, in spite of having OHO bond angles between 177-178°, ices VII and VIII are associated with the weakest H-bonds. One reason could be the repulsion between adjacent O···O contacts within the closely intertwined ice VII/VIII network.

3.2.2 Distribution of H-bond Strengths

A box-and-whisker diagram was employed to analyze the statistical distribution of H-bond strengths in each ice. It can be deduced from Figure 4 that there is a narrow to broad distribution of H-bond strengths considering all the ices. The variation of bond strength distribution can be explained in terms of local structural features of the H-bond topology in each ice, which we will delve into more in detail in Section 3.3. However, a comparison of bond strength distribution between the ices at face value leads to an important result. It can be inferred from the narrow distribution of hexagonal ice that the H-bond strengths are more or less evenly distributed throughout its structure instituting more balance to the H-bond network. This consistency in the bond strengths could be the reason why hexagonal ice is the most prevailing form of ice on Earth under ambient conditions.
3.2.3 Classification of H-bonds

A set of classification criteria based on $k_a$ were defined to assign all 264 H-bonds within the ten ice polymorphs we investigated into different classes; strong, intermediate, and weak. These criteria were selected such that the majority of H-bonds that can be associated with the same local environment in a particular ice structure, i.e. rings, can be classified into the same class. This led to the classification of H-bonds as follows: $>0.296$ mdyn/Å as strong, 0.296-0.237 mdyn/Å as intermediate, and $<0.237$ mdyn/Å as weak. Figure 5 shows the percentage of each type of H-bond present in each of the ten ice polymorphs. Accordingly,
ices Ih, III, and IX are comprised of strong H-bonds and conversely, ices VI and VIII are formed from entirely weak H-bonds. The H-bond networks of ices II and XIV correspond to an admixture of strong and intermediate H-bond classes, while all three types of H-bonds are present in ices VI, XIII, and XV.

3.3 $k_n^a$ and Local Structural Features of H-bond Topology of Ices

As mentioned in Section 3.2.2 the rationale behind the overall trends in the bond strength distribution of ices can be uncovered by focusing on the local details of the structure. In the following, the H-bond networks of each of the ten ices are discussed with a particular emphasis on identification and characterization of their local structural elements, which cohere to give different topologies to these networks. Table 2 reports average bond lengths $R_{avg}$, average local stretching force constants $k_{n,avg}$, and number of H-bonds incorporated in different local structural elements identified within the H-bond networks of each ice. The H-bond networks for the ten ices are shown in Figures 6-12 with respect to a $2\times2\times2$ supercell, where the primitive unit cell is also indicated. Since, disordered/ordered ice pairs correspond to the same H-bond topology, H-bond networks and primitive unit cells of only ordered ices are shown. In some cases, noteworthy local structural elements are also presented separately. Different H-bond environments as identified by $k_n^a$ are distinguished using different colors. The $2\times2\times2$ supercell is represented in two ways: (i) Each O⋯O contact around the central water molecule within the local tetrahedral environment is represented by a stick, (ii) Water molecules and H-bonds are represented by balls and sticks, and dashed lines, respectively.

In general, up to 0.3% and 6.5% deviations in $R_{avg}$ and $k_{n,avg}$, respectively, can be seen within the different H-bond environments of H-ordered ices. In the case of H-disordered ices this is even more pronounced. The considerable deviations in $k_{n,avg}$ in both instances might be caused by numerical errors during the evaluation of Hessian matrix from numerical differentiation of analytical gradients, and also by the large size of primitive unit cells which contain 8-28 water molecules. However, this precision is sufficient for our purpose of
capturing trends in the bond strength variations.

Table 2: Average bond length values $R_{avg}$ in Å, average local stretching force constant values $k_{a,avg}$ in mDyn/Å, each followed by respective standard deviation values $\pm \Delta s$, and number of bonds in each environment for different local H-bond environments, within primitive unit cell and across its boundaries, for the 10 ices calculated at the BLYP-D3(BJ) level.

<table>
<thead>
<tr>
<th>type</th>
<th>$R_{avg}$ ($\pm \Delta s$)</th>
<th>$k_{a,avg}$ ($\pm \Delta s$)</th>
<th>no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ice Ih hex</td>
<td>1.688 ($\pm$ 0.001)</td>
<td>0.346 ($\pm$ 0.006)</td>
<td>24</td>
</tr>
<tr>
<td>Ice II hex</td>
<td>1.724 ($\pm$ 0.013)</td>
<td>0.332 ($\pm$ 0.017)</td>
<td>12</td>
</tr>
<tr>
<td>pillar</td>
<td>1.758 ($\pm$ 0.004)</td>
<td>0.280 ($\pm$ 0.009)</td>
<td>6</td>
</tr>
<tr>
<td>bridge</td>
<td>1.781 ($\pm$ 0.006)</td>
<td>0.258 ($\pm$ 0.009)</td>
<td>6</td>
</tr>
<tr>
<td>Ice III helix</td>
<td>1.702 ($\pm$ 0.010)</td>
<td>0.333 ($\pm$ 0.026)</td>
<td>8</td>
</tr>
<tr>
<td>bridge-1</td>
<td>1.734 ($\pm$ 0.008)</td>
<td>0.317 ($\pm$ 0.015)</td>
<td>8</td>
</tr>
<tr>
<td>bridge-2</td>
<td>1.713 ($\pm$ 0.010)</td>
<td>0.314 ($\pm$ 0.010)</td>
<td>8</td>
</tr>
<tr>
<td>Ice IX helix</td>
<td>1.700 ($\pm$ 0.000)</td>
<td>0.343 ($\pm$ 0.003)</td>
<td>8</td>
</tr>
<tr>
<td>bridge-1</td>
<td>1.720 ($\pm$ 0.000)</td>
<td>0.323 ($\pm$ 0.008)</td>
<td>8</td>
</tr>
<tr>
<td>bridge-2</td>
<td>1.722 ($\pm$ 0.000)</td>
<td>0.299 ($\pm$ 0.002)</td>
<td>8</td>
</tr>
<tr>
<td>Ice VII cubic</td>
<td>1.930 ($\pm$ 0.000)</td>
<td>0.187 ($\pm$ 0.001)</td>
<td>32</td>
</tr>
<tr>
<td>VIII cubic</td>
<td>1.930 ($\pm$ 0.000)</td>
<td>0.174 ($\pm$ 0.001)</td>
<td>16</td>
</tr>
<tr>
<td>Ice XIII octa-1</td>
<td>1.706 ($\pm$ 0.000)</td>
<td>0.341 ($\pm$ 0.022)</td>
<td>8</td>
</tr>
<tr>
<td>octa-2</td>
<td>1.729 ($\pm$ 0.004)</td>
<td>0.318 ($\pm$ 0.010)</td>
<td>20</td>
</tr>
<tr>
<td>octa-3</td>
<td>1.757 ($\pm$ 0.001)</td>
<td>0.290 ($\pm$ 0.013)</td>
<td>12</td>
</tr>
<tr>
<td>octa-4</td>
<td>1.774 ($\pm$ 0.001)</td>
<td>0.272 ($\pm$ 0.005)</td>
<td>8</td>
</tr>
<tr>
<td>thread</td>
<td>1.802 ($\pm$ 0.002)</td>
<td>0.239 ($\pm$ 0.011)</td>
<td>8</td>
</tr>
<tr>
<td>Ice XIV connect-1</td>
<td>1.758 ($\pm$ 0.004)</td>
<td>0.303 ($\pm$ 0.004)</td>
<td>8</td>
</tr>
<tr>
<td>connect-2</td>
<td>1.746 ($\pm$ 0.000)</td>
<td>0.290 ($\pm$ 0.003)</td>
<td>4</td>
</tr>
</tbody>
</table>
### Table 2 continued

<table>
<thead>
<tr>
<th>type</th>
<th>$R_{avg}$ (±$\Delta s$)</th>
<th>$k_{n,avg}$ (±$\Delta s$)</th>
<th>no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>zigzag-1</td>
<td>1.745 (± 0.000)</td>
<td>0.280(± 0.002)</td>
<td>4</td>
</tr>
<tr>
<td>connect-3</td>
<td>1.783 (± 0.000)</td>
<td>0.256 (± 0.003)</td>
<td>4</td>
</tr>
<tr>
<td>zigzag-2</td>
<td>1.778 (± 0.000)</td>
<td>0.245 (± 0.003)</td>
<td>4</td>
</tr>
<tr>
<td>Ice VI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>connect</td>
<td>1.724 (± 0.003)</td>
<td>0.378 (± 0.004)</td>
<td>4</td>
</tr>
<tr>
<td>octahed-1</td>
<td>1.753 (± 0.000)</td>
<td>0.283 (± 0.003)</td>
<td>2</td>
</tr>
<tr>
<td>octahed-2</td>
<td>1.761 (± 0.005)</td>
<td>0.277 (± 0.000)</td>
<td>4</td>
</tr>
<tr>
<td>octahed-3</td>
<td>1.769 (± 0.005)</td>
<td>0.270 (± 0.008)</td>
<td>2</td>
</tr>
<tr>
<td>octahed-4</td>
<td>1.823 (± 0.019)</td>
<td>0.216 (± 0.016)</td>
<td>4</td>
</tr>
<tr>
<td>octahed-5</td>
<td>1.855 (± 0.015)</td>
<td>0.198 (± 0.008)</td>
<td>4</td>
</tr>
<tr>
<td>Ice XV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>connect</td>
<td>1.733 (± 0.001)</td>
<td>0.392 (± 0.016)</td>
<td>4</td>
</tr>
<tr>
<td>octahed-1</td>
<td>1.760 (± 0.000)</td>
<td>0.263 (± 0.005)</td>
<td>2</td>
</tr>
<tr>
<td>octahed-2</td>
<td>1.782 (± 0.003)</td>
<td>0.256 (± 0.008)</td>
<td>4</td>
</tr>
<tr>
<td>octahed-3</td>
<td>1.776 (± 0.000)</td>
<td>0.238 (± 0.000)</td>
<td>2</td>
</tr>
<tr>
<td>octahed-4</td>
<td>1.801 (± 0.003)</td>
<td>0.229 (± 0.013)</td>
<td>4</td>
</tr>
<tr>
<td>octahed-5</td>
<td>1.853 (± 0.002)</td>
<td>0.183 (± 0.005)</td>
<td>4</td>
</tr>
</tbody>
</table>

**references**

- **$R_{avg}$**
  - H$_2$O: O-H: 0.974
  - (H$_2$O)$_2$: OH···O: 1.928
  - (H$_2$O)$_6$: OH···O: 1.684
  - 2D H$_2$O$^a$: OH···O: 1.956

- $k_{n}$
  - H$_2$O: 7.683
  - (H$_2$O)$_2$: 0.170
  - (H$_2$O)$_6$: 0.340
  - 2D H$_2$O$^a$: 0.879

$^a$ See Ref. 79.

### 3.3.1 ice Ih

Hexagonal ice is a disordered ice phase where the H-bond framework is comprised of hexagonal rings in chair and boat conformations. The topology of H-bond network in Ih allows for hexagonal channels or chains of cavities which can entrap small atoms as helium.\textsuperscript{101}
We probed the hydrogen bond strengths in $I_h$ via $k_n^n$ and it was revealed that the primary structural elements of $I_h$, i.e. hexagonal rings, can be associated with the same H-bond environment (Figure 6). In other words, the H-bond strength in $I_h$ is uniformly distributed throughout the framework manifesting a highly balanced H-bond network.

![Figure 6: H-bond network of ice $I_h$.](image)

### 3.3.2 ice II

Ice II has a fully H-ordered structure composed of puckered hexagonal rings in the [111] direction of the rhombohedral unit cell. The hexagonal rings vary in their degree of puckering leading to flat- and chair-shaped rings. Four types of H-bonds can be identified within the H-bond network of ice II. $^{102}$ The H-bonds within the hexagonal rings are referred to as “flat bonds” and “chair bonds” while those linking hexagonal rings vertically and horizontally are called “pillar” and “bridge” bonds, respectively. In general, ice II structure can be apprehended as a collection of interlinked hexagonal columns, or in other words "ice nanotubes". $^8$ Nakamura et al. $^{23}$ determined the hydrogen bonding energies of these four types of H-bonds using pair interaction energies and their relative strengths were ranked in the following order, bridge<flat<chair<pillar. We also characterized the bond strengths of these four types of H-bonds and considering the marginal difference in $k_{n,avg}$ for "flat" and "chair" bonds, we identified that these hexagonal rings in fact belong to the same H-bond environment which is denoted by type hex. Accordingly, bond strengths in these three different H-bond environments were observed to vary in the order, bridge<pillar<hex, and a representation of these bonds within both the H-bond network and a pair of hexagonal columns is shown.
in Figure 7.

Figure 7: (a) H-bond network of ice II. Different H-bond environments are highlighted as hex (grey), pillar (yellow), and bridge (orange). (b) Side view of a pair of hexagonal columns in ice II.

3.3.3 ice III/IX

Ice IX has an almost completely H-ordered structure with a complex arrangement of five-membered rings in a tetragonal unit cell and its disordered counterpart is ice III. The structure contains two types of water molecules with oxygen atoms distinguished as O(1) and O(2). The topology of the H-bonded network is such that water molecules with atoms O(1) form helices along the c-axis and water molecules with atoms O(2) connect adjacent helices. By utilizing $k_n$ we could identify three types of H-bonds in total, one type corresponding to those connecting O(1) containing water molecules to each other along with two other different types of H-bonds that connect O(1) containing water molecules to those of O(2). In other words, first type of H-bonds connect the water molecules forming the helices,
and second and third types bridge those helices into a united framework. Referring to Figure 8 and Table 2, these three types of H-bonds termed as helix (in cyan), bridge-1 (in green), and bridge-2 (in grey) are found to have varying strengths in the order, bridge-2<bridge-1<helix, as observed for both ices III and IX.

![Figure 8: H-bond network of ice IX(III). Different H-bond environments are highlighted as helix (cyan), bridge-1 (green), and bridge-2 (grey).](image)

### 3.3.4 ice VII/VIII

Ices VII/VIII are composed of two independent and interpenetrating networks, each corresponding to ice Ic structure. Networks with these attributes are recognized as catenated or self-clathrate frameworks where two identical frameworks are inserted into each other with no H-bonds between the molecules that make up each framework. Similar to hexagonal ice, a uniform bond strength distribution can be identified pertaining to Ic structural elements of ices VII/VIII, which is termed as cubic H-bond type and depicted in Figure 9. In analogy to the hex type H-bonds that constitute the hexagonal rings of Ih, cubic type H-bonds in ices VII/VIII are much weaker.

### 3.3.5 ice XIII

The H-bond network in ice XIII is the most complicated among known ices and a lucid description of its distinct local features is not so straightforward. It can be perceived as a self-entangled network\(^{13}\) constituted of bands of fused eight-membered rings with H-bonded
chains threading through these rings. Varying ring sizes of four-, five-, six-, eight-membered rings are present within this framework. We could identify five different H-bond environments as depicted in Figure 10. The H-bonds, octa-1, octa-2, octa-3, and octa-4 which have decreasing bond strengths in the order mentioned, comprise eight-membered fused rings (Figure 10b), and in fact octa-1 and octa-2 type strong bonds combine these rings together. Meanwhile, the thread type H-bonds which have the lowest bond strengths are only visible

Figure 10: (a) H-bond network of ice XIII. Different H-bond environments are highlighted as octa-1 (cyan), octa-2 (green), octa-3 (grey), octa-4 (yellow), and thread (orange). (b) Fused eight-membered rings; (i) Top view, and (ii) Side view.
within the threads that shoot through eight-membered rings and besides thread type bonds octa-$n$ ($n = 2, 3, 4$) type bonds also merge to form these threads.

### 3.3.6 ice XIV

This is one of densest$^{104}$ ice forms among those with a simple H-bond topology without any interposition of structural elements. Three crystallographically different molecules$^{12}$ exist within this network, corresponding to five different types of H-bond strengths as revealed by $k_n$. As highlighted in Figure 11 in grey and orange, two types of H-bonds form "zig-zag" chains along the c-axis of the tetragonal unit cell while the other three types of H-bonds indicated by cyan, green, and yellow, connect these chains together furnishing a three-dimensional H-bond network. Their bond strengths tend to vary in the order, connect-1<connect-2<zigzag-1<connect-3<zigzag-2. Though it is implied from this particular projection that five- or six-membered rings could be present, in actuality only seven- or eight-membered rings are present within this framework.

![Figure 11: H-bond network of ice XIV. Different H-bond environments are highlighted as connect-1 (cyan), connect-2 (green), zigzag-1 (grey), connect-3 (yellow), and zigzag-2 (orange).](image)

### 3.3.7 ice VI/XV

The H-bond topology of this network also corresponds to the interpenetration of two identical networks with no H-bonds in between, as seen in ice VI/VIII framework. The structural
element of VI/XV network resembles a distorted octahedron, made up of six water molecules occupying the vertices of the octahedron, with only 8 H-bonds between the vertices. Each distorted octahedron is associated with five different H-bond environments as indicated in Figure 12 as octahed-n (n = 1, 2, 3, 4, 5) whereas the H-bonds termed as connect, which also turn out to be the strongest, link these octahedrons together.

Figure 12: (a) H-bond network of ice XV(VI). Different H-bond environments are highlighted as connect (cyan), octahed-1 (green), octahed-2 (grey), octahed-3 (yellow), octahed-4 (orange), and octahed-5 (brown). (b) A distorted octahedron.

3.3.8 General Remarks on the Strength of Local Structural Features

Overall, our study clearly reveals that the majority of the ices are composed of different types of H-bonds with varying strengths. In the case of ices IIh and VIII (VII), only one type of H-bond is present within the H-bond network of each ice, termed as hex and cubic types, respectively. Although, both types of bonds are arranged in a 6-membered ring, it
is interesting to see the striking difference between their bond strengths where hex type is significantly stronger than cubic type, by about 0.172 (0.159) mdyn/Å, considering ices VIII (VII). This could be a direct consequence of the topology of the H-bond network, which corresponds to the interposition of two identical Ic networks in the case of ices VIII (VII). In this setting, the repulsive nature between the oxygen atoms in close proximity leads to the weakening of the cubic type bonds. Apart from the mentioned, all other ices accommodate different types of H-bonds in their framework, which could be somewhat counterintuitive given the fact that all the ices originate from the same local tetracoordinated environment. This is due to the fact that according to the Bernal-Fowler ice rules the arrangement of these local environments gives rise to different local structural elements which in turn can be attributed with different H-bond strengths.

In addition, it is of interest to inquire the underlying physical phenomena that contributes to the different bond strengths in various types of H-bonds within a particular ice. As shown by Tao et al.\textsuperscript{74} in their investigation of different H-bond types in water clusters, when a donor and an acceptor water molecule participating in a H-bond, simultaneously accepts from and donates to other peripheral water molecules, the target H-bond is strengthened by the so called push-pull effect. The ideal situation is where the donor molecule accepts two peripheral water molecules which pulls electron density from the donor, meanwhile where the acceptor donates to two peripheral water molecules which pushes electron density towards the acceptor. This synchronized effect ultimately leads to a directional high polarization of the target H-bond. In analogy to H-bonds in ice, this optimal situation is perturbed where donor and acceptor can donate and accept an additional H-bond as dictated by the tetragonal coordination. In principle, when this perturbation significantly changes the directionality of the push-pull effect, the bond-strength of the target H-bond undergoes substantial changes. We tested this hypothesis by comparing the push-pull setting for different H-bond types in ice II as a case study.

Figure 13 depicts the directionality of the push-pull effect that can be observed in a typical
setting for a hex, pillar, and bridge H-bond in ice II. The black arrows indicate the same
directionality of the *push-pull effect* while dashed red arrows highlight the deviations from the
said direction. The peripheral water molecules that are responsible for these deviations are
marked as $P_x$ and $P_y$, and we monitored angles $P_xDA$ and $P_yAD$ with donor (D) and acceptor
(A) water molecules of the target bond for the three types of H-bonds. In comparison to
both hex and pillar, the bridge type H-bond corresponds to angles which are more in-plane
($125.5^\circ$ and $126.9^\circ$) with D and A water molecules. In contrast, for both hex and pillar type
H-bonds, one of the angles leading to the deviations is out-of-plane ($85.0^\circ$ and $85.4^\circ$). Thus,
it can be deduced that in the case of bridge type H-bonds, the in-plane angles lead to a higher
perturbation in the *push-pull* direction, thereby weakening the target H-bond. On the other
hand, when the peripheral water molecules supporting the *push-pull* direction are associated
with more in-plane angles, the consequent *push-pull effect* is high, which is reflected in higher
bond strengths. As compared for pillar and hex types, the latter corresponds to more in-
plane angles that are close to $120^\circ$, and the enhanced *push-pull effect* in this situation results
in stronger hex type bonds.

Figure 13: Schematic of *push-pull effect* for the hex, pillar, and bridge type bonds in ice
II. The target H-bond between the donor (D) and acceptor (A) water molecules is marked
in blue where dotted black lines show the H-bonds with peripheral water molecules. The
adjacent water molecules are in transparent. The direction of the *push or pull* is indicated by
black arrows and dashed red arrows indicate those which deviate from that direction. The
angles between the peripheral water molecules, $P_x$ and $P_y$, and D and A water molecules are
given by $P_xDA$ and $P_yAD$. 
3.4 $k_n^a$ and Water-Ice Phase Diagram

The water-ice phase diagram presents the domains of stability for known and predicted ices in the phase range up to 100 GPa. Several experimental and/or computational investigations have led to the determination of phase boundaries between ices $I_h$, II, III, V, VI, and VII and water while fields of stability for some ices like ice XIII and XIV have not been established yet. Ices Ic, IV, IX, XII are metastable since they do not have their own regions of stability but exist in the stability range of another ice phase. Thermodynamically stable ices can also exist beyond their fields of stability, but preferentially transforms to the most stable form under that P-T conditions, i.e. most high-pressure ices are metastable at ambient pressures and slowly transform to $I_h$. Exploration of the water-ice phase diagram has many facets; it can be analyzed in a thermodynamic aspect to obtain information about phase transitions and equations of state for coexisting phases; or can be studied for its chemical dimensions like the influence by acid/base dopants on H-ordering processes, or the focus can be on extreme conditions like terapascal (TPa) pressure regimes which extend our understanding on morphology and evolution of ice giants in the solar system.

An alternative route to traverse the water-ice phase diagram is to monitor the variation of lattice energies across the different ice phases. The experimental lattice energies by Whalley, which extrapolate phase coexistence lines measured at finite temperature and pressure to zero temperature and pressure, have been utilized to compare the stability of ices. As the lattice energy reflects the collective H-bond strengths of the H-bond network of a particular ice polymorph, we examined its variation across the phase diagram moving from ambient to high pressures based on the calculated lattice energies in this work. We could observe moving from left to right in the phase diagram, that for ices with unambiguously defined phase boundaries, the corresponding calculated lattice energies decreased in the order, $I_h > II > III > XV > VI > VIII > VII$. Not so surprisingly, we could see this trend duly reflected in the subsequent analysis based on the characterization of the H-bond network via $k_n^a$, which is explained in detail below.
Figure 14: H-bond strength variation across the water-ice phase diagram. Grey bars in the bottom, spanning three pressure regions (0-0.08, 0.8-1.5, 1.5-50 GPa), indicate upper and/or lower boundaries of $k_n^a$ values, followed by nature of scattering of $k_n^a$ values and dominating H-bond class. The latter is obtained via considering higher percentage of strong, intermediate, or weak H-bond class for ices included within a particular pressure region. Circles and squares distinguish between disordered and ordered ices respectively, and colors for indicating different ices are consistent with Figures 1-3.

As shown in Figure 14, the ices within gradually increasing pressure regimes of the phase diagram are associated with their dominating H-bond classes based on our strong, intermediate, and weak classification presented in Section 3.2.3. Accordingly, ices Ih, II, III, IX and XIII which can be found in the 0-0.8 GPa region of the phase diagram are dominated by strong H-bonds with a low scattering of bond strength values, latter higher than 0.28 mdyn/Å. The ices in the central region of the phase diagram (0.8-1.5 GPa) have a high percentage of intermediate H-bonds, with the largest scattering of bond strength values, varying between 0.15-0.40 mdyn/Å. The ices VII and VIII in the very high pressure end of the phase diagram (up to 50 GPa) are mostly comprised of weak H-bonds. Thus, going from left to right in the phase diagram, the predominant class of bond strengths associated with different H-bond network topologies of ices vary from strong to weak. Consequently, increasing pressures are required in order to stabilize these ices. Reverting to the similar trend observed with lattice energies, this is a confirmation that H-bond strength measured via $k_n^a$ can correctly capture the binding energy between water molecules as collectively
measured via lattice energy. Furthermore, as $k_n^a$ calculation does not impose pressure, it can exclusively characterize the pure electronic effects of H-bonding in ice.

4 Conclusions

Herein, we employed the extension of LVM theory to solid state systems to characterize H-bonds in several ice crystals. The local environment in the H-bond network of ice corresponds to a tetrahedral coordination at each water molecule and the global topology of ice can be perceived to emerge from the integration of these similar sites according to the Bernal-Fowler ice rules. Although, the general view so far was in support of a homogenous H-bond strength distribution within ices, we could clarify in this work that this is not the case. Based on the assessment of the intrinsic H-bond strength within the ten ices we investigated, we could observe an array of bond strength distributions. Apart from hexagonal ice, frequently occurring form of ice on the Earth surface, and ice VIII(VII), existing in the Earth interior, both of which displayed a uniform distribution of H-bond strengths as quantified via $k_n^a$, all the other ices exhibited disparate bond strength distributions. We could explain these broad bond strength distributions in terms of different local structural elements embedded within the H-bond framework of each ice. Different H-bond types with varying strengths can be associated with these elements and hence the disparity. The bond strengths in different H-bond types in a particular ice are governed by the push-pull effect. Moreover, based on the strong, intermediate, and weak H-bond classification presented in this work, we could shed light on the H-bond strength variation across the water-ice phase diagram. In summary, these new findings on the H-bond strengths of the known ices provide important new guidelines with regard to the verification of the existence of different H-bond networks, bolstering research in novel ice structure prediction.
Acknowledgement

This work was supported by the National Science Foundation (grant CHE 1464906). The authors thank Southern Methodist University for providing excellent computational resources.

Supporting Information Available

The following files are available free of charge.

- esi.pdf: Local mode properties of H-bonds and donor OH bonds of the ten ices (Tables S1-S2), Peripheral angles supporting the push-pull effect in a typical setting of different H-bond types in ice II (Figure S1)

- cif.zip: Optimized geometries in fractional coordinates and primitive unit cell parameters for the ten ices at the BLYP-D3(BJ)/6-31+G(d,2p) level of theory

References


(37) Konkoli, Z.; Larsson, J. A.; Cremer, D. A New Way of Analyzing Vibrational Spectra.


(56) Oliveira, V.; Kraka, E.; Cremer, D. The Intrinsic Strength of the Halogen Bond: 

(57) Oliveira, V.; Kraka, E.; Cremer, D. Quantitative Assessment of Halogen Bonding 

(58) Oliveira, V.; Cremer, D. Transition from Metal-Ligand Bonding to Halogen Bond-
ing Involving a Metal as Halogen Acceptor: A Study of Cu, Ag, Au, Pt, and Hg 

(59) Yannacone, S.; Oliveira, V.; Verma, N.; Kraka, E. A Continuum from Halogen Bonds 
to Covalent Bonds: Where Do $\lambda^3$ Iodanes Fit? *Inorganics* 2019, 7, 47.

(60) Oliveira, V. P.; Kraka, E.; Machado, F. B. C. Pushing 3c-4e Bonds to the Limit: A 
2019, 58, 14777–14789.

(61) Oliveira, V. P.; Marcial, B. L.; Machado, F. B. C.; Kraka, E. Metal-Halogen Bonding 

(62) Setiawan, D.; Kraka, E.; Cremer, D. Description of Pnicogen Bonding with the help of 

(63) Setiawan, D.; Kraka, E.; Cremer, D. Strength of the Pnicogen Bond in Complexes 

(64) Setiawan, D.; Cremer, D. Super-Pnicogen Bonding in the Radical Anion of the Fluor-

(65) Oliveira, V.; Cremer, D.; Kraka, E. The Many Facets of Chalcogen Bonding: Described 


(83) Dovesi, R.; Erba, A.; Orlando, R.; Zicovich-Wilson, C. M.; Civalleri, B.; Maschio, L.; Rérat, M.; Casassa, S.; Baima, J.; Salustro, S.; Kirtman, B. Quantum-Mechanical Con-


Graphical TOC Entry

Local Stretching Force Constant $k^2$ decreases $H$-bond types phase diagram
Appendix D

Paper on hydrogen release from water

Modeling Hydrogen Release from Water with Borane and Alane Catalysts: A Unified Reaction Valley Approach

Sadisha Nanayakkara, Marek Freindorf, Yunwen Tao, and Elfi Kraka*  

INTRODUCTION

Economic, energy-efficient, and environmentally benign hydrogen production is highly targeted in the transition from conventional fossil fuels to clean and renewable hydrogen power.1–4 However, the intended benefits of future hydrogen economy are vastly limited by the fact that we are still relying on fossil fuels as the primary source of hydrogen production. When switching to cleaner hydrogen production methods, water splitting is a promising approach, which can be done by various methods like electrolysis, thermolysis, thermochemical reactions, and photocatalysis. There are significant efforts underway for developing techniques associated with water splitting methods to achieve energy-efficient hydrogen production, as highlighted in several reviews5 and publications.6–17

Water splitting in the presence of catalysts has presented itself as a viable option to accelerate hydrogen release from water at a lower energy cost.18,19 The choice of catalysts for this purpose is largely governed by the ability to form dihydrogen interactions with water. It has been shown that dihydrogen bonding20–22 of the type X–H···H···Y plays an important role in hydrogen evolution from hydrogen storage systems such as borane amines and alane amines by reducing the reaction barrier.23–26 Some other aspects to be considered for a viable hydrogen carrier, specifically rendering them as attractive systems for on-board hydrogen supply, are low molecular weight and high hydrogen content.27 Accordingly, main-group-element hydrides like boranes and alanes have been identified as promising candidates for facilitating hydrogen release from water, as documented in several experimental28–30 and theoretical31–32 studies. For example, Nguyen and co-workers33 investigated the catalytic effects of monomers and dimers of boranes and alanes, BH$_3$, AlH$_3$, B$_2$H$_6$, Al$_2$H$_6$, and B–Al mixed compound, AlH$_3$BH$_3$, with water. They used high-accuracy potential energy surface (PES) calculations to determine energetic parameters. In addition, they analyzed the stationary points on the PES with atom-in-molecule (AIM)33,34 and electron localization function (ELF)35 approaches.

While the results of this study help to evaluate the overall performance of these catalysts, more in-depth insights need a detailed mechanistic analysis, which is not only limited to energetics or stationary points on the PES but also examines the entire reaction pathway in terms of different reaction parameters. An analysis of this scope is possible through the unified reaction valley approach (URVA) of Kraka and
Cremer,36−38 which will be described in more detail in the Computational Methods section. Mainly, our study focused on addressing the missing gaps in knowledge regarding hydrogen release by those catalysts answering questions such as: What are the major mechanistic features along the way leading to dihydrogen bond formation? What is/are driving force(s) behind these reactions?

The major objectives of this study were the following: (i) to elucidate the mechanism of hydrogen release from water, (ii) to investigate solvent effects by describing the aqueous solution using polarizable continuum model39 (reaction oR1PCM), (iii) to investigate the role of the catalyst and identify mechanistic differences between B and Al, and (iv) to derive guidelines for improving borane and alane catalysts. For this purpose, reactions oR1−oR9 with one-water molecule (see Figure 1) and tR1−tR12 with two-water molecules (see Figure 2) were investigated. The connection between reaction models with one- and two-water molecules is shown in Table 1.

A key distinction has to be made between all reactions, based on whether hydrogen release is accomplished during the reaction prior to understanding their reaction mechanism in depth. Reactions oR3 and oR7 and the two-water analogues of the latter, tR8 and tR9, fall into the category (marked within the respective reaction schemes) where hydrogen release is not possible during the reaction, which deem these catalysts as unsuitable options to be used in H2 production.

## COMPUTATIONAL METHODS

The major tools applied in this work are the unified reaction valley approach (URVA)36−38 and the local vibrational mode analysis of Konkoli and Cremer,40−46 which will be summarized in the following text. URVA is based on analyzing the curving of the reaction path traced by the reaction complex (RC, e.g., the union of reacting molecules) on the potential energy surface (PES) from the entrance to exit channel via the TS. As the reaction proceeds, the RC undergoes electronic structure changes that are directly reflected by the scalar reaction path curvature calculated at each path point s.38 The maxima and minima of the curvature profile can be used to retrieve significant chemical information. The curvature maxima correspond to locations where chemical events take place such as bond formation and cleavage, charge transfer, charge polarization, and rehybridization. Contrarily, the curvature minima represent locations of minimal electronic structure change, which can be associated with hidden intermediates.47,48 A reaction phase47,48 is defined as the path region between two minima enclosing a curvature peak. In this way, each chemical reaction is uniquely attributed with a signature pattern of curvature maxima and minima with a different number of reaction phases, i.e., a fingerprint of the reaction. Decomposition of the reaction path curvature into internal coordinate components reveals which parameter(s) dominate(s) a chemical event. A positive component supports, while a negative component resists the chemical change. For a detailed mathematical derivation and recent advances of URVA, the interested readers are referred to ref 38.
The other tool utilized in this work is the local vibrational mode analysis, which is based on vibrational spectroscopy. In particular, local vibrational force constants $k_a$ are a quantitative measurement of the intrinsic bond strength. Local stretching force constants have been extensively applied to assess the intrinsic strength of both covalent bonding $^{49-63}$ and non-covalent $^{64-85}$ interactions. For a detailed mathematical derivation of the local vibrational mode theory, the interested readers are referred to a recent review article $^{46}$ and recent advances involving the local vibrational mode theory can be found in refs $^{63}$ and $^{86}$.

For simplification, the local mode force constants $k_a$ are converted into relative bond strength orders (BSOs $n$) by employing an extended Badger rule $^{48,87}$ according to which the BSO $n$ is related to the local mode force constants $k_a$ via the following power relationship $^{87}$

$$n_a k_a BSO ( ) = (1)$$

The constants $a$ and $b$ are fully determined by the $k_a$ values of two reference molecules with well-defined bond orders along with the requirement that for a zero-force constant, the BSO $n$ value becomes zero. In this work, B–H and Al–H bond cleavages as well as H–H dihydrogen bond formation were investigated. For BSO $n$(BH), we used as reference the BH bond in BH$_3$ and B$_2$H$_7^-$, having $k_a$ values of 4.128 and 1.901 mdyn/Å, respectively, and the corresponding Mayer bond orders $^{88}$ of 0.991 and 0.485. For BSO $n$(AlH), we used the AlH bonds in AlH$_3$ and Al$_2$H$_7^-$, having $k_a$ values of 2.318 and 1.143 mdyn/Å, respectively, and the corresponding Mayer bond orders of 0.931 and 0.423. For BSO $n$(BO), we used the BO bond in BO$_2^-$ and BO$_3^3-$, having $k_a$ values of 6.308 and 5.485 mdyn/Å, respectively, and

Table 1. Relation between Reactions Involving One-Water Molecule and Their Analogous Reactions with Two-Water Molecules Investigated in This Work

<table>
<thead>
<tr>
<th>catalyst</th>
<th>one-water$^a$</th>
<th>two-waters$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BH$_3$</td>
<td>oR1</td>
<td>tR1</td>
</tr>
<tr>
<td></td>
<td>tR2</td>
<td>tR3</td>
</tr>
<tr>
<td></td>
<td>tR4</td>
<td>tR5, step 1</td>
</tr>
<tr>
<td>AlH$_3$</td>
<td>oR2</td>
<td>tR3</td>
</tr>
<tr>
<td></td>
<td>tR4</td>
<td>tR5, step 2</td>
</tr>
<tr>
<td>B$_2$H$_6$</td>
<td>oR5, step 1</td>
<td>tR6, step 1</td>
</tr>
<tr>
<td></td>
<td>oR5, step 2</td>
<td>tR6, step 2</td>
</tr>
<tr>
<td></td>
<td>oR6</td>
<td>tR7</td>
</tr>
<tr>
<td>Al$_2$H$_6$</td>
<td>oR7</td>
<td>tR8</td>
</tr>
<tr>
<td></td>
<td>oR8</td>
<td>tR9</td>
</tr>
<tr>
<td></td>
<td>oR9</td>
<td>tR10</td>
</tr>
<tr>
<td>AlH$_3$BH$_3$</td>
<td>oR11</td>
<td>tR11</td>
</tr>
</tbody>
</table>

$^a$Reactions involving one-water molecule start with prefix o; $^b$Reactions involving two-water molecules start with prefix t.

The other tool utilized in this work is the local vibrational mode analysis, which is based on vibrational spectroscopy. In particular, local vibrational force constants $k_a$ are a quantitative measurement of the intrinsic bond strength. Local stretching force constants have been extensively applied to assess the intrinsic strength of both covalent bonding $^{49-63}$ and non-covalent $^{64-85}$ interactions. For a detailed mathematical derivation of the local vibrational mode theory, the interested readers are referred to a recent review article $^{46}$ and recent advances involving the local vibrational mode theory can be found in refs $^{63}$ and $^{86}$.

For simplification, the local mode force constants $k_a$ are converted into relative bond strength orders (BSOs $n$) by employing an extended Badger rule $^{48,87}$ according to which the BSO $n$ is related to the local mode force constants $k_a$ via the following power relationship $^{87}$

$$n_a k_a BSO ( ) = (1)$$

The constants $a$ and $b$ are fully determined by the $k_a$ values of two reference molecules with well-defined bond orders along with the requirement that for a zero-force constant, the BSO $n$ value becomes zero. In this work, B–H and Al–H bond cleavages as well as H–H dihydrogen bond formation were investigated. For BSO $n$(BH), we used as reference the BH bond in BH$_3$ and B$_2$H$_7^-$, having $k_a$ values of 4.128 and 1.901 mdyn/Å, respectively, and the corresponding Mayer bond orders $^{88}$ of 0.991 and 0.485. For BSO $n$(AlH), we used the AlH bonds in AlH$_3$ and Al$_2$H$_7^-$, having $k_a$ values of 2.318 and 1.143 mdyn/Å, respectively, and the corresponding Mayer bond orders of 0.931 and 0.423. For BSO $n$(HH), we used the HH bond in H$_2$ and H$_2^+$, having $k_a$ values of 6.308 and 1.740 mdyn/Å, respectively, and the corresponding Mayer bond orders of 0.931 and 0.423. For BSO $n$(BO), we used the BO bond in BO$_2^-$ and BO$_3^3-$, having $k_a$ values of 9.793 and 5.485 mdyn/Å, respectively, and

Figure 2. Reactions TR1–TR12 with two-water molecules investigated in this work. Reactions shown within blue boxes represent those that do not release hydrogen during the reaction.
the corresponding Mayer bond orders of 1.474 and 1.197. For BSO \((\text{AlO})\), we used the AlO bond in \((\text{AlO})_2^-\) and \((\text{AlO})_4^-\), having \(k\) values of 5.326 and 2.574 mdyn/Å, respectively, and the corresponding Mayer bond orders of 1.666 and 1.150. All values refer to an MP2/6-31+G(d,p) model chemistry.\(^{89-93}\)

A dual-level approach was utilized in this work, which has been proven to be effective in studying larger reaction systems in conjunction with obtaining accurate energetics.\(^{94-97}\) The reaction valley was evaluated at the MP2 level with Pople’s 6-31+G(d,p) basis set for all reactions \(\text{oR1} \rightarrow \text{oR9}\) and \(\text{tR1} \rightarrow \text{tR12}\). The energetics were calculated at the DLPNO-CCSD(T) level of theory\(^{98,99}\) utilizing MP2 geometries. The MP2 calculations were performed with a pruned ultrafine (99,590) integration grid and a tight convergence criterion. The aqueous solution applied in reaction \(\text{oR1}_{\text{PCM}}\) was described by the polarizable continuum solvent model (PCM) of Tomasi and co-workers.\(^{39}\)

As the reaction path, the intrinsic reaction coordinate (IRC) of Fukui\(^{100}\) was used with a step size of \(s = 0.03\) amu\(^{1/2}\) b, applying the improved reaction path following the algorithm of Hratchetian and Kraka, enabling us to follow a chemical reaction far out into the entrance and exit channels.\(^{101}\) The URVA analysis was carried out with the program pURVA,\(^{102}\) and local mode analysis was carried out with the program COLOGNE.\(^{20,103}\)

The quantum chemical calculations were performed using Gaussian99\(^{9,104}\) natural bond orbital (NBO) charges were calculated with the program NBO 6,\(^{105,106}\) and DLPNO-CCSD(T) calculations as well as the Mayer bond order analyses were performed using the ORCA\(^{98}\) program package.

### RESULTS AND DISCUSSION

First, the energetics of reactions \(\text{oR1} \rightarrow \text{oR9}, \text{tR1} \rightarrow \text{tR12},\) and \(\text{oR1}_{\text{PCM}}\) in aqueous solution are discussed. This is followed by the URVA analysis focused on discussing the major chemical events along the reaction path of monomers, dimers, and mixed compounds of B and Al in the presence of one-water (reactions \(\text{oR1} \rightarrow \text{oR9}\)) and two-water (reactions \(\text{tR1} \rightarrow \text{tR12}\)) molecules. Then, a local mode analysis describing the strength of the major bonds formed and cleaved during the reactions is presented taken at the reactant (R), transition state (TS), and product (P) of each reaction. The results are complemented by an investigation of charge transfer and a ring puckering analysis.

**Reaction Energetics.** The activation energies and reaction energies discussed below refer to DLPNO-CCSD(T)/aug-cc-pVQZ calculations, which are summarized in Table 2. The energetics of the hydrogen release in the presence of one- and two-water molecules reveal large variations in the activation barriers, 6–47 kcal/mol for one-water, 11–56 kcal/mol for two-water. Most of the reactions involving one-water molecule are exothermic (4–30 kcal/mol) except for reactions \(\text{oR3}\) and \(\text{oR7}\) (\(\text{oR3}: 10.2\) kcal/mol, \(\text{oR7}: 23.0\) kcal/mol). Interestingly, \(\text{oR9}\) also is an endothermic reaction where \(\text{H}_2\) is duly liberated during the reaction. The overall trend in energetics is also true for the

---

**Table 2. Energetics (kcal/mol) for oR1–oR9 and tR1–tR12 in Gas Phase and oR1_{\text{PCM}} in Solution**

<table>
<thead>
<tr>
<th>Reaction</th>
<th>MP2/6-31+G(d,p)</th>
<th>DLPNO-CCSD(T)/aug-cc-pVQZ/</th>
<th>MP2/6-31+G(d,p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta E^\text{a})</td>
<td>(\Delta E_R^\text{a})</td>
<td>(\Delta H^\text{a})</td>
<td>(\Delta H_R^\text{a})</td>
</tr>
<tr>
<td>One-Water</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>oR1</td>
<td>24.9</td>
<td>-17.9</td>
<td>24.6</td>
</tr>
<tr>
<td>oR2</td>
<td>17.2</td>
<td>-11.2</td>
<td>16.5</td>
</tr>
<tr>
<td>oR3</td>
<td>25.7</td>
<td>12.5</td>
<td>24.1</td>
</tr>
<tr>
<td>oR4</td>
<td>9.5</td>
<td>-6.0</td>
<td>9.8</td>
</tr>
<tr>
<td>oR5, step 1</td>
<td>5.9</td>
<td>0.4</td>
<td>5.1</td>
</tr>
<tr>
<td>oR5, step 2</td>
<td>3.6</td>
<td>-29.8</td>
<td>4.6</td>
</tr>
<tr>
<td>oR6</td>
<td>8.8</td>
<td>-27.6</td>
<td>8.3</td>
</tr>
<tr>
<td>oR7</td>
<td>47.5</td>
<td>24.1</td>
<td>47.4</td>
</tr>
<tr>
<td>oR8</td>
<td>5.1</td>
<td>-30.2</td>
<td>6.2</td>
</tr>
<tr>
<td>oR9</td>
<td>16.9</td>
<td>15.5</td>
<td>16.4</td>
</tr>
<tr>
<td>Two-Waters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tR1</td>
<td>32.2</td>
<td>-10.1</td>
<td>32.6</td>
</tr>
<tr>
<td>tR2</td>
<td>29.3</td>
<td>-13.1</td>
<td>29.6</td>
</tr>
<tr>
<td>tR3</td>
<td>26.6</td>
<td>-3.6</td>
<td>26.8</td>
</tr>
<tr>
<td>tR4</td>
<td>23.3</td>
<td>-3.1</td>
<td>23.2</td>
</tr>
<tr>
<td>tR5, step 1</td>
<td>21.6</td>
<td>5.4</td>
<td>20.4</td>
</tr>
<tr>
<td>tR5, step 2</td>
<td>15.0</td>
<td>-1.3</td>
<td>15.0</td>
</tr>
<tr>
<td>tR6, step 1</td>
<td>5.2</td>
<td>-1.9</td>
<td>4.8</td>
</tr>
<tr>
<td>tR6, step 2</td>
<td>9.9</td>
<td>-24.6</td>
<td>11.6</td>
</tr>
<tr>
<td>tR7</td>
<td>11.9</td>
<td>-22.4</td>
<td>11.4</td>
</tr>
<tr>
<td>tR8</td>
<td>55.5</td>
<td>24.6</td>
<td>56.2</td>
</tr>
<tr>
<td>tR9</td>
<td>53.1</td>
<td>23.7</td>
<td>52.9</td>
</tr>
<tr>
<td>tR10</td>
<td>12.8</td>
<td>-28.2</td>
<td>13.4</td>
</tr>
<tr>
<td>tR11</td>
<td>10.6</td>
<td>-25.4</td>
<td>11.6</td>
</tr>
<tr>
<td>tR12</td>
<td>22.0</td>
<td>21.8</td>
<td>22.3</td>
</tr>
</tbody>
</table>

### Aqueous Solution

<table>
<thead>
<tr>
<th>Reaction</th>
<th>MP2/6-31+G(d,p)</th>
<th>DLPNO-CCSD(T)/aug-cc-pVQZ/</th>
<th>MP2/6-31+G(d,p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>oR1_{\text{PCM}}</td>
<td>30.6</td>
<td>-11.0</td>
<td>30.3</td>
</tr>
</tbody>
</table>

\(^a\)Activation energies \(\Delta E^\text{a}\), enthalpies \(\Delta H^\text{a}\), and free energies \(\Delta G^\text{a}\); reaction energies \(\Delta E_{\text{R}}\) enthalpies \(\Delta H_{\text{R}}\) and free energies \(\Delta G_{\text{R}}\). For the coupled cluster enthalpies and free energies, MP2 thermochemistry was used. \(^b\)Using the PCM model.
analogous reactions with two-water molecules, where the
exothermicity is generally reduced (1−26 kcal/mol), whereas
the endothermicity is increased for tr12 (two-water analogue of
or9, 22.1 kcal/mol) and decreased for tr5 (or3 analogue, 4.0
cal/mol) and tr9 (or7 analogue, 22.7 kcal/mol). The high
activation barrier in the presence of two-water molecules are
obviously a result of enhanced stabilization of the RC by
hydrogen bonds in the presence of an additional water molecule.
In Figure 3, the energy profiles for reactions with one- and
two-water molecules in the gas phase (GP) and in aqueous solution
are shown. The activation barrier in the aqueous solution
(reaction or1pcm) is increased by about 5 kcal/mol indicating
that H2 evolution is impeded in the aqueous solution.

Curvature Profiles and Reaction Mechanism. In the
following, curvature profiles of or1–or9 and tr1–tr12 will be
discussed with an emphasis on the mechanistic differences
between reactions involving one-water molecule and their
analogous reactions with an additional water molecule. Addition-
alily, the influence of aqueous solution is analyzed via reaction or1pcm. In Figures 4 and 5, the scalar reaction path curvatures and their decomposition into most important bond
length components for these reactions are shown. Prechemical
phases, which are associated with reorientations of the RC, and
postchemical phases, which do not contribute to the reaction
barrier, are omitted from the discussion to focus on the actual
chemical events that occurred via bond formation and cleavage.
Thus, the reaction path is narrowed down to only show the
chemical phases. The uncurbed plots showing the entire
reaction path (see Figures S3 and S4) and the respective plots
for curvature decomposition into angle and dihedral com-
ponents (see Figures S5–S8) are included in the Supporting
Information.

An important observation is that except reactions or3 and
or7, all of the other reactions follow a similar curvature pattern
within their respective chemical phases (highlighted with a gray
bar). The chemical phases of these reactions (except or3 and
or7) are characterized by three curvature peaks K1, K2, and K3,
recognizable as well-separated distinct peaks or small curvature
enhancements in some cases. These indicate the major chemical
processes of OH cleavage, H2 formation, B/H/AlH4 component (i = x
for or1, or2, and or6 reactions and i = y for all other reactions)
cleavage, and OB/OAl formation. In all cases, the sequence of
chemical events starts with the OH cleavage as characterized by
peak K1 before the TS. This peak is dominated by the OH contribution accompanied by a resisting contribution from
H2/H2. In addition, K1 has small supporting contributions from
OB/OAl and B/H/AlH4 components, where the latter can also be slightly resisting in some cases.

For the majority of reactions, in the next chemical phase after
the TS, the H2/H2 bond is formed as seen by the peak K2. Since
H2 formation occurs after the TS, it does not cost energy in these
cases. In or9, or5 (step 2), and or8, peak K2 occurs at or
slightly before TS, shifting the latter toward the exit channel
leading to a late transition state. The B/H/AlH4 bond cleavage
and OB/OAl bond formation processes span throughout the
entire chemical region, and the finalization of both bond
cleavage and formation is achieved at peak K3 where OB/OAl
shows a supporting contribution and B/H/AlH4 can show
either a small supporting or a resisting contribution. The
prolonged bond cleavage of B/H/AlH4 bonds, which expands
over several peaks, closely resembles that of a metal−hydrogen
bond cleavage process, which is cost-effective, and it is
interesting to observe a stepwise bond cleavage in the presence
of nonmetallic B. As emphasized earlier, during reactions or3
and or7 (see Figure 4d), no hydrogen is released, i.e., there is
no contribution from the H2/H2 component in the exit channel,
and thus, the dominating OH contribution is absent in the
entrance channel. These reactions are energy-demanding as
evident by their curvature profiles, which reveal that the B/H2
bond cleavage before the TS consumes energy. In comparison to
the two-step reaction or5 with alkane dimer, this unfavorable
situation is circumvented by cleaving the AlH4 first in the first
step under a low-energy penalty and proceeding into the actual
chemical events in the second step.

Reactions tr1–tr12 are characterized by the presence of an
additional water molecule. In reference to Table 2, only reactions with a reasonable activation barrier less than 25
cal/mol as well as those leading to H2 will be further discussed,
i.e., reactions tr4, tr6, tr7, and tr10–tr12. The remaining
reactions are included for completeness in the Supporting
Information (see Figure S4).

The chemical phases for reactions with two-water molecules
retain similar characteristics as in the one-water-molecule case
with three curvature peaks describing the major chemical events,
which are OH cleavage, H2 formation, B/H/AlH4 cleavage (i = x
for tr4 and tr7 reactions and i = y for all of the other
reactions), and OB/OAl formation. In the case of tr6 (step 2;
see Figure 5c) and tr10 (see Figure 5e), peak K2 corresponding
to the H2 formation appears as a small shoulder, and in tr12
(see Figure 5g), it becomes a small curvature enhancement.
Similar to their one-water analogues, reactions tr6 (step 2),
tr11, and tr12 are also associated with a late TS.
Figure 4. continued

(a) Reaction oR1, PCM

(b) Reaction oR1

(c) Reaction oR2

(d) Reaction oR3

(e) Reaction oR4

(f) Reaction oR5, Step 1

(g) Reaction oR5, Step 2
In addition to the major components, there are also other components that contribute significantly to the reaction mechanism. In the two-step reactions, oR5 and tR6, the first step goes through cleavage of an Al-Ha bond that is crucial to acquire the correct orientation of the RC in order for the two reacting hydrogens to come in close contact. Also in reactions oR5-step 2, oR6, and oR8, along with their two-water analogues tR6-step 2, tR7, tR10, and tR11, the formation of an OAl interaction is observed as indicated by the subsequent peak after the chemical phase (see Figure S4 in the Supporting Information). This leads to the stabilization of the byproduct, which is reflected in the high exothermicity associated with these reactions.

Relative Bond Strength Orders. O--H Cleavage. As the reaction sequence starts off with the cleavage of the OH bond in all reactions except those with no hydrogen release, we investigated how the strength of the OH bond is influenced by the borane and alane catalysts. We applied the local mode analysis to examine the relative change in $k_a$ from the R to TS, as shown in Table 3. In line with the curvature data, reactions oR3, oR7, tR8, and tR9 have the weakest H--H bonds at the TS, further confirming that H$_2$ evolution is impossible during these reactions. All of the other reactions demonstrate moderate-to-strong H--H bonds at the TS. Additionally, we wanted to clarify whether stabilizing electrostatic interactions between hydridic H$\delta^-$ from B(Al) and protonic H$\delta^+$ from water can contribute to the activation barrier as postulated in previous works involving borane and alane hydrides. As $k_a$ is able to capture all electronic and steric effects, it is an ideal measure of the strength of the stabilizing interactions within the dihydrogen bond. Although there is no direct correlation between the activation energy and $k_a$ (see Figure 7), three trends emerge (marked within differently colored dashed ovals). The region encircled in red corresponds to the reactions with high activation energies and weak H--H bond strengths. The other two regions encircled in purple and green correspond to all of the other reactions with lower activation energies and moderate-to-strong H--H bonds. Interestingly, all reactions within the green region proceed through a six-membered cyclic TS, compared to the majority of reactions within the purple region, which have a four-membered cyclic TS except tR10. The distinction between the two regions most probably originates from the steric strain inherent to the four-membered cyclic TS.

B(Al)--H Cleavage. As revealed by the curvature peaks, the B(Al)--H cleavage process starts before the TS and proceeds into the subsequent phases, leading to a gradual weakening of the B(Al)--H bond for all reaction resulting in hydrogen...
Figure 5. Scalar curvature as a function of the reaction path parameter $s$ (solid black line) for the hydrogen release in the presence of two-water molecules for $tR4$, $tR6$, $tR7$, and $tR10$–$tR12$. The reaction path is narrowed down to show only the chemical phases (gray bar). The decomposition of the scalar curvature into the most important bond components is given. The borders of the reaction phases are indicated by the vertical dashed lines at curvature minima $M1$, $M2$, $M3$, etc. Curvature maxima are indicated by $K1$, $K2$, $K3$, etc. The TS at $s = 0$ amu$^{1/2}$ b is also indicated by a vertical dotted line.
evolution. This process is analyzed via the BSO \( n \) values determined at the R, TS, and P, as shown in Figure 8. Since BSO \( n \) values were determined using two different power relationships for \( B-H \) (solid curve) and \( Al-H \) (dashed curve) bonds, the corresponding data points are separated into their respective curves. In the R (red points), \( B-H \) bonds appear to be generally stronger than \( Al-H \) bonds with BSO \( n \) values between 0.8 and 1.0. The BSO \( n \) values decrease at the TS (blue points) but still remain moderately strong with values between 0.4 and 0.6, indicating that the bond cleavage has not been finalized yet. In the P (purple points), all except \( oR4 \) and \( oR9 \) have BSO \( n \) values close to zero, those with marginally strong \( B-H \) bonds implying a weak \( B-H \) interaction in the P. A similar analysis was done for reactions with two-water molecules and is included within the Supporting Information along with local mode properties at R, TS, and P for all reactions (see Figure S10 and Tables S1 and S2).

O–B(Al) Formation. A critical consideration in the catalytic \( H_2 \) production is the regeneration of the catalyst from the borate and aluminate byproducts, which is required to close the hydrogen cycle. This is very challenging due to the inefficiency in the process of breaking an O–B(Al) bond in the byproduct to generate a B–H bond.\(^{108-110}\) Utilizing the local mode analysis to inspect the intrinsic bond strength of O–B(Al) bond(s) in the byproduct, we were able to shed some new light on the regenerability of the catalyst, which is a complex process and a comprehensive analysis is required, however, to further understand its proper mechanism.

As seen from Figure 9, the strongest O–B(Al) bonds are formed in the byproducts from the reactions \( oR1, oR2, oR4, \)
or6, and or9 and their two-water analogues tr1−tr5, tr7, and tr12, having BSO n values between 1.0 and 1.3. Reactions or6 and tr7 lead to relatively weak secondary (numbered as 2, BSO n: 0.6) Al−O bonds; however, having strong primary (numbered as 1, BSO n: 1.1) Al−O bonds restricts the efficiency of these byproducts in the regeneration process. Interestingly, the byproducts from reaction or8 along with its two-water analogues tr10 and tr11 emerge to have very weak B−O bonds (BSO n: 0.5), in addition to moderately strong primary Al−O bonds (BSO n: 0.8−0.9), rendering it as a desirable candidate to be considered for catalyst regeneration. In addition, similar to the B(Al)−H cleavage, the O−B(Al) formation also follows a stepwise procedure with moderately strong O−B(Al) bonds at the TS (see Tables S3 and S4 of the Supporting Information).

**Charge Transfer.** We performed an NBO charge analysis along the reaction path, as shown in Figure 10, for the catalyst (Figure 10a,b) and the two hydrogen atoms (Figure 10c,d) taking part in the dihydrogen bond.

We observed that the charge transfer starts already in the van der Waals region, i.e., in the entrance channel, where the catalyst accepts charge from water in all reactions involving one- and two-water molecules. In the case of one-water molecule, the net charge acquired by the catalyst in the beginning goes to a maximum of 300 me, and in the two-water case 334 me. It seems that the introduction of a second water molecule to the initial RC has little influence on the charge-transfer process, which primarily occurs between the catalyst and the water molecule in direct contact. After passing through a minimum (see the dip in the curves in Figure 10a,b), the charge transfer switches direction before the TS for the majority of the reactions, revealing the bifunctional nature of borane and alane catalysts. It should be mentioned that in the first step of or5 and similarly in tr6, the charge-transfer switching point is observed after the TS, but in both cases, the major chemical events occur in step 2. However, or3 and or7 and their two-water analogues tr8 and tr9 are in contrary where they continue to act as charge acceptors as the reaction proceeds.

After switching direction, the catalyst starts to act as a charge donor with the above-mentioned exceptions, and for the majority of the reactions, it is observed that the net charge donated reaches a maximum during the chemical phases corresponding to the OH cleavage and HH formation. In fact, the highest charge donated within the chemical phases varies between 70 and 470 me in the one-water scenario and 60−700 me with two-water molecules. In all reactions where the catalyst acts as a charge donor in the chemical phases, an energy barrier of around 5−25 kcal/mol is observed with one-water molecule and 10−30 kcal/mol for two-water molecules. This increases up to 55 kcal/mol in the cases where the catalyst reverses its role and acts as a charge acceptor in the chemical phases. This is an indication that the preferred role for the catalyst is as a charge donor for a feasible hydrogen release.

We further analyzed in reaction or3 the inability to act as a donor, in comparison to or5 reaction where the catalyst is bifunctional, via a molecular orbital approach as provided by the NBO analysis. It was seen in or3, initially, that the electron density of the s-orbital in the bridged H atom is delocalized into nonbonding p orbitals of both B and H, as suggested by the strong interaction energies (second-order perturbation energies, E2: 817 and 878 kcal/mol). Meanwhile, the electron acceptor ability from the lone pair of the incoming O atom into the nonbonding p-orbital of B is greatly reduced (E2: 0.85 kcal/mol). As the reaction proceeds from the R to TS, the acceptor ability of the latter increases (E2: 142 kcal/mol), but the electron donor capacity from the B2−H2 sigma bond into the O−H2 sigma antibonding orbital remains very low (E2: 0.25 kcal/mol), consequently restricting the O−H2 bond cleavage and B2−H2 sigma and H2 s-orbital overlap that need to promote the dihydrogen interaction. In contrast, in or5, the nonbonding p-orbital of Al strongly accepts electron density from the O lone pair in the beginning (E2: 48 kcal/mol), and going toward the TS, the Al2−H2 sigma bond weakly donates (E2: 0.84 kcal/mol) to the O−H2 sigma antibonding orbital, which substantially increases (E2: 45 kcal/mol) in the intermediate after the first step, thereby promoting OH bond cleavage in the second step and subsequent H2 formation.

In addition, the charges on the two hydrogen atoms, H from catalyst and H2 from water, were analyzed as indicated in Figure 10c,d. It can be seen that H in the catalyst is more negatively charged than H2 in water in the beginning, and the charge on both H atoms deteriorates rapidly after TS, in all complete reactions as the H2−H2 bond is formed. The electrostatic attractions between the oppositely charged H atoms drive the H−H bond formation process, which is, however, unattainable in or3 and or7 and the two-water analogues of the latter, tr8 and tr9.

Considering the reaction in the aqueous solution, or1PCM (see Figure 10e), it can be seen that shortly before the TS, the catalyst changes the charge-transfer direction and proceeds as a charge donor, donating a maximum of 655 me. Altogether, the charge-transfer mechanism seems more or less similar to its gas-phase counterpart. However, in comparison to or1, the charge separation between the catalyst and water in the van der Waals complex appears to be higher, which is a consequence of the polar medium.

As charge transfer is a crucial factor in the reaction mechanism, we were interested in deriving an explicit relationship between the charge-transfer process and the activation energy ΔE∗. By taking into account that O−H and X−H (X = B, Al) bond cleavage processes occur or in the case of the latter initiates before the TS and these cost energy, it was worth analyzing how the change in the charge separation, going from the R to TS for the respective atoms of the cleaving bond, i.e., charge separation between the O and H atoms for the O−H bond, ...
bond and between the X and H atoms for the X−H bond, correlates with the ΔEa. It was immediately observed that there is a distinct division in the data set based on the predominant type of correlation as shown by Figure 10 for reactions oR1−oR9. For reactions oR1, oR2, oR4, and oR8, there is an inverse relationship with ΔEa for the cleaving O−H bond where a small change in the charge separation between the O and H atoms going from the R to TS results in a higher ΔEa (see Figure 11a). In contrast, reactions oR3, oR5, oR6, oR7, and oR9 show a direct correlation with the change in charge separation for the cleaving X−H bond where a large charge separation in the X−H bond as it goes from the R to TS leads to a higher ΔEa (see Figure 11b). For each of the reactions, by considering the predominant O−H or X−H contribution to the ΔEa and its nature, one can derive useful guidelines to modify the catalysts so that the ΔEa is reduced. Electron-withdrawing substituents on the catalyst will result in higher charge separation for the O−H bond, thereby reducing the ΔEa. Likewise, electron-donating substituents in the cases of oR3, oR5, oR6, oR7, and oR9 can reduce the charge separation in the X−H bond and lead to lower ΔEa. An analysis of similar nature was conducted for the two-water case, which shows parallel trends and is included in the Supporting Information (see Figure S9).
Puckering Analysis. It has been shown in previous studies that the catalyst actively participates in the hydrogen transfer through a six-membered cyclic TS, which has been observed for several reactions in this study. However, a quantitative measurement of the ring puckering and its influence on the hydrogen transfer mechanism has not been discussed so far. We used the Cremer–Pople puckering analysis to evaluate the puckering of the intermediate six-membered rings, which is discussed in detail in the following.

The Cremer–Pople puckering analysis probes the conformational changes in the RC during the reaction and provides a quantitative description of the intermediate rings. The puckering analysis for a six-membered ring is based on the following parameters: puckering amplitudes $q_1$ and $q_2$, the pseudo-rotation phase angle $\Phi_2$, the hyperspherical angle $\Theta$, and the total puckering amplitude $Q$. The combination of the pseudo-rotational coordinate pair $(q_2, \Phi_2)$ describes the pseudo-rotation of the boat and twistboat conformers of the six-membered ring, while the single-crown puckering amplitude $(q_1)$ describes the chair (positive $q_1$) and inverted chair forms (negative $q_1$). The contributions from each of the conformations, chair, boat, and twistboat forms, can be quantified using the values of the puckering coordinates as given by the following formulas:

- Chair: $100\% \times \frac{q_1^2}{Q^2}$
- Boat: $100\% \times \frac{q_1^2 \cos^2(3\Phi_2)}{Q^2}$
- Twistboat: $100\% \times \left[1 - \cos^2(3\Phi_2)\right] \frac{q_2^2}{Q^2}$

In Table 4, the percentage contribution from each of the chair, boat, and twistboat conformers is summarized for the TS of reactions involving a six-membered ring, which are $oR3$–$oR5$ and $oR7$–$oR9$ and their two-water analogous reactions $tR5$, $tR6$, and $tR8$–$tR12$. According to the puckering analysis, the TS of the majority of the reactions has a predominant twistboat character, the contributions varying from 36 to 93%. The transition state of reactions $oR3$ and $oR4$ along with $tR8$, $tR10$, and $tR11$ is predominated by the boat form with contributions 73.8, 86.5, 47.4, 66.5, and 56.6%, respectively.

Table 4. Puckering Analysis for the TS Involving a Six-Membered Ring Formation (% Contributions)

<table>
<thead>
<tr>
<th>reaction</th>
<th>chair</th>
<th>boat</th>
<th>twistboat</th>
</tr>
</thead>
<tbody>
<tr>
<td>$oR3$</td>
<td>15.8</td>
<td>73.8</td>
<td>10.4</td>
</tr>
<tr>
<td>$oR4$</td>
<td>12.4</td>
<td>86.5</td>
<td>1.0</td>
</tr>
<tr>
<td>$oR5$, step 2</td>
<td>14.0</td>
<td>0.1</td>
<td>85.9</td>
</tr>
<tr>
<td>$oR7$</td>
<td>27.5</td>
<td>6.6</td>
<td>65.9</td>
</tr>
<tr>
<td>$tR8$</td>
<td>28.0</td>
<td>47.4</td>
<td>24.6</td>
</tr>
<tr>
<td>$tR9$</td>
<td>28.7</td>
<td>34.9</td>
<td>36.3</td>
</tr>
<tr>
<td>$tR10$</td>
<td>18.5</td>
<td>66.5</td>
<td>15.0</td>
</tr>
<tr>
<td>$tR11$</td>
<td>9.4</td>
<td>56.6</td>
<td>34.0</td>
</tr>
<tr>
<td>$tR12$</td>
<td>36.9</td>
<td>19.2</td>
<td>43.9</td>
</tr>
</tbody>
</table>

We further analyzed how the character of the TS could influence the activation barrier, as presented in Figure 12. A quadratic relationship was observed between the percentage of the twistboat character in the TS and the activation energy for the set of reactions shown in Figure 12a. It can be deduced from this correlation that a twistboat character of 70–75% would ideally be suitable for lowering the activation barrier, and in this connection, the TS of $oR8$ has the most favorable conformation. For the remaining set of reactions, we could observe a quadratic relationship between an admixture of boat and twistboat characters in the TS and the activation energy for the set of reactions shown in Figure 12b. Higher admixture percentages contributed to lower activation barriers, with the TS character in $tR11$ being the most favorable.

CONCLUSIONS

In this work, we probed the mechanistic features of the hydrogen evolution reaction from water in the presence of small hydrides such as boranes and alanes, in the gas phase and in solution, by employing URVA complemented with the local mode analysis and the ring puckering analysis. This led to the following conclusions:

- For all reactions, except the reactions $oR3$ and $oR7$ and their two-water analogous reactions $tR8$ and $tR9$, a similar
The curvature pattern in the chemical phases was found consisting of three curvature peaks. These peaks correspond to the major chemical events occurring during reaction, i.e., O–H cleavage, H–H formation, B(Al)–H cleavage, and O–B(Al) formation. In all cases, the O–H cleavage precedes the other chemical events. Interesting to note, the B(Al)–H cleavage spans over a range of peaks in a cost-effective fashion, resembling the metal–H bond cleavage.

- The local mode force constants of the O–H, H–H, B(Al)–H, and O–B(Al) calculated for R and TS complement the curvature data. The O–H cleavage process is characterized by a relative change in k between the R and TS of more than 75%. The same holds for the B(Al)–H cleavage. Furthermore, there is a correlation between the activation energy and the change in the charge separation for the cleaving O–H and B(Al)–H bonds going from the R to TS. For reactions with a larger change in the charge separation for the O–H bond corresponding to lower activation energies, electron-withdrawing substituents on the catalyst will reduce the activation barrier. On the other hand, for reactions with a smaller change in the charge separation for the B(Al)–H bond associated with lower activation energy, electron-donating substituents on the catalyst will further reduce the activation barrier. This provides useful guidelines for the modification of these catalysts.

- We observe moderate-to-strong H–H bonds at the TS for reactions leading to H2 liberation. Most importantly, reactions without hydrogen release are characterized by a weak H–H bond (i.e., small k values) at the TS, thereby discarding any practical utilization of these catalysts for H2 production. Trends between k(H–H) at the TS and activation energy reflect TS stabilization via dihydrogen bond formation. The stepwise cleavage of the O–B(Al) bond is indicated by moderately strong O–B(Al) bonds at the TS. The O–B(Al) bond strength of the byproduct provides insights into the regenerability of the catalyst. Accordingly, oR8 and its two-water analogues tR10 and tR11 are the most promising candidates for catalyst regeneration.

- The driving force behind these reactions is the charge transfer between the catalyst and water. Oppositely charged H atoms of the catalyst and water participate in a dihydrogen interaction, which eventually leads to the release of a H2 molecule. Except in oR3 and oR7 and their two-water analogous reactions tR8 and tR9, the catalyst acts as a charge donor in the chemical phases. In reactions with no hydrogen release, the catalyst acts as a charge acceptor within the chemical phases. A molecular orbital analysis of oR3 further confirmed the inability of the catalyst to act as a charge donor. The donation from the B–H sigma orbital to the O–H sigma antibonding orbital is extremely low, as revealed by the orbital interaction energies. As examined for oR1PCM, the overall charge-transfer mechanism in solution remains unaltered; however, it leads to a larger charge separation in the RC compared to the gas phase due to the presence of the polar medium.

- Conformational changes also contribute to the reaction mechanism and the activation barrier. In most cases, the reactions proceed through an intermediate six-membered ring. The puckering analysis suggests that the latter is predominantly of boat and twistboat characters. The puckering character of the TS correlates with the activation barrier. Two subsets of reactions were identified with lower barriers, favoring either 65–75% of the twistboat character or a high percentage of the boat and twistboat admixtures.

Inspired by the new insights, we are currently investigating how to best optimize these catalysts for efficient hydrogen evolution.

# ASSOCIATED CONTENT

* Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpca.0c07244.

 Reaction schemes with original numbering, calculated bond lengths and local mode properties for selected bonds for all reactions; curvature decomposition into bond, angle, and dihedral components for all reactions; correlation analysis between the activation energy and change in charge separation for selected bonds for reactions tR1–tR12; local mode analysis of the B(Al)–H cleavage for reactions tR1–tR12; description of puckering coordinates; and coordinates of the stationary points of all reactions (PDF)
AUTHOR INFORMATION

Corresponding Author
Elfi Kraka − Computational and Theoretical Chemistry Group (CATCO), Department of Chemistry, Southern Methodist University, Dallas, Texas 75275-0314, United States; orcid.org/0000-0002-9658-5626; Email: ekraka@smu.edu

Authors
Sadisha Nanayakkara − Computational and Theoretical Chemistry Group (CATCO), Department of Chemistry, Southern Methodist University, Dallas, Texas 75275-0314, United States
Marek Freindorf − Computational and Theoretical Chemistry Group (CATCO), Department of Chemistry, Southern Methodist University, Dallas, Texas 75275-0314, United States
Yunwen Tao − Computational and Theoretical Chemistry Group (CATCO), Department of Chemistry, Southern Methodist University, Dallas, Texas 75275-0314, United States

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpca.0c07244

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (grant CHE 1464906). The authors thank Southern Methodist University for providing excellent computational resources.

REFERENCES


(38) Kraka, E.; Zou, W.; Tao, Y.; Freindorf, M. Exploring the Mechanism of Catalysis with the Unified Reaction Valley Approach (URVA) - A Review. Catalysts 2020, 10, No. 691.


Appendix E

Paper on PyVibMS

Y. Tao, W. Zou, S. Nanayakkara, and E. Kraka, PyVibMS: A PyMOL plugin for visualizing vibrations in molecules and solids,
Journal of Molecular Modeling (2020) 26:290
PyVibMS: a PyMOL plugin for visualizing vibrations in molecules and solids

Yunwen Tao1 · Wenli Zou2 · Sadisha Nanayakkara1 · Elfi Kraka1

Received: 30 June 2020 / Accepted: 4 August 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract
Visualizing vibrational motions calculated with different ab initio packages requires dedicated post-processing tools. Here, we present a PyMOL plugin called PyVibMS for visualizing the vibrational motions for both molecular and solid systems calculated by mainstream quantum chemical computer programs including Gaussian, Q–Chem, VASP, and CRYSTAL. Benefiting from the continuing development of the PyMOL platform, PyVibMS provides powerful functionalities and user-friendly interface. PyVibMS was written in Python and its open-source nature makes it flexible and sustainable. As an example, the motions of the Konkoli-Cremer local vibrational modes are shown in this work for the first time. PyVibMS is freely available at https://github.com/smutao/PyVibMS.

Keywords Vibrational spectroscopy · PyMOL · Phonon · Normal mode · Harmonic approximation · Local vibrational mode theory

Introduction
Vibrational spectroscopy is a fundamental analytical tool to study structural properties and dynamics of molecules and solids [1–4]. Infrared (IR) and Raman spectroscopy can be employed to determine the microscopic structure by assigning observed frequency peaks to the vibrations of specific functional groups [5, 6] or to characterize intermolecular forces, e.g., hydrogen bonding [7–9], in both molecular and solid systems [10–14].

With the rapid development in the field of computational chemistry, the theoretical prediction of vibrational spectra of molecular systems within the harmonic approximation based on analytic Hessian matrices is nowadays a routine procedure, which has been implemented in various quantum chemical packages [15, 16]. Extracting vibrational spectra for liquid and solution systems with time-correlation functions of velocities from molecular dynamics simulation trajectories is also on the rise [17–20]. Theoretical modeling of solids, e.g., crystals, although lagging behind the modeling of isolated molecules, has witnessed the development of several first-principles calculation packages (e.g., VASP, CASTEP, CP2K, Quantum ESPRESSO, WIEN2k, DMol3, and CRYSTAL) that allow the calculation of vibrational frequencies of solids [21, 22].

After the lattice vibrational modes and associated frequencies have been calculated for a crystal or other periodic system, it is necessary to visualize the atomic vibrational motions if assigning a vibrational frequency to a certain functional group is desired. When modeling the heterogeneous catalysis on metal surfaces, it would be helpful to visualize the vibrational mode with imaginary frequency in order to check whether the atomic movement in this mode connects the reactant and product geometries. However, only a few software packages have the functionality to visualize lattice vibrations in solids. One is the wxDragon package [23] which was designed mainly for visualizing the geometries calculated by various ab initio programs. The other is the XCrystDen program [24, 25], which supports the animation of vibration with given trajectory files as a workaround.

There exist also a few online tools that can display the vibration motions of solids. The Phonon website created by Miranda [26] takes one step further. It allows the user to visualize the vibrational motions besides Γ...
(q = 0) point by clicking at the points in a phonon dispersion plot. This online applet has been embedded into two material databases: Materials Project (https://materialsproject.org/) and Materials Cloud (https://www.materialscloud.org/). Recently, Noël designed a webpage [27] as part of the CRYSPLOT [28] suit for visualizing the solid vibrations calculated by CRYSTAL program. However, these online tools require the user to upload the output files and their closed-source nature prohibits the flexibility to meet different needs of the users.

In this work, we present a novel easy-to-use graphic user interface (GUI) named PyVibMS as a plugin for the PyMOL visualization system to visualize vibrations for both molecular and solid systems. PyVibMS is designed to visualize the geometry and vibrational modes for a number of ab initio packages, including Gaussian, Q-Chem, VASP, and CRYSTAL. It is written in Python language and is released as an open-source software, which makes PyVibMS easy to maintain and extend for more possibilities.

Methods

Overall architecture

PyMOL is written in the Python programming language [29, 30]. It offers an easy-to-use Python interface which allows users being familiar with Python to import PyMOL as a module, and to directly manipulate structural objects in the viewer window with the cmd functions in an external script. Based on the Python platform, PyMOL has access to a variety of modules and libraries that perform scientific computing, and it fosters many plugins developed by the scientific user community [31–68]. Starting from version 2.0 released in 2017, PyMOL replaces Tkinter with PyQt as its GUI engine and plugin developers could then harness PyQt’s more abundant widgets and design the plugin’s GUI in a straightforward way.

The overall architecture of the PyVibMS plugin created in this work is shown in Fig. 1. It contains a PyQt GUI window as is shown in Fig. 2. The plugin is compatible with PyMOL 2.x and works on all operating systems including Windows, Linux, and MacOS.

In the following subsections, the functionalities of PyVibMS are described in more detail.

Functionalities

View 3D structures of molecules and solids

PyVibMS reads the input geometry from up to five different file formats:

- XYZ file
- Gaussian 09/16 output file
- Q-Chem output file
- VASP 5.0 output with POSCAR/CONTCAR file
- CRYSTAL17 output file

The user can choose the desired file format from a drop-down list located in Section “Introduction” from the GUI window of PyVibMS. If the geometry has been obtained from the calculation with the program currently not supported by PyVibMS, one can extract the Cartesian
Fig. 2 The graphical user interface (GUI) window of the PyVibMS plugin. It is divided into four major sections according to functionalities, including (1) loading input geometry, (2) setting up periodic boundary conditions (PBC), (3) loading extra vibrational mode file, and (4) tweaking the visualization of vibrations.

Besides molecular structures, the PyVibMS plugin can also be used to show the 3D structure of periodic systems of one, two, and three dimensions. The structure of any periodic system can be described by the unit cell or primitive cell with translational symmetry [69], which is specified by lattice vectors \( \mathbf{a}_i \) \((i = 1, 2, 3)\). If the unit cell is translated along a lattice vector, it will completely overlap with the neighboring unit cell.

Importing solid structure into PyVibMS is a two-step process. (i) The geometry for the unit cell needs to be loaded as it can be done for molecular geometry. (ii) The user needs to enter the information about the lattice vectors in Section “Methods” of the GUI window (see Fig. 2). After selecting and confirming the dimensionality of the imported periodic system, the lattice vector information can be entered. A lattice vector has three components in \( x, y, \) and \( z \) directions separated by comma, and each component takes the unit of angstrom (Å) in PyVibMS. However, for VASP and CRYSTAL output files, the lattice information is automatically recognized by PyVibMS. Then if the Unit/Primitive Cell button is clicked, the lattice vectors will be shown in red, green, and blue colors for \( \mathbf{a}_1 \), \( \mathbf{a}_2 \), and \( \mathbf{a}_3 \) respectively. The structure of the supercell can be shown by clicking the Make Supercell button. Atoms within supercell adjacent to the central unit cell will be shown with smaller atomic radii as demonstrated in Fig. 3.

### Animate vibrations

Molecular vibrations or lattice vibrations in solids are frequently treated as harmonic oscillations about the equilibrium geometry. In PyVibMS, the animation of a vibration is realized by showing a series of interpolated displacements as a movie, being played with a constant frame rate. The trigonometric sine function is used within the interpolation in order to simulate the movement of the harmonic oscillator.

A vibrational mode displacement vector specifies the direction and amplitude of the movement for each atom. However, it can have different names or even definitions in the output files of various calculation packages, which can be confusing. In this work, we provide a short summary for clarification.

The harmonic normal vibrational modes and frequencies of a molecular system being composed of \( N \) atoms are obtained by solving the Wilson equation of vibrational spectroscopy [10]:

\[
\mathbf{f}' \mathbf{L} = \mathbf{M} \mathbf{L} \quad \text{(1)}
\]

where \( \mathbf{f}' \) is the Hessian matrix in \( 3N \times 3N \) dimension. \( \mathbf{M} \) is the mass matrix collecting the atomic masses for \( N \) atoms in three directions as its diagonal elements. The diagonal matrix \( \mathbf{L} \) with the dimension of \( N_{\text{vib}} \times N_{\text{vib}} \) collects \( N_{\text{vib}} \)
vibrational eigenvalues $\lambda_\mu$ ($\mu = 1, ..., N_{vib}$ with $N_{vib} = 3N - K$) where $N_{vib}$ is the total number of molecular vibrations and $K$ equals $5/6$ for linear/non-linear molecules. The harmonic vibrational frequencies $\omega_\mu$ (normally in the unit of cm$^{-1}$) are related to the eigenvalue $\lambda_\mu$ by:

$$\lambda_\mu = 4\pi^2 c^2 \omega_\mu^2$$  \hspace{1cm} (2)

where $c$ is the speed of light. Matrix $L$ with the dimension of $3N \times N_{vib}$ collects $N_{vib}$ column vectors $l_\mu$, which are renormalized and orthogonal to each other. As a result, one obtains $L^T M L = M^D$ and $L^T \Gamma L = K$, which defines the reduced mass matrix $M^D$ and the diagonal normal force constant matrix $K$ in normal coordinates, respectively. In Gaussian 09/16 and Q-Chem, the $l_\mu$ vectors collected in $L$, termed normal coordinates, are printed out as the atomic displacement for each vibration [16].

In a periodic system, i.e., solid, the vibrational analysis usually takes the ansatz of phonon spectroscopy. After the Hessian matrix $F$ of the unit cell has been obtained, a dynamical matrix (mass-weighted Hessian matrix) $D$ at $\Gamma$ point ($q = 0$) is calculated by mass-weighting [70, 71].

$$D_{\alpha l, \beta j}(q = 0) = \frac{f_{l j}}{\sqrt{m_\alpha m_\beta}}$$  \hspace{1cm} (3)

where $m_\alpha$ and $m_\beta$ are the masses for atoms $a$ and $b$ associated with the $i$-th and $j$-th Cartesian coordinates respectively. Noteworthy is that the term “unit cell” mentioned above arises from periodic boundary conditions (PBC) for calculation. This cell can be either a primitive cell or in order to calculate the vibrational motions leading to infrared (IR) and Raman spectra or a supercell used for calculating phonon dispersion relations [72, 73].

By diagonalizing matrix $D$ for the unit cell containing $N$ atoms, $3N$ eigenvalues $\lambda$ and their eigenvectors $c$ are obtained (collected in $A$ and $C$ respectively).

$$DC = CA$$  \hspace{1cm} (4)

For a 2D/3D periodic system, it is expected to have three zero eigenvalues, which correspond to the translation of the whole system in three directions. For a 1D periodic system, four zero eigenvalues are produced from diagonalization of $D$. The fourth zero eigenvalue is linked to the overall rotation of the system around the principal axis parallel to the lattice vector. The remaining $3N - 3/4$ vibrational frequencies $\omega$ can be calculated via Eq. 2. In VASP 5.x, the normalized eigenvectors $c$ are printed in the output file as the atomic displacements for both vibrations and translations/rotations. However, the CRYSTAL09/17 package takes a different path to obtain the atomic displacements for each vibration, while the vibrational frequencies are calculated in the same way as VASP 5.x.

In CRYSTAL09/17, the Hessian matrix $F'$ is first diagonalized:

$$F'C' = C'\Lambda'$$  \hspace{1cm} (5)

and each eigenvector $c'$ is then mass weighted by:

$$c_{ai}^{m,w.} = \frac{c'_{ai}}{m_a}$$  \hspace{1cm} (6)

where $m_a$ is the mass for atom $a$ with $i$-th Cartesian coordinates leading to a new vector $c_{ai}^{m,w.}$ whose norm (vector length) is denoted as $u_0$. Then, vector $c_{ai}^{m,w.}$ is normalized according to the equation of the time average kinetic energy of a phonon mode [74]:

$$\frac{1}{8}\rho V \omega^2 u_0^2 = \frac{1}{2} \left( \frac{1}{2} \hbar \omega \right)$$  \hspace{1cm} (7)

which can be re-written as:

$$\rho V \omega u_0^2 = 2\hbar$$  \hspace{1cm} (8)

where $\omega$ is the vibrational frequency, $\rho$ is the mass density, $V$ is the crystal volume, and $\rho V$ does mass-weighting (with $m_a$) on the square of each element in $c_{ai}^{m,w.}$. As the right-hand side of Eq. 8 is two times the Planck constant leading to a constant, vector $c_{ai}^{m,w.}$ can be then normalized according to a determined value of $u_0$ as the mode amplitude in Eq. 8 and thereby leading to a new vector as $c_{ai}^{m,w.}$. In the CRYSTAL09/17 output file, vectors $c_{ai}^{m,w.}$ are printed as normal modes normalized to classical amplitudes in the unit of Bohr [75].
Although different packages use different flavors for the definition of vibrational mode displacements, these differences are obscured when the displacement of the atoms is translated into the animation on the screen. Therefore, PyVibMS directly takes the displacement information from output files. There is no need for conversion into a consistent definition as explained above.

The user needs to select the Has Vib. Info. checkbox in the GUI window before the input geometry file is loaded if it also contains vibrational analysis information. After the vibration information is properly loaded, the vibrational frequency and symmetry (if available) for each vibration will be listed in the table region of the GUI window. Clicking a row in the table selects a target vibration. If the Start Animation button is then clicked, the movement of the selected vibration is shown in PyMOL’s viewer window.

In addition, PyVibMS provides two slider bars in the GUI window for the user to adjust the animation speed and displacement amplitude of vibrations.

**Show displacement vectors**

PyVibMS allows the user to visualize the vibrational displacement vectors as arrows by selecting the Displacement Vectors checkbox in the GUI window. The slider bar on the right-side scales the length of these arrows. We have used in PyVibMS a script written by Holder to show the arrow object in PyMOL [76].

If an arrow of the displacement vector overlaps with a chemical bond, the user can choose to unselect and then reselect the Displacement Vectors checkbox so that all displacement vector arrows will be shown in an opposite direction for better visualization.

**Visualize local vibrational modes**

In the past few years, our research group has made many contributions to the theory extension [77–82] and application [83–93] of the local vibrational mode theory originally proposed by Konkoli and Cremer in 1998 [94–100]. We have been employing extensively the local mode force constant $k_n^b$ and local mode frequency $\omega_n^b$ in our previous work, but we have not visualized so far how the local vibrational modes look like, although it has been proved that the Konkoli-Cremer local modes are the only local counterpart of normal vibrational modes [77, 88]. With PyVibMS, it is now possible to visualize how the atoms in molecules and solids move during a local vibration.

A local vibrational mode associated with an internal coordinate $q_n$ (e.g., bond length) in a molecular system is defined based on the leading parameter principle [94] which describes a process that this internal coordinate is first displaced infinitesimally followed by the relaxation of other parts of this molecule.

The solution of the Wilson equation (Eq. 1) leads to a diagonal Hessian matrix in normal coordinates $Q$:

$$ f^Q = K = L^T f^L $$

where $L^T$ is the transpose of $L$.

The internal coordinate $q_n$ leading a local vibrational mode can be defined via the Wilson $B$-matrix [10], connecting internal coordinate $q_n$ with Cartesian coordinates $x$ via partial derivatives:

$$ b_n = \frac{\partial q_n}{\partial x} $$

Row vector $b_n$ of length of $3N$ converts the $N_{vib}$ vibrational modes collected in $L$ from Cartesian coordinates into internal coordinates via:

$$ d_n = b_n L $$

Row vector $d_n$ of the length of $N_{vib}$ is then used to derive the local vibrational mode vector $a_n$ led by the internal coordinate $q_n$:

$$ a_n = \frac{K^{-1}d_n}{d_n^T K^{-1} d_n} $$

However, before visualizing local mode $a_n$ of dimension $N_{vib}$ with PyVibMS, it has to be transformed into Cartesian coordinates via:

$$ a_n^L = L a_n $$

so that $a_n^L$ has the proper length of $3N$. Besides, it can be easily proved that $b_n a_n^L = 1$.

In order to visualize a local vibrational mode, the file containing the mode vector $a_n^L$ needs to be first loaded with Section “Computational details” of the plugin’s GUI window. Then, the corresponding local vibrational modes will be added into the table region with corresponding local mode frequencies $\omega_n^b$. The format of the mode file storing the local vibrational mode vectors is elaborated in Appendix 2.

**Produce publication-quality images and movies**

In order to make publication ready figures with PyVibMS, e.g., the displacement vectors of a vibration, one can benefit from PyMOL’s powerful ray tracing function, which renders quality images for publication use.

To export the animation of a vibration as a movie file, it is recommended to use PyMOL’s Export Movie utility as long as necessary encoders are properly installed. Before recording the movie, the target vibration needs to be selected in the table region. According to our tests, a movie file
in (1) animated GIF or (2) QuickTime MOV format has satisfactory publication quality.

**Computational details**

The geometry of the carbon dioxide (CO$_2$) molecule was optimized using the B3LYP density functional [101] with Dunning’s aug-cc-pVTZ basis set [102, 103] in the Gaussian 16 package [104]. The density functional theory (DFT) calculation was conducted on a pruned (75, 302) grid [105]. The Hessian matrix was evaluated at the same level of theory. The local mode analysis was carried out with the program package COLOGNE2020 [106].

The monocrystalline silicon was modeled in the VASP 5.4.4 package [21, 107–110] using the Perdew-Burke-Ernzerhof (PBE) functional [111, 112] with projector-augmented wave (PAW) potential [113, 114]. The kinetic energy cutoff for basis set is 900 eV. A 12×12×12 k-point grid according to the Monkhorst-Pack scheme was employed to sample the Brillouin zone [115] as a primitive cell model was used in this work. A tight convergence criterion ($<10^{-8}$ eV/Å) was used for cell relaxation. The analytic Hessian matrix for the primitive cell was calculated with density-functional perturbation theory (DFPT) [116].

The two-dimensional (2D) water layer was calculated with the CRYSTAL17 program [117, 118] at the M06-2X/6-31+G(d,p) level [119–122]. In order for sufficient sampling of k-points in reciprocal space, a 12×12 grid was employed according to the Pack-Monkhorst method [115]. Besides, a pruned XXLGRID (99,1454) integration grid was used for DFT calculation. The Hessian matrix was evaluated in a semi-analytical approach by taking numerical derivatives of analytic gradients according to a central-difference formula with the stepsize of 0.001 Å [118].

**Results**

**Carbon dioxide molecule**

As a demonstration how a local vibrational mode looks like, the local vibration of a CO bond stretching in the carbon dioxide (CO$_2$) molecule is shown in Fig. 4.

According to the definition of internal coordinates in Eq. 10, the C1=O2 bond length is defined as the partial derivatives of the Cartesian coordinates for these two atoms in the X direction as shown in the components of the $b_n$ vector. However, the component for the third atom of the local vibrational mode vector $a_n^x$ is not zero. Instead, with the lengthening of the C1=O2 bond, the O3 atom has a displacement in the +X direction. This is because the local vibrational mode as defined with the leading parameter principle starts with an infinitesimal displacement of the C1=O2 bond and the O3 atom follows effortlessly.

**Monocrystalline silicon**

The monocrystalline silicon is an important player in materials science. It shares a similar lattice structure with diamond. In the primitive cell of monocrystalline silicon, there are two atoms connected with a Si–Si single bond and in total $3 \times 2 - 3 = 3$ vibrations, which are triply degenerate due to the unique diamond cubic lattice structure as shown in Fig. 5.

All three lattice vibrational modes in the monocrystalline silicon are orthogonal to each other. However, if the perspective is changed into the direction along the Si–Si
Fig. 5  Three lattice vibrations in a primitive cell of monocristalline silicon denoted as (I), (II), and (III) with the same vibrational frequency of 502 cm$^{-1}$. The cell with red, green, and blue edges is a primitive cell containing two silicon atoms with relatively larger size. For each vibration, the structure in two different perspectives are shown.

bond axis, one can see that the vibrational mode vectors can be projected onto a planar hexagon and the displacement vectors of two atoms point to a pair of opposite vertices.

**Water layer**

The 2D water layer is a model system which has been investigated in our previous work to determine the intrinsic strength of its hydrogen bonds [82]. As shown in Fig. 6, all water molecules in this layer have the same orientation and each water molecule donates and accepts two hydrogen bonds simultaneously.

The primitive cell in the 2D water layer can be chosen to contain one water molecule with three atoms. Therefore, in total, six lattice vibrations of the primitive cell are expected as shown in Fig. 7.
Fig. 7 Six lattice vibrations denoted as (I)∼(VI) in the 2D water layer with the vibrational frequencies as 435, 667, 750, 1642, 3718, and 3799 cm$^{-1}$, respectively.

These six lattice vibrations can be divided into two groups. The first group (I∼III) is basically the overall rotations of the water molecule. Different from molecular water in gas phase, each water molecule in the 2D layer forms hydrogen bonds with neighboring water molecules. Therefore, its rotation causes a change in the total energy, leading to non-zero vibrational frequencies. The three vibrations of the second group (IV∼VI) are the same as the internal vibrations of an isolated water molecule including the angle bending, symmetric O-H bond stretching, and asymmetric stretching. Noteworthy is that the asymmetric O-H stretching (VI) has also higher vibrational frequency than the symmetric O-H stretching (V) as in the case of an isolated water molecule.

Conclusions

PyVibMS is a freely available PyMOL plugin that can be obtained from GitHub (https://github.com/smutao/PyVibMS) and it works on major operating system platforms including Linux, Mac OS, and Windows.

With PyVibMS, one can visualize the vibrational modes of molecular and solid systems calculated by standard computational packages including Gaussian 09/16, VASP 5.x, and CRYSTAL09/17 in the PyMOL system. Benefiting from PyMOL’s powerful visualization capability, high-quality figures and movies can be easily generated.

We expect that PyVibMS will be a useful tool for all colleagues who are interested in the vibrational motions in molecules and solids.

Outlook

For future versions of PyVibMS, we plan to implement a functionality of identifying the symmetry (point group/space group and related irreducible representations) of the input geometry and the vibrational modes.

Besides, an interface to the phonopy package [123] is under development to visualize the phonons for other wave vectors $\mathbf{q}$ besides the $\Gamma$ point ($\mathbf{q} = \mathbf{0}$).

PyVibMS will also be interfaced to more ab initio/first-principles modeling packages. With the release of our local mode program LMODEA [100], users will be able to visualize local vibrational mode in both molecules and solids.

Acknowledgments We thank SMU for providing supercomputing resources. Y.T. thanks Yue Qiu and Xin Chen for valuable comments.

Funding This work was financially supported by National Science Foundation Grants CHE 1464906.

Appendix 1: Installation guide

The following instructions describe the installation of the latest version of PyVibMS:

1. Install the latest version of PyMOL 2.x from pre-compiled binaries or source code;
2. Download the latest version of PyVibMS from GitHub repository;
3. Open PyMOL and navigate to the Plugin Manager menu under the Plugin button;
4. On the Install New Plugin tab, select the \_init\_.py file in the PyVibMS folder after clicking the Choose file button. This step loads PyVibMS into PyMOL;
5. The PyVibMS item will be added to the Plugin menu if properly installed. Clicking PyVibMS opens its GUI window.

**Appendix 2: Format of the user-provided mode file**

If the user wants to visualize the local vibrational modes calculated by LMODEA program or molecular/lattice vibrations calculated by a different package outside the supported ones listed in this work, an additional text file can be read like the following.

```
L 3000.82834 0 CH_stretch
-0.00350461823
-0.071010202331
0.000003904593
0.00156831709
0.928990134727
0.000008762757
0.003646002972
-0.005846446492
0.00001286935
-0.000206076488
0.013893423928
0.000001632069
```

The 1st line contains two integer numbers: the first is the number of atoms \( N \) in a molecule or in a primitive/unit cell for solid systems, while the second number specifies the number of additional vibrations this text file has.

The 2nd and 16th lines are the blank lines before the information of each vibration. The 3rd and 17th lines are the header of each vibration, and each line has four fields. The first field takes either \( L \) or \( N \), representing local vibration and normal vibration respectively. The second field gives the vibrational frequency in wavenumbers. The last field takes a string which will show up in the comment column in the table section of the GUI window.

The 30th line of END following the displacement information of the last vibration denotes the end of this text file.

**References**


Zhao Y, Truhlar DG (2008) The m06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four m06-class functionals and 12 other functionals. Theor Chem Acc 120(1):215–241


Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Appendix F

Paper on GSVA

A revised formulation of the generalized subsystem vibrational analysis (GSVA)

Yunwen Tao1 · Wenli Zou2 · Sadisha Nanayakkara1 · Marek Freindorf1 · Elfi Kraka1

Received: 9 June 2020 / Accepted: 8 February 2021 © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

In this work, a simplified formulation of our recently developed generalized subsystem vibrational analysis (GSVA) for obtaining intrinsic fragmental vibrations (J Chem Theory Comput 14:2558, 2018) is presented. In contrast to the earlier implementation, which requires the explicit definition of a non-redundant set of internal coordinate parameters to be constructed for the subsystem, the new implementation circumvents this process by employing massless Eckart conditions to the subsystem fragment paired with a Gram–Schmidt orthogonalization to span the same internal vibration space indirectly. This revised version of GSVA (rev-GSVA) can be applied to equilibrium structure as well as transition state structure, and it has been incorporated into the open-source package UniMoVib (https://github.com/zorkzou/UniMoVib).

Keywords  Vibrational spectroscopy · Normal mode analysis · Subsystem · Harmonic approximation · GSVA · UniMoVib

Molecular vibrations described by normal modes are in general delocalized over the molecular structure of interest [1]. This delocalization hampers the direct comparison between the vibrations of a molecule in different environments (e.g., gas phase and solution) or when this molecule binds to a host compound. Over the past two decades, enormous effort has been put into finding “localized” normal vibrational modes (which can also be called “localized modals” in analogy to “localized orbitals” [2]), belonging to a molecule within a noncovalently bonded complex or a fragment of one molecule, including partial Hessian diagonalization [3], partial Hessian vibrational analysis (PHVA) [4, 5], mobile block Hessian (MBH) [6–10], vibrational subsystem analysis (VSA) [11–13], local Hessian transformation [14], just to name a few [15]. However, all these approaches share a common deficiency—the unphysical partitioning of the full Hessian matrix, which causes information loss about the interaction between the subsystem and its environment.

Recently, we proposed the generalized subsystem vibrational analysis (GSVA) as a new solution to obtain intrinsic fragmental vibrations [16, 17]. The key feature of GSVA compared to its predecessors lies in avoiding the partitioning of the full Hessian matrix. Instead, GSVA extracts for the subsystem a unique effective Hessian matrix $F^e_{\text{sub}}$

$$F^e_{\text{sub}} = B_{\text{sub}}^\dagger (B^\dagger F^\dagger B^\dagger + B_{\text{sub}}^\dagger)^{-1} B_{\text{sub}}$$

(1)

where $F^\dagger$ is the full Hessian matrix expressed in Cartesian coordinates of dimension $(3N \times 3N)$ for the whole molecular system being composed of $N$ atoms including $n$ atoms of the target subsystem and $N-n$ atoms of the environment. The Wilson B-matrices [1] $B^\dagger$ and $B_{\text{sub}}^\dagger$ define a non-redundant set of $(3n - k_{\text{sub}})$ internal coordinates for the subsystem fragment in rows with full $3N$ columns and truncated $3n$ columns (excluding the environment atoms), respectively. $k_{\text{sub}}$ is the total number of rotations and translations for the subsystem being 5 or 6 depending on whether the subsystem geometry is linear or nonlinear. $(F^\dagger)^\dagger$ is the Moore–Penrose inverse [18] of $F^\dagger$, which is singular. $F^\dagger_{\text{sub}}$ on the left-hand side is a symmetric matrix of dimension $(3n \times 3n)$, and it has exactly $k_{\text{sub}}$ zero eigenvalues. The $\dagger$ superscript denotes matrix transpose.

With the effective Hessian matrix $F^e_{\text{sub}}$ expressed in Cartesian coordinates, the conventional normal mode analysis (NMA) machinery [19, 20] which is widely implemented in...
most quantum chemical packages can be employed to calculate for the subsystem a new type of localized normal modes, which we coined \textit{intrinsic fragmental vibrations}. The reason why these normal vibrations are called “intrinsic” is due to the fact that the effective Hessian matrix $F^x_{\text{sub}}$ retains the curvature of the potential energy surface (PES) in the direction defined by any internal coordinate within the subsystem \cite{16, 17}. In other words, the subsystem fragment “feels” exactly the same curvature of the PES as the whole system being described with the full Hessian matrix $F^x$. This property of $F^x_{\text{sub}}$ endows our GSVA method with a solid physical basis \cite{16}.

As discussed in our earlier work on the original GSVA implementation \cite{16}, GSVA requires a complete and non-redundant set of internal coordinate parameters, and its Wilson-B matrices (see Eq. 1) to span the internal vibration space of the subsystem. However, the construction of the non-redundant parameter set is nontrivial and it needs either judicious selection of parameters mutually with expert knowledge, or a dedicated algorithm which automatically selects the non-redundant parameter set from a series of redundant set of parameters in a trial-and-error manner.

In this work, we propose an alternative formulation of GSVA, which can save the effort of constructing the non-redundant parameter set for the subsystem. The new implementation replaces the Wilson-B matrices in Eq. 1 with a different matrix, which also spans the internal vibrational space of the subsystem via the following procedure.

First, we apply to the subsystem fragment with its Cartesian coordinates collected in a $3n \times 1$ column vector $\mathbf{R}_{\text{cart}}$ the massless (assuming all atomic masses are identical) Eckart conditions \cite{21, 22} to generate a set of five or six translational and rotational vectors which are orthonormal to each other; see Eq. 2:

$$\mathbf{R}_{\text{tr, ro}} = \{ \mathbf{r}_1, ..., \mathbf{r}_i, ..., \mathbf{r}_k \} \quad (2)$$

where $k$ equals 5 or 6 depending on whether the subsystem is linear or not and $\mathbf{r}_1$ is a column vector of length $3n$.

Next, a Gram–Schmidt orthonormalization is conducted on $\mathbf{R}_{\text{tr, ro}}$ to generate $n_{\text{vib}} = 3n - k$ remaining vectors collected in $\mathbf{V}$

$$\mathbf{V} = \{ \mathbf{v}_1, ..., \mathbf{v}_j, ..., \mathbf{v}_{n_{\text{vib}}} \} \quad (3)$$

where $\mathbf{v}_j$ is a column vector of length $3n$. It has to be noted that matrix $\mathbf{V}$ is equivalent to matrix $\mathbf{B}^\dagger_{\text{sub}}$ in spanning the internal coordinate/vibration space of the subsystem. In order to obtain the equivalent matrix of $\mathbf{B}^\dagger$, we pad each column vector $\mathbf{v}_j$ with $3(N - n)$ zeros associated with environmental atoms, resulting in the matrix $\mathbf{V}_{\text{full}}$ with the dimension $(3N \times n_{\text{vib}})$.

In this way, the effective Hessian matrix $F^x_{\text{sub}}$ for the subsystem can be written as

$$F^x_{\text{sub}} = V(V^\dagger_{\text{full}}(F^x_{\text{full}})^+V_{\text{full}})^{-1}V^\dagger. \quad (4)$$

Then, the conventional NMA machinery is applied to obtain the intrinsic fragmental vibrations for the subsystem as in our earlier formulation \cite{16}. The new formulation has two major advantages for practical implementation. (1) It avoids the complicated process of finding the complete and non-redundant internal coordinate parameter set for the subsystem; (2) the code for finding translation/rotation vectors from the Eckart conditions and conducting Gram–Schmidt orthonormalization in a modern quantum chemical package can be reused, which facilitates implementing GSVA.

As a showcase example, we have employed rev-GSVA to calculate the intrinsic fragmental vibrations of the methane (CH$_4$) molecule in (1) methane-intercalated B$_{36}$N$_{36}$ complex (Fig. 1a), (2) methane-intercalated C$_{60}$ structure (Fig. 1b) \cite{23} and (3) gas phase as reference. Unlike the methane-intercalated C$_{60}$ complex, the methane-intercalated B$_{36}$N$_{36}$ system \cite{24} has not been synthesized experimentally so far, and it is interesting to compare the intrinsic fragmental vibrations of the methane molecule in B$_{36}$N$_{36}$ and C$_{60}$ in order to explore the different encapsulation effect. These three molecular systems were modeled at the M06-2X/6-31G(d,p) level \cite{25-27} with Grimme’s D3(0) dispersion correction \cite{28} using the Gaussian 16 package \cite{29}.

The results in Table 1 show that the methane molecules encapsulated inside the two cages retain $T_d$ symmetry, as the reference methane molecule in gas phase. The non-degenerate $A_1$ vibration describes the symmetric stretching of four C–H bonds. The doubly degenerate E modes specify the relative turnstile twisting motions of two H–C–H fragments. The triply degenerate $T_2$ modes (1–3) with lower frequencies denote the bending of methane, while the other triply degenerate $T_2$ modes (1’–3’) with higher frequencies are antisymmetric stretching motions of four C–H bonds. We found the

\textbf{Fig. 1} Structure of a methane molecule encapsulated in B$_{36}$N$_{36}$ cage with $T_d$ symmetry and b methane encapsulated in fullerene (C$_{60}$) with $T$ symmetry.
largest deviation in the vibrational frequencies relative to methane in gas phase for the $A_1$ mode and the $T_2(1'–3')$ modes which are concerned with the C–H bond stretching. For methane molecule contained in the $B_{36}N_{36}$ cage, above two vibrations are redshifted by 29 and 38 cm$^{-1}$ respectively compared to gas phase. However, these two vibrations of methane in fullerene are blueshifted by 70 and 45 cm$^{-1}$, respectively. This means the fullerene cage could strengthen the C–H bonds of methane, while the $B_{36}N_{36}$ weakens the C–H bonds of the contained methane molecule. One might argue that comparing the normal mode frequencies of the whole system (the way spectroscopists usually adopt) could lead to similar conclusion because the methane molecule is well separated from the cage structure; however, one needs to note that only the (rev-)GSVA method could provide the localized normal modes and frequencies which can be legitimately comparable across different systems containing the same target subsystem.

In addition to local minima on the PES, we have also tested a first-order saddle point (i.e., transition state, TS) structure. As it has been proven in our earlier work [16] that GSVA retains the curvature of the PES, it is of interest to explore whether (rev-)GSVA can retain the imaginary vibrational mode specifying the bond breaking/forming in the subsystem. In this pilot study, we investigated the TS of the chemical reaction involving a potential $\alpha$-ketoamide inhibitor of the SARS-CoV-2 main protease (Mpro), which is assumed to inhibit the activity of SARS-CoV-2 virus by blocking viral replication [30]. According to a recent X-ray structure of the $\alpha$-ketoamide SARS-CoV-2 main protease (Mpro) complex [30], ketoamide and enzyme are linked via a cysteine side chain of the enzyme. According to the suggested catalytic mechanism, the chemical reaction starts with a nucleophilic attack of the sulfur atom onto a C=O carbon atom of $\alpha$-ketoamide moiety, which is followed by proton transfer from the –SH group to a nearby oxygen atom of the inhibitor, as shown in Fig. 2. Work is in progress to model the reaction in the enzyme.

We started with a minimal subsystem of 3 atoms containing the proton and its donor/acceptor atoms, and rev-GSVA was applied to calculate the corresponding intrinsic fragmental vibrations. Surprisingly, no imaginary frequency exists for this small 3-atom subsystem. This is probably due to the fact that the reaction center should also include the carbon atom to draw a more complete picture about bond forming and breaking. However, one imaginary frequency starts to emerge when more surrounding atoms are included into the subsystem (see Table 2) and the imaginary frequency value quickly converges to that of the full system when the subsystem contains 20 atoms. This result indicates that if the normal vibration is localized in a particular part of the molecule (e.g., bond breaking/forming or C=O bond stretching), (rev-)GSVA

### Table 1

<table>
<thead>
<tr>
<th>Symmetry(No.)</th>
<th>$\text{CH}<em>4@B</em>{36}N_{36}$</th>
<th>$\text{CH}<em>4@C</em>{60}$</th>
<th>Gas phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$</td>
<td>3051</td>
<td>3150</td>
<td>3080</td>
</tr>
<tr>
<td>E(1)</td>
<td>1560</td>
<td>1583</td>
<td>1584</td>
</tr>
<tr>
<td>E(2)</td>
<td>1560</td>
<td>1583</td>
<td>1584</td>
</tr>
<tr>
<td>$T_2(1')$</td>
<td>3167</td>
<td>3250</td>
<td>3205</td>
</tr>
<tr>
<td>$T_2(2')$</td>
<td>3167</td>
<td>3250</td>
<td>3205</td>
</tr>
<tr>
<td>$T_2(3')$</td>
<td>3167</td>
<td>3250</td>
<td>3205</td>
</tr>
<tr>
<td>$T_2(1)$</td>
<td>1326</td>
<td>1328</td>
<td>1356</td>
</tr>
<tr>
<td>$T_2(2)$</td>
<td>1326</td>
<td>1328</td>
<td>1356</td>
</tr>
<tr>
<td>$T_2(3)$</td>
<td>1326</td>
<td>1328</td>
<td>1356</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Subsystem size (n)</th>
<th>Imaginary frequency ($\omega_i$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>n/a</td>
</tr>
<tr>
<td>5</td>
<td>$-789$</td>
</tr>
<tr>
<td>8</td>
<td>$-778$</td>
</tr>
<tr>
<td>15</td>
<td>$-729$</td>
</tr>
<tr>
<td>20</td>
<td>$-727$</td>
</tr>
<tr>
<td>Full system</td>
<td>$-725$</td>
</tr>
</tbody>
</table>

Fig. 2 Transition-state structure of proton transfer from methanethiol to an $\alpha$-ketoamide inhibitor [30]. The methanethiol group is a simplified model of cysteine in SARS-CoV-2 main protease. The minimal 3-atom subsystem is highlighted in green color. The 5-atom subsystem is highlighted with green and blue. The 8-atom subsystem includes the 5-atom subsystem and atoms highlighted with orange. The 15-atom subsystem includes the 8-atom subsystem and atoms highlighted in purple. The 20-atom subsystem includes the 15-atom subsystem and atoms highlighted with cyan. This system was modeled at B3LYP/6-31G(d,p) level in Gaussian 16.
is expected to reproduce this vibration using a subsystem containing this segment and a few surrounding atoms. This valuable feature of (rev-)GSVA can lead to important first insights into the role of surrounding atoms for the reaction mechanism by analyzing the TS, before starting a more complex reaction path following procedure.

We have implemented the new formulation of GSVA (rev-GSVA) introduced in this work into the open-source package UniMoVib [20] for interested readers to use in their own research.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00214-021-02727-y.

Acknowledgements This work was financially supported by the National Science Foundation Grants CHE 1464906. W. Z. was supported by the National Natural Science Foundation of China (Grant No. 21673175) and the Double First-Class University Construction Project of Northwest University. We thank SMU for providing supercomputing resources.

References

Appendix G

Paper on H⋯H interactions

A. Hazrah, S. Nanayakkara, N. Seifert, E. Kraka and W. Jáger, Non-covalent interactions and structural study of 1- and 2- naphthol: insights into H-H interactions,
Under review in Physical Chemistry Chemical Physics
Non-covalent Interactions and Structural Study of 1- and 2-Naphthol: Insights into H-H Interactions

Arsh H. Hazrah, Sadisha Nanayakkara, Nathan A. Seifert, Elfi Kraka, and Wolfgang Jäger

The microwave spectra of the 1-naphthol and 2-naphthol monomers were measured using a chirped-pulse Fourier transform microwave spectrometer in the 2-6 GHz range. Spectroscopic assignments of the cis- and trans-conformers of 1-naphthol and 2-naphthol are supported by results from electronic structure calculations. Spectra of ten 13C isotopologues were identified and analyzed for all four species, which enabled the determination of substitution structures and so-called semi-experimental effective structures. The theoretical results show that the OH group of cis-1-naphthol points ~6° out of plane, which is consistent with the difference in inertial defect between cis- and trans-1-naphthol. For cis-1-naphthol, NCI and QTAIM analyses resulted in a bond critical point between the hydroxyl H-atom and the close H-atom at the neighbouring ring, suggesting a bonding interaction. To confirm or deny this hypothesis, local mode analyses were carried out to obtain a more in-depth assessment of the O-H bond strengths. Compared to trans-1-naphthol, the strength of the O-H bond is much stronger in cis-1-naphthol, reflected in a blue shift of the O-H stretching frequency, which is consistent with the traditional notion of a steric repulsion rather than a bonding interaction.

Introduction

1- and 2-naphthol, hydroxy-derivatives of naphthalene, the simplest polycyclic aromatic hydrocarbon, are used as precursors in the dye, perfume, insecticide, and pharmaceutical industries and are metabolites of naphthalene itself, which has uses as pesticide. The naphthols can also be considered as naphthalene homologues of phenol and are, from an intermolecular interactions point of view, of interest because they contain a hydrophobic part, i.e. the bicycle with an extended π-electron system, and an OH group. Indeed, a number of weakly bound complexes involving the naphthols have been studied using spectroscopic techniques. The Leutwyler group in particular has studied 1-naphthol – (H2O)N (N=1 to 50) and 2-naphthol – (NH3)N (N=1-10) complexes and clusters with laser spectroscopic techniques in an effort to determine the solvation threshold for excited state proton transfer from naphthol to water, 1-naphthol – alkane complexes. 1-naphthol – rare gas and N2 complexes, hydrogen-bonded complexes of naphthol, and most recently complexes of 1-naphthol with linear molecules. Naphthol – water clusters have also been studied by the Fujii group, by Knochenmuss and Smith, and by Pratt et al. The 1-naphthol dimer has been studied by infrared dip spectroscopy and more recently the microwave spectrum of the 1-naphthol dimer has been measured and assigned and interpreted, with the help of theoretical calculations, in terms of a structure that is dominated by π-π stacking interactions over canonical hydrogen bonding.

There are a number of earlier spectroscopic studies of the naphthol monomers. Pratt’s group has studied the fluorescence spectra of 1- and 2-naphthol and established the existence of cis- and trans-conformers for both monomers. This was followed by a microwave spectroscopic study of cis- and trans-1-naphthol by Brown and co-workers, who identified a close contact between the hydroxyl H and the neighbouring ring H-atom. Saeki et al. reported then on infrared dip spectra of trans- and cis-1-naphthol. Recently, Goubet et al. measured rotational and vibrational spectra of cis-2-naphthol and re-examined the spectra of trans-1-naphthol.

Here, we describe the rotational spectrum of trans 2-naphthol measured with a chirped-pulse Fourier transform microwave spectrometer in the 2 to 6 GHz range. Spectra of all singly substituted 13C isotopologues were also recorded, in addition to those of cis and trans 1-naphthol and cis 2-naphthol. The rich isotopic data were used to derive heavy atom Kraitchman substitution coordinates and semi-experimental structures for both conformers of both isomers. The close H-H contact in cis 1-naphthol is a main focus of this study. A Quantum Theory of Atoms-in-Molecules (QTAIM) analysis
revealed a bond critical point between the two hydrogen atoms and a Non-Covalent Interactions (NCI) analysis resulted in a reduced electron density gradient isosurface with positive sign of the second Hessian eigenvalue, which is often an indication for a bonding interaction. We applied several approaches to provide further insights into the weak intramolecular interactions within each conformer. The local vibrational mode theory originally introduced by Konkoli, Cremer et al. was utilized to assess and compare the intrinsic strength of the OH bond in both isomers of 1- and 2-naphthol. We also used Charge Model 5 (CMS) and Natural Bond Orbitals (NBO) approaches and determined that there is an attractive component to the H-H interaction, but that the steric repulsive component outweighs it.

Materials and Methods

Theoretical Methods

Utilizing density functional theory (DFT) and the Gaussian 16 program suite geometry optimizations and harmonic frequency calculations were performed at the B3LYP-D3(BJ)\textsuperscript{[13, 34, 35]} level of theory with the def2-TZVP basis set\textsuperscript{36} and at the Møller-Plesset second order perturbation theory (MP2)\textsuperscript{[37]} level with the aug-cc-pVTZ basis set. Analyses of the molecular electron density distribution based on Bader’s quantum theory of atoms in molecules (QTAIM)\textsuperscript{38} including non-covalent interactions (NCI) analyses were done using the AIMAll\textsuperscript{[39]} and MultiWfn\textsuperscript{[40]} programs, then visualized using UCSF Chimera.\textsuperscript{41}

Normal vibrational modes are generally delocalized as a result of mass coupling.\textsuperscript{42-45} This implies that if one considers a particular normal stretching mode between two atoms of interest, it can be coupled to other normal modes such as bending or torsion, which hampers the direct correlation between stretching frequency and bond strength as well as the comparison between stretching modes of similar nature. As a consequence, the normal stretching force constant cannot be used as a direct bond strength measure, which results in the need to derive a local counterpart that is free from any mode-mode coupling. Konkoli, Cremer et al.\textsuperscript{[26-30]} approached this problem by solving the mass-decoupled analogue of Wilson’s equation of vibrational spectroscopy leading to local vibrational modes, associated local mode frequencies, and local mode force constants. Zou and Cremer showed that the local stretching force constant reflects the curvature of the PES in the direction of the bond stretching.\textsuperscript{46} This important result qualifies the local stretching force constants $k_n$ as a unique quantitative measure of the intrinsic strength of a chemical bond and/or weak chemical interaction based on vibrational spectroscopy, which has been extensively applied in previous work.\textsuperscript{25} (For some recent work see also Refs. \textsuperscript{47-50}.) Another important feature of the local vibrational mode theory is that any complete set of non-redundant local modes can be transformed into the corresponding set of normal modes via an adiabatic connection scheme.\textsuperscript{51} This unique one-to-one correspondence has led to a new comprehensive analysis of infrared/Raman spectra via the characterization of normal modes (CNM) procedure\textsuperscript{[28, 52]} which allows the decomposition of each normal mode into its respective local mode counterparts, thereby facilitating the identification of their individual contributions. The local mode analysis (LMA) was applied in this work to assess the strength of the O-H bonds in cis- and trans-1- and 2-naphthol and to evaluate the local character of the corresponding O-H normal modes utilizing the LModeA program package.\textsuperscript{25} For convenience, $k_n$ was converted into more chemically intuitive bond strength order (BSO) values by utilizing an extended Badger rule.\textsuperscript{29, 53} The latter relates BSO $\pi$ to $k_n$ via a power relationship, which is fully determined based on two well-known references and the requirement that for a zero-force constant, the corresponding BSO $\pi$ is zero:

$$\text{BSO } \pi = a(k_n)^b$$ (1)

Here, the constants $a$ and $b$ were evaluated using F-H and [F-H-F] as the references, which have BSO $\pi$ values 1 and 0.5, respectively, which resulted in $a=0.518$ (0.490) and $b=0.291$ (0.319) (values for MP2/aug-cc-pVTZ calculations are shown first, followed by those of B3LYP-D3(BJ)/def2-TZVP calculations in parentheses). The corresponding $k_n$ values for F-H and [F-H-F] are 9.584 (9.367) mDyn/Å and 0.884 (1.064) mDyn/Å, respectively. Since we are interested in O-H bonds, we scaled the BSO $\pi$ values obtained via equation (1), according to which the O-H bond in H$_2$O has BSO $\pi$ values of 0.961 (0.957) by a factor of 1.041 (1.045), so that BSO $\pi$ OH in H$_2$O is 1.
Experimental Methods

Rotational spectra of the naphthols were recorded with a pulsed nozzle chirped-pulse Fourier transform spectrometer in the frequency range between 2 and 6 GHz. Our instrument operates analogously to the spectrometer designed by the Pate group, with some variances in component specifications. The 2-6 GHz frequency region is well suited to study 1- and 2-naphthol as they are relatively large molecules and many low J rotational transitions fall into this range. Per molecular pulse, we recorded six free induction decays (FIDs). About 3.8 M (850 k) FIDs were averaged and then Fourier transformed to generate the broadband rotational spectrum for 1-naphthol (2-naphthol). 3.8 M averages were necessary to measure $^{13}$C transition for the cis-1-naphthol conformer with sufficient signal to noise ratio.

The 1- and 2-naphthol (≥ 99 %) samples were purchased from Millipore-Sigma and used without any further modification or purification. Both 1- and 2-naphthol are solids with melting points of about 95 and 120 °C, respectively. To generate the vapor pressure needed to bring a sufficient number of molecules into the gas phase, we utilized a special attachment to a General Valve, Series 9, pulsed valve which contains a sample reservoir and can be heated. For 1-naphthol the attachment was heated to 100 °C and for 2-naphthol to 120 °C. Neon (helium) was used as backing gas at pressures of about 3 atm for 1-naphthol (2-naphthol). Neon backing gas improved the signal to noise ratio for 1-naphthol and was used to measure its $^{13}$C transitions.

Results and Discussion

Experimental Results

Sections of the broadband rotational spectra for 1-naphthol and 2-naphthol are shown in Figure 1. The rotational and centrifugal distortion constants from Whitham et al. were used to locate transitions of cis- and trans-1-naphthol. For cis-trans-1-naphthol, additional strong a-type and b-type transitions were recorded. Transitions of the cis-isomer are weaker by about 11% than those of trans-1-naphthol, consistent with their theoretical relative energies (vide infra). The transition frequencies are given in Tables S1 and S2 of the electronic supplementary information (ESI) together with the quantum number assignments. The frequencies were used in a fitting procedure using the SPCAT/SPFIT program suite and Watson’s A-reduction Hamiltonian in its $^{13}$C representation to determine experimental rotational and quartic centrifugal distortion constants, which are given in Table 1. For cis-2-naphthol, the spectroscopic constants of Goubet et al. were used and for trans-2-naphthol, the assignment procedure was based on the theoretical (B3LYP-D3(BJ)/def2-TZVP, see below) rotational constants and dipole moment components in Table 1. We were able to measure 14 (1) weaker a-type and 27 (35) strong b-type transitions for cis-trans-2-naphthol. In this case, transitions of cis- and trans-isomers have similar strengths, also consistent with the theoretical energies and dipole moment components. The transition frequencies with the associated quantum numbers are in Tables S3 and S4 of the ESI and the resulting rotational and centrifugal distortion constants are in Table 1. The transitions were strong enough that all ten singly substituted $^{13}$C isotopologues for both conformers of 1- and 2-naphthol could be detected in their natural abundances. The transition frequencies are in Tables S5 to S8 of the ESI and the corresponding spectroscopic constants are given in Tables S9 to S12 of the ESI.

Theoretical Results

Geometry optimizations and harmonic frequency calculations were done at the B3LYP-D3(BJ)/def2-TZVP and MP2/aug-cc-pVTZ level for both trans- and cis-conformers of 1- and 2-naphthol. The resulting structural parameters are given in Tables S13 to S16 of the ESI and the corresponding rotational constants, centrifugal distortion constants, and dipole moment components are in Table 1; the structures are shown in Figure 2. The dihedral angle $\tau$(C2-C1-O1-H1) for 1-naphthol and $\tau$(C1-C2-O1-H2) for 2-naphthol were scanned to obtain potential energy curves (Figure 2) that connect the trans- to the cis-isomers via barriers of 2.4 kcal/mol (1-naphthol) and 3.1 kcal/mol (2-naphthol). At the B3LYP-D3(BJ)/def2-TZVP level of theory, the relative zero-point energy differences between the trans- and cis-isomers are 0.8 kcal/mol (1-naphthol) and 0.5 kcal/mol (2-naphthol). In 1-naphthol, the trans-conformer is...
This apparent bonding interaction between two close contact H atoms in pure hydrocarbons, such as phenanthrene, where the two close contact H-atoms are connected by a bond path with bond critical point. This apparent bonding interaction between two equally or similarly charged hydrogen atoms was termed hydrogen-hydrogen, or H-H, bonding, in contrast to dihydrogen bonding, where the two interacting H-atoms have a charge difference of about 1 e or more. Bader and coworkers have considered atomic energies within the framework of QTAIM to determine atomic energy differences within the framework of QTAIM and to study the energetics of hydrogen-hydrogen bonding.

### Table 1. Spectroscopic parameters for the naphthol isomers from theory and from fits of experimental data.

<table>
<thead>
<tr>
<th></th>
<th>cis-1-naphthol</th>
<th>trans-1-naphthol</th>
<th>cis-1-naphthol</th>
<th>trans-1-naphthol</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(B3LYP-D3(BJ)/def2-TZVP)</td>
<td>(B3LYP-D3(BJ)/def2-TZVP)</td>
<td>(experiment)</td>
<td>(experiment)</td>
</tr>
<tr>
<td>ΔE / kcal mol⁻¹</td>
<td>0.05262</td>
<td>0.05331</td>
<td>[0.0136]</td>
<td>0.0170(61)</td>
</tr>
<tr>
<td>ΔΔ / kHz</td>
<td>-0.08807</td>
<td>-0.09152</td>
<td>[0.0409]</td>
<td>0.021(14)</td>
</tr>
<tr>
<td>δ / amu Å</td>
<td>0.04032</td>
<td>0.04306</td>
<td>[0.006]</td>
<td>0.070(45)</td>
</tr>
<tr>
<td>N</td>
<td>0.01201</td>
<td>0.01177</td>
<td>[-0.006]</td>
<td>0.009(12)</td>
</tr>
<tr>
<td>σ / kHz</td>
<td>-0.01278</td>
<td>-0.01482</td>
<td>[-0.0018]</td>
<td>0.023(12)</td>
</tr>
<tr>
<td>N</td>
<td>0.0</td>
<td>0.0</td>
<td>26</td>
<td>63</td>
</tr>
<tr>
<td>ΔE / kcal mol⁻¹</td>
<td>0.8</td>
<td>0.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Δ / amu Å</td>
<td>-0.0443</td>
<td>0.0001</td>
<td>-0.2925(2)</td>
<td>-0.2124(2)</td>
</tr>
<tr>
<td></td>
<td>µ</td>
<td>/ D</td>
<td>µ=1.4, µ=0.3, µ=0.1</td>
<td>µ=1.2, µ=0.1, µ=0.0</td>
</tr>
</tbody>
</table>

### Figure 3. Results from QTAIM analyses for the four experimentally observed isomers of naphthol. The green spheres denote bond critical points, while the red spheres denote ring critical points.

lower in energy, while in 2-naphthol the cis-conformer is the lower energy conformer. This ordering is consistent with previous literature determinations. To gain insights into the intramolecular interactions in the naphthol monomers, we performed QTAIM (Figure 3) and NCI analyses, Figure S1, ESI) of the molecular electron density distributions. Interestingly, we found, apart from the expected ring critical points near the centres of the aromatic rings, a bond path with bond critical point between the hydroxyl H-atom and the neighbouring H-atom at the adjacent ring in cis-1-naphthol (see Figure 3). This appears to be indicating a bonding interaction between those two H-atoms, in stark contrast with the traditional notion of a steric repulsion at a separation of only 1.9 Å, much shorter than the sum of the van der Waals radii of ~2.4 Å. Accompanying the H-H bond critical point is a ring critical point that is associated with the 6-ring formed by the H-H interaction. One can follow the evolution of these two critical points along the cis-trans conversion coordinate and finds that they coalesce and disappear if the dihedral angle τ becomes greater than 21°. Similar effects have been found by the Bader group in pure hydrocarbons, such as phenanthrene, where the two close contact H-atoms are connected by a bond path with bond critical point. This apparent bonding interaction between two equally or similarly charged hydrogen atoms was termed hydrogen-hydrogen, or H-H, bonding, in contrast to dihydrogen bonding, where the two interacting H-atoms have a charge difference of about 1 e or more. Bader and coworkers have considered atomic energies within the framework of QTAIM to determine atomic energy differences within the framework of QTAIM and to study the energetics of hydrogen-hydrogen bonding.

---

* fixed at the values from Whitham et al.²¹

rationalize the existence of H-H bonding. The atomic energies are obtained by partitioning the molecular kinetic electronic energy among the atomic basins whose extents are defined by the topology of the electron density. Bader and co-workers note that the presence of the H-H bond path is accompanied by a lowering of the collective H-atom energy, in the case of phenanthrene versus its linear isomer anthracene, for example. The concurrent increase in carbon skeleton energy is less in magnitude, such that overall phenanthrene is lower in energy than anthracene.

In some of the hydrocarbons considered by Bader and co-workers, the H-atoms are forced into close contact by the rigid structure of the C-atom skeletons. The case of cis-1-naphthol is interesting because of the flexibility in the OH-group orientation; the close H-H contact is adopted ‘voluntarily’. We carried out analyses analogous to those by Bader and co-workers for cis- and trans-1-naphthol using the AIMA11 program. We find that the atomic energy, \( E_\alpha \), of the close-contact C-bonded H-atoms in cis-1-naphthol is lower by 11.6 kcal/mol compared to the average energy of all other C-bonded H-atoms. In contrast, the same H-atom is higher in energy by 5.3 kcal/mol in trans-1-naphthol. The total H-atom energy is lower by 7.0 kcal/mol in cis-1-naphthol and the total heavy atom energy is higher by 8.1 kcal/mol than in trans-1-naphthol. In total, trans-1-naphthol is more stable by 1.1 kcal/mol, consistent with the theoretical calculations. A compilation of relevant data from the AIMA11 calculations is in Table S17 of the ESI.

**Structures**

The experimental and theoretical[B3LYP-D3(BJ)/def2-TZVP level of theory] calculated rotational constants for cis- and trans-isomers of both 1- and 2-naphthol in Table 1 differ by at most 0.8 %, which is indicative of reasonable agreement between theoretical structural parameters and those underlying the experimental rotational constants. For a meaningful comparison, however, one needs to take into account the effects of zero-point vibrational motions that cause the difference between equilibrium \((\alpha_\text{eq})\) structure and equilibrium \((\alpha_\text{eq})\) rotational constants on one hand and effective \((\alpha_\text{eff})\) structure and ground state \((\alpha_0, B_0, C_0)\) rotational constants. The anharmonic vibrational contributions to the rotational constants were calculated using vibrational perturbation theory (VPT2) at the B3LYP-D3(BJ)/def2-TZVP level and subtracted from the experimental constants to arrive at the so-called semi-experimental rotational constants, given in Table 2 and Tables S18 to S22 in the ESI. The largest difference to the theoretical equilibrium constants is now less than 0.1 %.

To further assess the quality of the semi-experimental rotational constants, one may look at the inertial defects, defined as \( \Delta_\text{I} = I_\text{E} - (I_\alpha + I_\beta) \). For a planar rigid molecule \( \Delta_\text{I} \) is zero. The inertial defects for the naphthol isomers (Table 1) range from -0.212 to -0.293 amu Å\(^2\) and are consistent with bicyclic heterocycles whose \( \Delta_\text{I} \) values have been discussed in detail in Ref. 59. The negative values are a result of greater contributions by out-of-plane zero-point vibrational motions to the inertial defect compared to the in-plane vibrations. The cis-1-naphthol isomer has the largest magnitude \( \Delta_\text{I} \) value. The difference in \( \Delta_\text{I} \) between cis- and trans-1-naphthol is -0.080 amu Å\(^2\), in reasonable agreement with the theoretical value of \( \Delta_\text{I} \) for cis-1-naphthol (-0.047 amu Å\(^2\)). The non-zero theoretical value of \( \Delta_\text{I} \) for cis-1-naphthol is attributable to the out-of-plane location of the OH H-atom with a dihedral angle \( \tau \) of 6°. The experimental difference between \( \Delta_\text{I} \) values of 0.080 amu Å\(^2\) corresponds to a dihedral angle \( \tau \) of 10°. The \( \Delta_\text{I} \) values calculated from the semi-experimental constants are reduced by about 80% and have magnitudes on the order of 0.05 amu Å\(^2\) or smaller. It is interesting to note that there is an overcorrection for \( \Delta_\text{I} \) in cis-1-naphthol compared to the other isomers. This leads to positive semi-experimental \( \Delta_\text{I} \) values that are not consistent anymore with the out-of-plane OH H-atom in cis-1-naphthol and the other planar naphthol isomers. It is likely that the large amplitude motion of the OH H-atom from above to below the heavy atom plane in cis-1-naphthol is not captured properly in the anharmonic calculations, thus leading to the inconsistency.

Additionally, we compare our inertial defect values to that of Jahn et al.\(^{59}\) to examine how naphthol compares to other heterocyclic molecules. A plot showing the difference between the experimental inertial defect and calculated inertial defect are shown in Figure 4 with the corresponding fit shown in equation 2 of Jahn et al. The calculated values for the four isomers of naphthol were computed by summing Oka’s equation\(^{60}\) over the lowest out-of-plane vibrations. As outlined by Jahn et al. the number of out-of-plane modes used in the sum is equal to the number of rings present, or for molecules with an extra low wavenumber out-of-plane vibration (<100 cm\(^{-1}\)) is equal to the number of rings +1. For trans 1-naphthol and both isomers of 2-naphthol the lowest two out-of-plane vibrations were used, while the three lowest out-of-plane vibrations were used for cis 1-naphthol due to the existence of an extra low wavenumber out-of-plane vibration. The vibrational modes for each isomer were obtained from the B3LYP-D3(BJ)/def2-TZVP results. The calculated and experimental inertial defect values and the difference between the two are shown in Table S23 of the ESI. For cis 1-naphthol there is a difference of ~ 0.49 amu Å\(^2\), while the differences for trans 1-naphthol and both isomers of 2-naphthol range from 0.21-0.23 amu Å\(^2\). Comparing these results to the fit in Figure 4 in Jahn et al. there is a discrepancy in the \( I_{\text{c}} \) values determined from the plot using the inertial defect difference and the experimental \( I_{\text{c}} \) values. For trans 1-naphthol the fit slightly underestimates the experimental \( I_{\text{c}} \) value (~26.6 amu Å\(^2\)) by several amu Å\(^2\). The fit also underestimates the experimental \( I_{\text{c}} \) values for both isomers of 2-naphthol, where the averaged 28.1 amu Å\(^2\) values does not correspond to an average inertial defect difference of 0.22 amu Å\(^2\). The largest discrepancy between the fit data and our results are for cis 1-naphthol where the inertial difference of 0.49 amu Å\(^2\) severely overestimates the \( I_{\text{c}} \) value, the experimental value is 26.6 amu Å\(^2\). Even using the two lowest out-of-plane vibration instead of three the difference (0.32amu Å\(^2\)) still overestimates the value of \( I_{\text{c}} \). One potential reason for the discrepancy may be the molecules selected by the
authors to construct the fit, which contain a range of sizes of heterocyclic compounds, but do not have any conformational flexibility. Although the skeleton structure of naphthol itself is not very flexible it does have the ability to interconvert between two conformers. This added flexibility may be a factor in the discrepancy between the fit in results reported herein.

The semi-experimental rotational constants of normal and singly substituted $^{13}$C isotopologues of both conformers of 1- and 2-naphthol were used in a fitting procedure to produce the structural parameters in Tables S13 to S16 of the ESI. Comparison with the theoretical structure in same tables shows, in general, very good agreement with the differences on the order of 0.1 Å and 1° for bond lengths and angles, respectively.

Another way to minimize the effect of zero-point vibrational motions on structural parameters that is purely based on experimental data is a Kraitchman substitution analysis. Since rotational constants are available for all singly substituted $^{13}$C isotopologues, we were able to produce substitution coordinates for all carbon atoms of cis- and trans- 1- and 2-naphthol. The resulting substitution, $r_n$, structural parameters are also in Tables S13 to S16 of the ESI and are in excellent agreement with the theoretical and semi-experimental values. The largest differences are on the order of 0.03 Å and 2° for bond lengths and angles, respectively.

**Conformer conversion pathways**

![Figure 4](image-url)

**Figure 4.** Power relationship between bond strength order (BSO) $n$ and $k^a$ of O-H bonds in cis- and trans- 1- and 2-naphthol (in purple) and reference molecules (in red) for B3LYP-D3(BJ) (orange line) and MP2 (blue line) calculations. (BSO) $n$ has been scaled by a factor of 1.045 and 1.041 for B3LYP-D3(BJ) and MP2 calculations, respectively.

naphthol > cis-2-naphthol. This confirms that there is a repulsive H-H interaction in cis-1-naphthol and refutes the possibility of a bonding interaction as one would speculative based on the bond critical point between the H-atoms. This observed trend in the O-H bond strengths for the naphthol compounds can also be visualized as a blue shift in the IR spectrum provided that the OH normal mode is not coupling with the other modes, e.g. bending modes. We investigated this via the CNM procedure for cis and trans conformers of 1-naphthol calculated at the B3LYP-D3(BJ) level of theory. In Figures S2 and S3 of the ESI, the decomposition plots for all normal modes (51 modes) into a non-redundant set of 51 local modes, for cis- and trans-conformers of 1-naphthol, are shown where the OH normal mode is highlighted in yellow. It is observed that the OH normal stretching mode (3817 cm$^{-1}$) in the cis conformer has 99.9% contribution from O-H local stretching mode with up to 0.06% contributions from the two O-C-C local bending modes. Also, in the trans conformer, the O-H local stretching mode dominates with 99.93% contribution to the OH normal mode (3804 cm$^{-1}$).

As the O-H normal modes in both conformers possess the character of an almost pure O-H stretching mode, one can directly compare the corresponding O-H normal modes. Thus,

**Table 2.** Experimental, semi-experimental, and theoretical rotational constants of cis-1-naphthol.

<table>
<thead>
<tr>
<th>Constant</th>
<th>Experiment</th>
<th>Semi-Experimental</th>
<th>B3LYP-D3(BJ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B / MHz</td>
<td>1124.30739</td>
<td>1130.679</td>
<td>1130.8340</td>
</tr>
<tr>
<td>C / MHz</td>
<td>713.09734</td>
<td>717.334</td>
<td>717.128</td>
</tr>
</tbody>
</table>

Local Mode Analysis

The power relationship between BSO $n$ and $k^a$ of O-H bonds in naphthol and reference molecules are shown in Figure 4 and their respective local mode force constants and local mode frequencies are reported in Table 3. An in-depth assessment of the O-H bond strengths in naphthol is imperative, in particular to evaluate the nature of the C-H - H-O interaction in cis-1-naphthol. In case of 1-naphthol, a bonding interaction between H-atoms would lead to a weakening of the O-H bond for the cis conformer compared to the trans conformer as electron density is moved from the O-H bond region. Conversely, a stronger O-H bond in cis-1-naphthol would be consistent with the traditional notion of a steric repulsion between the H-atoms. In comparison to the reference molecules, it can be seen that the O-H bond strengths in naphthol compounds tend to be weaker than those in water, propen-2-ol, and methanol while they are comparable to those in phenol and 2-propanol. Importantly, it is revealed that the O-H bond strengths in naphthol compounds vary in the order, cis-1-naphthol > trans-2-naphthol > trans-1-naphthol > cis-2-naphthol. In accord with the OH normal mode is not coupling with the other modes, e.g. bending modes. For 1-naphthol, the cis-conformer is higher in energy by 0.8 kcal/mol than the trans form (see Figure 2), in accord with the relative intensities of the recorded transitions. The trans-to-cis barrier is 3.3 kcal/mol, which is not so easily surmountable in a molecular beam expansion. From the relative intensities in the experimental spectrum, we estimate a conformational temperature of about 150 K. The situation for 2-naphthol is similar, except that the relative energies of cis and trans forms are reversed, with the cis form more stable by 0.5 kcal/mol. In this case, the conformational temperature is estimated to be 170 K. In the case of cis-1-naphthol, there is a barrier at the planar configuration and there is the possibility for a tunneling motion to occur in which the OH hydrogen atom moves from above to below the C-skeleton plane. However, consideration of the zero-point vibrations shows that the zero-point energy level lies above the barrier (by 0.1 kcal/mol), indicating that the OH hydrogen undergoes a large amplitude vibrational motion.
as reflected by the blue shift, which has also been confirmed experimentally,\textsuperscript{15} of cis-1-naphthol we can quantify the repulsive nature of the close contact H-atoms based on vibrational spectroscopy.

**Interpretation of the close H-H contact in cis-1-naphthol**

The interpretation by Bader and co-workers of the bond path and bond critical point between the close-contact H-atoms in terms of bonding, i.e. H-H bonding, interactions has been criticized, for example by Bickelhaupt and co-workers.\textsuperscript{61} In particular, they point out that the physical meaning of the atomic energy, \( E_a \), is not particularly clear. They consider that the reduction of the H-atom \( E_a \) when two H-atom come into close contact may be a result mainly of a reduction in atomic basin volume, for the case of insignificant charge transfer. Grimme et al.\textsuperscript{62} have also voiced their doubts about interpreting the bond critical point between the close-contact H-atoms in terms of a bonding interaction. They analysed the experimental splitting between symmetric and antisymmetric stretching normal modes of vibration involving close-contact D-atoms in dideuterophenanthrene. From this analysis and by comparison with theoretical calculations, they conclude that the observations are inconsistent with a bonding interaction between the close-contact D-atoms.

Cremer, Kraka et al. introduced a new aromaticity delocalization index assessing \( \pi \)-delocalization in aromatic systems based on vibrational frequencies.\textsuperscript{63,64} They could clarify that the larger stability of phenanthrene relative to anthracene predominantly results from its higher resonance energy, a direct consequence of the topology of ring annellation and not from a maximum electron density path between the bay H atoms.\textsuperscript{63} Furthermore, they pointed out that the close spatial proximity of ring and bond critical points in the bay region of phenanthrene and their low electron densities, as well as the positive energy density at the bond critical point, are indicative of an electrostatic, destabilizing interaction, confirming the findings of Grimme et al.\textsuperscript{62}

In an effort to elucidate the interaction between the two close contact H-atoms, a Charge Model 5 (CMS) charge analysis was carried out, which derives partial atomic charges from a Hirshfeld population analysis. By measuring the CMS atomic charges as the C-C-O-H dihedral angle of cis 1-naphthol is varied from 0° to 90°, one can ultimately infer how the electron density changes as the two hydrogens approach each other. Because of the absence of a bond critical point (QTAIM plots) and non-covalent interactions (NCI plots) between the two hydrogens, we used \textit{trans}-1-naphthol, \textit{cis}-2-naphthol, and \textit{trans}-2-naphthol as control molecules. We examined the charges of the five atoms which play the most prominent role in this interaction: the hydroxy hydrogen (H'), oxygen (O'), the carbon bound to the oxygen (C'), the adjacent hydrogen (H''), and the carbon bound to the adjacent hydrogen (C''). The labelling is summarized in Figure S9 of the ESI. The charge as a function of dihedral angle for all four isomers of naphthol, and the net charge change for each atom are summarized in Figure S10 and Table S27 of the ESI, respectively. Based on Figure S10 and Table S27, no anomalies are observed for atoms C'' and C' of cis 1-naphthol as the charge change generally follows the same pattern as in the other isomers. For the most part the O' atom follows the trends of the other isomers, with the net change (Table S27) only slightly smaller than in its counterparts. Interestingly, for the two hydrogens the charge pattern for cis-1-naphthol deviates significantly from the other isomers. For the H' atom the charge, and thus the electron density at the atom, does not change significantly in cis-1-naphthol, while it increases for the other isomers with increasing dihedral angle. For the H'' atom the charge decreases (increasing electron density) as the dihedral angle approaches 0°, while it stays almost constant for the other isomers. Despite this clear increase in electron density for H', attributing this to a bonding or repulsive interaction is not so clear. On one hand the slightly smaller change in net charge of O', and therefore less electron density being moved away from the O' atom, may indicate a steric repulsive H-H interaction. On the other hand, an increase in electron density for the H'' atom may be indicative of a bonding interaction as electron density is being transferred to the H' atom from the H'' atom. In this case the H'' atom would be serving as an electron donor while the H' atom as the acceptor.

To further examine this electron transfer and clarify the CMS results, a Natural Bond Orbital (NBO) analysis was carried out. An NBO analysis allows for the decomposition of non-covalent and covalent interactions into localized bonding and antibonding orbitals. In the framework of NBO a bonding interaction is present when electron density is transferred from the bonding orbital (BD) to the antibonding orbital (BD*). For example, if a bonding interaction was to be present in cis-1-naphthol one would observe an electron density transfer from the O'-H' BD to the C''-H'' BD* or from the C'-H' BD to the O'-H' BD*. Table S28 of the ESI presents the electron occupancy for the O'-H' and the C''-H'' BD and the O'-H' and the C'-H' BD*. From table S28, there is a slightly lower occupancy in both the O'-H' BD and C''-H'' BD and a higher occupancy in both the O'-H' BD* and C'-H' BD* in cis-1-naphthol compared to the other isomers, supporting the notion of a bonding interaction. Using a second order perturbative treatment of the Fock matrix, the two interactions (C'-H' BD --- O'-H' BD* and O'-H' BD --- C''-H'' BD*) can be quantified in terms of stabilization energies as a function of dihedral angle (Tables S29-S32). The sum of the two BD-BD* interactions for each isomer is plotted in Figure S11, panel b, of the ESI. From the figure one can clearly see that the stabilization energies are several kJ mol\textsuperscript{-1} higher for cis-1-naphthol than the other isomers, demonstrating an attractive force between the two hydrogen atoms. To examine the repulsive components of the interaction we looked at the pairwise steric exchange energies between the C''-H'' BD and O'-H' BD. The results are plotted in Figure S11, panel a, of the ESI. The steric exchange energy for cis-1-naphthol increases far more compared to the other isomers as the dihedral angle approaches 0° and outweighs the stabilization energy by a factor of almost 3. Therefore a large steric or repulsive interaction between the two hydrogens is present. Finally, comparing the steric (repulsion) energies to the stabilization
Table 3. For each local mode, the local mode force constant $k^a$(mDyne/Å for stretching and mDyne/Å/Å for bending vibrations) and local mode frequency $\omega^a$ (cm$^{-1}$) are given. Reported values are for MP2/aug-cc-pVTZ calculations followed by values for B3LYP-D3(Bj)/def2-TZVP calculations in parentheses.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>cis-1-naphthol</th>
<th>trans-1-naphthol</th>
</tr>
</thead>
<tbody>
<tr>
<td>O1-H1</td>
<td>8.169 (8.085)</td>
<td>8.097 (8.057)</td>
</tr>
<tr>
<td>C1-O1</td>
<td>5.825 (5.609)</td>
<td>5.790 (5.673)</td>
</tr>
<tr>
<td>H1-O1-C1</td>
<td>0.743 (0.735)</td>
<td>0.745 (0.749)</td>
</tr>
<tr>
<td>O1-C1-C2</td>
<td>1.716 (1.694)</td>
<td>1.756 (1.804)</td>
</tr>
<tr>
<td>O1-C1-C3</td>
<td>1.771 (1.728)</td>
<td>1.727 (1.777)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>cis-2-naphthol</th>
<th>trans-2-naphthol</th>
</tr>
</thead>
<tbody>
<tr>
<td>O1-H1</td>
<td>8.091 (8.035)</td>
</tr>
<tr>
<td>C1-O1</td>
<td>5.851 (5.735)</td>
</tr>
<tr>
<td>H1-O1-C1</td>
<td>0.749 (0.754)</td>
</tr>
<tr>
<td>O1-C1-C2</td>
<td>1.599 (1.647)</td>
</tr>
<tr>
<td>O1-C1-C3</td>
<td>1.548 (1.600)</td>
</tr>
</tbody>
</table>

References
- water methanol 2-propanol propen-2-ol phenol
<table>
<thead>
<tr>
<th>Parameter</th>
<th>k*</th>
<th>$\omega^a$</th>
<th>k*</th>
<th>$\omega^a$</th>
<th>k*</th>
<th>$\omega^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>O1-H1</td>
<td>8.348 (8.151)</td>
<td>3865.9 (3819.9)</td>
<td>8.288 (8.091)</td>
<td>3851.9 (3805.9)</td>
<td>8.158 (8.045)</td>
<td>3821.5 (3795.0)</td>
</tr>
<tr>
<td>C1-O1</td>
<td>4.775 (4.628)</td>
<td>1087.3 (1070.4)</td>
<td>4.417 (4.220)</td>
<td>1045.6 (1022.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1-O1-C1</td>
<td>-</td>
<td>0.731 (0.736)</td>
<td>1241.6 (1245.2)</td>
<td>0.718 (0.723)</td>
<td>1227.0 (1232.1)</td>
<td></td>
</tr>
<tr>
<td>O1-C1-C2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.280 (1.301)</td>
<td>609.3 (613.6)</td>
<td></td>
</tr>
<tr>
<td>O1-C1-C3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.323 (1.340)</td>
<td>608.8 (613.0)</td>
<td></td>
</tr>
</tbody>
</table>

(11) one can clearly see that although the stabilization energy is considerably larger when compared to the isomers, the steric energy is even larger and simply outcompetes the stabilization.

QTAIM and NCI analyses evaluate the electron density distribution and the presence of a bond critical point means that there is a minimum of electron density between the two atoms and maxima in the two perpendicular directions. Chemical bonds are characterized by such electron density topology and it is fair to say that there is some weak bonding contribution to the H-H interaction in cis-1-naphthol. Perhaps not only in the case of cis-1-naphthol, it is necessary to also consider the repulsive interactions, which outweigh the weak H-H attraction.

The major ‘attractive’ contribution that keeps the OH H-atom in cis-1-naphthol in such close contact with the neighbouring ring H-atom is provided by the covalent chemical bonds, in particular a maximization of $\pi$-electron delocalization, as already pointed out by Whitham et al.21

Conclusions

Rotational spectra of the 1-naphthol and 2-naphthol monomers were recorded using a CP-FTMW spectrometer in the 2-6 GHz range. We assigned the spectrum of trans-2-naphthol and extended the measurements for the other isomers. In addition to the parent species, spectra of ten $^{13}$C isotopologues were assigned for each conformer. The $^{13}$C isotopologues were then used to determine carbon-skeleton substitution structures. The corresponding bond lengths and bond angles are in excellent agreement with the theoretical results.

For cis-1-naphthol, a QTAIM analysis gives a bond path with bond critical point between the close-contact OH and ring H-atom. Local mode analyses were carried out for the naphthols and the O-H bond strengths were compared for all four conformers. The results show that cis-1-naphthol has strongest O-H bond among the four isomers. This is consistent with the traditional notion of a steric repulsion between the two close-contact H-atoms, in accord also with the blue-shift of the OH stretching frequency in cis-1-naphthol. Charge Model 5 and, in particular, Natural Bond Orbital analyses of cis-1-naphthol do support the notion of a weak bonding H- H interaction, which, however, is outweighed by steric repulsion. The major part of the attraction force is presumably provided by a maximization of $\pi$-electron delocalization, as already pointed out by Whitham et al.21

Conflicts of interest

There are no conflicts to declare.

Acknowledgements
This research was funded by the University of Alberta and the Natural Sciences and Engineering Research Council (NSERC) of Canada. We gratefully acknowledge access to the computing facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET: www.sharcnet.ca), the Western Canada Research Grid (Westgrid) and Compute/Calcul Canada. SN and EK thank the Center for Research Computing at Southern Methodist University for providing generous computational resources and the National Science Foundation for support (grant CHE 1464906).

Notes and references
35. F. Weigend, Physical chemistry chemical physics, 2006, 8, 1057-1065.
36. C. Moller and M. S. Plesset, Physical review, 1934, 46, 618.
Appendix H

Paper on halogen bonds in crystals

In Situ Assessment of Intrinsic Strength of X-I· · ·OA-Type Halogen Bonds in Molecular Crystals with Periodic Local Vibrational Mode Theory

Yunwen Tao 1,†, Yue Qiu 2,†, Wenli Zou 3, Sadisha Nanayakkara 1, Seth Yannacone 1 and Elfi Kraka 1,*

1 Department of Chemistry, Southern Methodist University, 3215 Daniel Avenue, Dallas, TX 75275-0314, USA; yunwent@smu.edu (Y.T.); snanayakkara@smu.edu (S.N.); syannacone@smu.edu (S.Y.)
2 Grimwade Centre for Cultural Materials Conservation, School of Historical and Philosophical Studies, Faculty of Arts, University of Melbourne, Parkville, VIC 3052, Australia; yueq1@student.unimelb.edu.au
3 Institute of Modern Physics, Northwest University, and Shaanxi Key Laboratory for Theoretical Physics Frontiers, Xi’an 710127, China; wzou@smu.edu
* Correspondence: ekraka@smu.edu
† These authors contributed equally to this work.

Received: 14 February 2020; Accepted: 24 March 2020; Published: 30 March 2020

Abstract: Periodic local vibrational modes were calculated with the rev-vdW-DF2 density functional to quantify the intrinsic strength of the X-I· · ·OA-type halogen bonding (X = I or Cl; OA: carbonyl, ether and N-oxide groups) in 32 model systems originating from 20 molecular crystals. We found that the halogen bonding between the donor dihalogen X-I and the wide collection of acceptor molecules OA features considerable variations of the local stretching force constants (0.1–0.8 mdyn/Å) for I· · ·O halogen bonds, demonstrating its powerful tunability in bond strength. Strong correlations between bond length and local stretching force constant were observed in crystals for both the donor X-I bonds and I· · ·O halogen bonds, extending for the first time the generalized Badger’s rule to crystals. It is demonstrated that the halogen atom X controlling the electrostatic attraction between the σ-hole on atom I and the acceptor atom O dominates the intrinsic strength of I· · ·O halogen bonds. Different oxygen-containing acceptor molecules OA and even subtle changes induced by substituents can tweak the n → σ∗(X-I) charge transfer character, which is the second important factor determining the I· · ·O bond strength. In addition, the presence of the second halogen bond with atom X of the donor X-I bond in crystals can substantially weaken the target I· · ·O halogen bond. In summary, this study performing the in situ measurement of halogen bonding strength in crystalline structures demonstrates the vast potential of the periodic local vibrational mode theory for characterizing and understanding non-covalent interactions in materials.

Keywords: halogen bonding; dihalogen; local vibrational mode theory; local stretching force constant; molecular crystal; chemical bond strength; vibrational spectroscopy; crystal engineering; generalized Badger’s rule; VASP

1. Introduction

Halogen bonding (D-X· · ·Y) is one important type of non-covalent interaction between a donor halogen atom X (F, Cl, Br or I) and an electron-rich atom/group Y (e.g., atoms with lone pair electrons including N, O, P and S) [1]. With its great bond strength tunability, halogen bonding has gained popularity in drug design [2–5], enzyme engineering [6], material science [7–9], catalysis [10–13] and crystal engineering [14–16]. The currently well-accepted understanding of halogen bonding focuses on the interplay of the following contributions: (1) charge transfer from nucleophile Y to the σ∗
anti-bonding orbital of D-X, (2) attractive electrostatic forces, (3) dispersion interaction, and (4) a repulsive term arising from Pauli exclusion principle. The weight of the above individual terms varies for different types of halogen bonding [17–21].

During the process of conceptualizing halogen bonding, theoretical and computational chemistry played an indispensable role. Politzer and co-workers first noticed the anisotropic electron density distribution around the donor halogen atom X [22] and they found in the elongation of D-X bond a small region with surprisingly positive electrostatic potential (ESP), which was then named $\sigma$-hole to explain the attraction between halogen atom X and nucleophile Y [23–25]. To understand the physical nature of the stabilization energy of halogen bonding, natural energy decomposition analysis (NEDA) [26] was employed by Infante and co-workers [27] to demonstrate that the $\sigma$-hole theory itself could not explain all halogen bonding situations and the contribution from $n \rightarrow \sigma^*$ charge transfer is substantial. Frontier orbital analysis [28] and natural bond orbital (NBO) analysis [29] were employed by Rosokha and co-workers to confirm the importance of charge transfer [30–32]. With symmetry-adapted perturbation theory (SAPT) [33,34], Stone came to similar conclusions that induction (which contains charge transfer) and dispersion are essential for halogen bonding besides the electrostatic term [35].

While many contributions have been made to obtain more accurate density functionals and more computationally attainable wavefunction theory methods for describing halogen bonding theoretically in terms of geometry and binding energy [17], the quest for theoretical tools emphasizing on bonding analysis is equally crucial for studying halogen bonding due to two reasons.

- The binding energy is a reaction parameter, summarizing all effects leading to bonding in a cumulative way. Even in a simple dimer the binding energy cannot serve as a measure for the intrinsic strength of a bond; it is contaminated with the stabilization energy of the two fragments caused by geometry relaxation and reorganization of the electron density of the fragments upon bond breakage [36]. This applies even more to complex systems with halogen bonding at work (e.g., a halogenated drug in a protein or halogen bonding in crystals);
- We need a bond strength measure that is not based on bond breaking and that follows Levine’s suggestion that chemistry is local [37].

Besides the above-mentioned analytical tools (e.g., NBO, SAPT), a few other tools have been employed in halogen bonding studies. The quantum theory of atoms in molecules (QTAIM) [38] can be employed to obtain the bond critical points (BCPs) of halogen bonding, showing that the electron density at BCPs correlates well with the interaction energy [39]. At (3, -1)-type BCPs, the local energy density by Cremer and Kraka determines whether a halogen bond is more covalent or ionic [40,41]. Methods like reduced density gradient (RDG) [42], electron localization function (ELF) [43], density overlap regions indicator (DORI) [44], and independent gradient model (IGM) [45] can identify the existence of halogen bonding graphically [46,47].

An important theoretical tool which has considerably contributed to a deeper understanding of halogen bonding is the local vibrational mode theory originally introduced by Konkoli and Cremer [48–53]. They derived local vibrational modes (associated with individual internal coordinates such as bond lengths, bond angles, etc.) directly from the normal vibrational modes (whose fundamental frequencies $\omega_{\mu}$ can be measured by infrared and Raman spectroscopy or calculated theoretically) by solving mass-decoupled Euler-Lagrange equations leading to a mass-decoupled analogue of Wilson’s equation of vibrational spectroscopy. Normal vibrational modes of polyatomic molecules are generally delocalized because of mass coupling [54–57] and therefore cannot directly be used as bond strength measure, which is an important fact but often overlooked.

Each local mode is associated with a corresponding local mode frequency $\omega_{\mu}^a$ and local mode force constant $k_{\mu}^a$. Zou and co-workers demonstrated that there is a one-to-one relationship between the local and normal vibrational modes that can be verified by an adiabatic connection scheme (ACS), providing a solid physical fundament for the local vibrational mode theory [52]. Zou and Cremer
further proved that the local stretching force constant \( k_a \) is directly related to the intrinsic strength of chemical bonds [58]. The underlying physical nature of this important proof results from the fact that \( k_a \) equals the curvature of the potential energy surface (PES) in the direction of the bond stretching, determined via an infinitesimal change of the bond length and followed by the relaxation of all other atoms in the molecule [59,60]. In this way, \( k_a \) absorbs all electronic effects contributing to an individual chemical bond and serves as unique measure of the intrinsic strength of a chemical bond and/or weak chemical interaction based on vibrational spectroscopy. In combination with other analytical tools (e.g., NBO analysis or electron density analysis) and the knowledge about well-studied systems, local stretching force constants offer a deeper and more comprehensive understanding of complex bonding situations, in particular determining whether the presence/absence of a specific electronic effect (e.g., \( \pi \) conjugation) is strengthening/weakening a chemical bond of interest in a comparative manner.

So far, the local mode analysis has been successfully applied to characterize covalent bonds [58,61–66] and weak chemical interactions such as intra- and inter-molecular hydrogen bonding in various forms and systems [67–73], chalcogen [74–76], pnictogen [77–79] and tetrel interactions [80], and in particular halogen bonding [81–84]. Recently, we extended the local vibrational mode theory from molecular to periodic one-dimensional (1D) through three-dimensional (3D) systems [85]. We consider it as an important step forward as it enables us to quantify and compare intrinsic bond strength in both materials/crystals and molecules, in particular considering (i) that currently only a few analytical tools are available for bonding analysis in periodic systems [86–94] due to the complication arising from the use of plane-wave basis set in first-principle modeling; (ii) the often reduced functionality of periodic versions of analysis tools originally designed for molecular systems (e.g., periodic NBO [95] does not provide interaction energies from second-order perturbation theory); and (iii) that the lattice structure intrinsically prohibits the calculation of bond dissociation energies.

In this work, we will apply our periodic local vibrational mode theory [85] to explore the X-I⋯O-A-type halogen bonding (where X-I is the donor bond and I⋯O is the halogen bond) in molecular crystals in a systematic and comprehensive way. We chose this specific type of halogen bonding based on the following considerations.

1. The I⋯O halogen bonds account for a large portion of all halogen bonds ever discovered;
2. The oxygen acceptor atom is more common in molecular crystals than the higher chalcogens (e.g., S, Se and Te);
3. Dihalogen/interhalogen compounds X-I consist of only two atoms and therefore are the simplest halogen bond donors;
4. Recently, Rosokha and co-workers investigated the I-I⋯O-N-type halogen bonding in crystals with N-oxide acceptors via crystallography and theoretical calculations [96,97]. Their analysis based on molecular dimer models suggests that the charge transfer is a key factor in the I⋯O halogen bonding besides electrostatic attraction. Our work here based on a collection of crystal structures should provide a more detailed and comprehensive understanding of halogen bonding in materials.

The following objectives were pursued in this work.

- To create a comprehensive set of local stretching force constants for X-I⋯O-A halogen bonds in different crystals describing the intrinsic halogen bond strength in these systems;
- To derive a more realistic model description considering the crystal packing effect explicitly and to understand factors that affect the solid state halogen bond strength by an in situ investigation of halogen bonding in a crystalline environment;
- To test the validity of the Badger’s rule [98–100], in particular the validity of the generalized Badger’s rule based on local stretching force constants [61] in crystals.

As the first systematic investigation of a series of non-covalent interactions using our periodic local vibrational mode theory, this paper will set an example for following projects, especially in terms
of properly obtaining local mode force constants for periodic systems from first-principle calculations with sufficient accuracy.

The paper is structured in the following way: First, Computational Details are given. In the Results and Discussion section, the calculated crystal structures are discussed followed by a comparison of key structural features between computational and experimental results. Then we quantify the intrinsic strength of both donor and donor-acceptor bonds of X-I···O halogen bonding, as determined with periodic local vibrational mode theory. Important findings of this work are summarized in the Conclusions section and a future perspective is given.

2. Computational Details

All calculations in this work were carried out with the VASP 5.4.4 package [101–105] using van der Waals density functional rev-vdW-DF2 [106] with projector-augmented wave (PAW) potential [107,108]. The kinetic energy cutoff for basis set is 1000 eV. The rev-vdW-DF2 functional is a recently revised version of vdW-DF2 functional [109] with improved accuracy in describing van der Waals interactions (i.e., non-covalent bonding). According to a timely benchmark by Tran and co-workers, the rev-vdW-DF2 functional was shown to be the best choice for describing non-covalent interactions in molecular crystals [110].

All crystal systems investigated in this work were initially converted into primitive unit cells, as required by the definition of local vibrational modes in periodic systems [85]. The Monkhorst-Pack scheme [111] was used to sample the Brillouin zone with a k-point grid, where each lattice constant a times its number of k-points k is close to 28 Å, i.e., a · k ∼ 28 Å. Noteworthy is that the settings for the energy cutoff and k-point sampling were based on delicate convergence testing of total energy and ionic forces. All halogen-bonded systems investigated in this work were optimized to local minima, verified via vibrational frequency calculations to ensure that there were no imaginary frequencies. A relatively tight criterion for geometry optimization as well as cell relaxation was adopted with the maximum ionic force less than 10⁻⁷ eV/Å. The force constant matrices were calculated numerically with analytic gradients using two displacements (±Δs) for each degree of freedom, where the step size Δs was set to 0.005 Å.

The molecules of diiodine (I₂) and iodine monochloride (ICl) were calculated in a cubic box with the size of 20 Å and 24 Å, respectively, to simulate non-periodic calculations. Dipole corrections to the total energy were added for ICl along its dipole moment. Only the Γ point (k = 0) was considered for sampling the Brillouin zone.

The intrinsic bond strength of the halogen bonds was quantified by their local stretching force constants kₐₙ derived from the local vibrational mode theory for periodic systems [85]. The periodic local mode analysis (whose time cost for local mode analysis in either molecules or crystals is comparable to normal mode analysis as long as the Hessian matrix has been calculated beforehand) was carried out with COLOGNE2019 package [112].

3. Results and Discussion

3.1. Selection of Molecular Crystals

This work was focused on the X-I···OA halogen bonding in crystals, of which the structures were retrieved from two major databases collecting crystal structures — the Cambridge Structural Database (CSD) [113] and the Crystallography Open Database (COD) [114] by performing searches with the keywords “dihalogen”, “interhalogen”, “diiodine” or their synonyms. 62 unique crystal structures with halogen bonding, matching the search keywords were obtained and then manually screened based on the following criteria:

- The molecular crystal should contain only the elements C, N, H, O/S/Se and X while excluding metal atoms;
• The total number of atoms in the primitive cell of the crystal should be preferably smaller than 80 to save computational cost;
• The dihalogen/interhalogen X-I should exist as neutral diatomic molecules instead of trihalogen cations.

20 crystal structures (as listed in the first column of Table 1) passed the screening procedure and they contain different dihalogen/interhalogen molecules including I2, IBr, ICl or Br2. To simplify the investigation, in particular with regard to a more consistent and straightforward analysis, we limited our investigation to I2 and ICl molecules with iodine atom as donor participating in the halogen bond. For the sake of enriching the data set, we carried out in silico crystal engineering by replacing the dihalogen/interhalogen molecules within those 20 crystals with either I2 or ICl molecules, leading to 20 × 2 = 40 model systems. The reasons why we chose I2 and ICl as the halogen bond donors are due to the following:

1. Iodine as the donor atom has relatively large polarizability and will more easily form a σ-hole than chlorine, bromine or fluorine;
2. Chlorine is more electronegative than bromine and it generally forms stronger X-I···O halogen bonds according to the σ-hole theory [23–25];
3. Although stronger halogen bonding is expected for iodine monofluoride (IF) as the donor molecule, this species is unstable and cannot form co-crystals under ambient conditions [115].

Furthermore, all sulfur and selenium atoms within the acceptor molecules of these 40 model systems were replaced with oxygens for consistency. These model systems were then used for VASP calculation to relax their lattice structures. The primitive cell structure of 33 model systems could be optimized while the remaining 7 systems either underwent drastic structural changes via isomerization or failed to meet the desired convergence criteria (see supplementary materials). It is worth noting that the overall arrangement of X-I donors and acceptor molecules in the optimized structures was retained compared to the original crystal structures without element replacement.

The calculations of the Hessian matrices confirmed that 32 model systems were optimized to local minima on the PES, while only one model system has a negative eigenvalue of its Hessian matrix and therefore was removed from our investigation. Subsequently, we obtained 34 unique X-I···O halogen bonds out of 32 model systems derived from the 20 crystal structures, as listed in Table 1. For a more systematic discussion, the 32 model systems were then organized according to the 16 acceptor molecules, shown in Figure 1, labeled A–Q.

![Figure 1. Structures of the 16 halogen bonding acceptor molecules investigated in this work. In any structure with two or more oxygen atoms, the oxygen participating in I···O halogen bonding is colored in red. Below each 2D molecular structure is the COD/CSD id number for the crystal structure(s) from which the acceptor molecule was extracted or adapted. The colored dot in front of the COD/CSD id number indicates that the model crystal structure containing the present acceptor molecule associated with diiodine (I2; purple dot) or iodine monochloride (ICl; green dot) was calculated in this work.](image-url)
Table 1. Summary of 34 X-I···O halogen bonding interactions investigated in this work.

<table>
<thead>
<tr>
<th>ID a</th>
<th>Label b</th>
<th>N_atom c</th>
<th>Donor Bond f</th>
<th>r c</th>
<th>r_exp. d</th>
<th>k_abc e</th>
<th>Halogen Bond f</th>
<th>r c</th>
<th>r_exp. d</th>
<th>k_abc e</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSD-1522625 [116]</td>
<td>A</td>
<td>24</td>
<td>H-I</td>
<td>2.7598</td>
<td>1.257</td>
<td>1..O=C(C2)</td>
<td>2.8133</td>
<td>0.087</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COD-1543603 [117]</td>
<td>B1-1</td>
<td>30</td>
<td>I-I</td>
<td>2.7994</td>
<td>1.049</td>
<td>2.5281</td>
<td>0.309</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COD-1543604 [117]</td>
<td>B2</td>
<td>56</td>
<td>H-I</td>
<td>2.7753</td>
<td>1.199</td>
<td>2.7886</td>
<td>2.8195</td>
<td>0.143</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSD-1201775 * [118]</td>
<td>C1-1</td>
<td>38</td>
<td>I-I</td>
<td>2.7715</td>
<td>1.134</td>
<td>2.5170</td>
<td>0.370</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSD-1272661 * [119]</td>
<td>C2-1</td>
<td>38</td>
<td>I-I</td>
<td>2.7714</td>
<td>1.186</td>
<td>2.5195</td>
<td>0.368</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COD-7228662 * [119]</td>
<td>D</td>
<td>88</td>
<td>H-I</td>
<td>2.4447</td>
<td>1.335</td>
<td>2.4616</td>
<td>0.416</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COD-7027471 * [120]</td>
<td>E1-1</td>
<td>36</td>
<td>I-I</td>
<td>2.7317</td>
<td>1.364</td>
<td>2.7340</td>
<td>0.149</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COD-7027472 * [120]</td>
<td>E1-2</td>
<td>36</td>
<td>I-I</td>
<td>2.4147</td>
<td>1.526</td>
<td>2.4958</td>
<td>0.412</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COD-4323206 * [121]</td>
<td>F1</td>
<td>60</td>
<td>I-I</td>
<td>2.7667</td>
<td>1.146</td>
<td>2.6855</td>
<td>0.159</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COD-1552728 [97]</td>
<td>G1-1</td>
<td>60</td>
<td>I-I</td>
<td>2.7724</td>
<td>1.176</td>
<td>2.5575</td>
<td>0.321</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSD-1270637 * [122]</td>
<td>G2-1</td>
<td>60</td>
<td>I-I</td>
<td>2.4553</td>
<td>1.326</td>
<td>2.4145</td>
<td>0.561</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSD-1457584 [123]</td>
<td>H1</td>
<td>16</td>
<td>H-I</td>
<td>2.7685</td>
<td>2.6926</td>
<td>2.7565</td>
<td>2.8078</td>
<td>0.203</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSD-1145571 * [124]</td>
<td>H2-1</td>
<td>36</td>
<td>I-I</td>
<td>2.7718</td>
<td>1.115</td>
<td>2.6949</td>
<td>0.212</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSD-1543603</td>
<td>N</td>
<td>26</td>
<td>H-I</td>
<td>2.8073</td>
<td>2.7340</td>
<td>1.036</td>
<td>2.7565</td>
<td>2.8078</td>
<td>0.203</td>
<td></td>
</tr>
<tr>
<td>CSD-152252730 [97]</td>
<td>O-1</td>
<td>44</td>
<td>I-I</td>
<td>2.8629</td>
<td>2.7952</td>
<td>2.3857</td>
<td>2.3587</td>
<td>0.565</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSD-19129389 [97]</td>
<td>P-1</td>
<td>40</td>
<td>H-I</td>
<td>2.8188</td>
<td>2.7512</td>
<td>2.4470</td>
<td>2.4637</td>
<td>0.435</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSD-1588334 [96]</td>
<td>Q</td>
<td>40</td>
<td>H-I</td>
<td>2.8118</td>
<td>2.7328</td>
<td>2.4850</td>
<td>2.4990</td>
<td>0.198</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diiodine</td>
<td>I2</td>
<td>2</td>
<td>I-I</td>
<td>2.6919</td>
<td>2.6660</td>
<td>1.667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iodine monochloride</td>
<td>ClI</td>
<td>2</td>
<td>Cl-I</td>
<td>2.3431</td>
<td>2.3207</td>
<td>2.233</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a The id number of the crystal structure from COD/CSD database. The "*" symbol indicates that sulfur, selenium, or tellurium atoms have been replaced with oxygen atoms in this crystal structure. The label for each halogen bond is formatted as X-i·ii y, where letter "X" denotes a specific acceptor molecule as shown in Figure 1, number "i" denotes the first or second COD/CSD structure, number "ii" denotes whether the donor molecule is diiodine or iodine monochloride and letter "y" denotes the first or second halogen bonding in the same crystal model system.

b Calculated bond length r and local stretching constant k_abc in the unit of Angstrom (Å) and mdyn/Å, respectively.

c Experimentally measured bond length r_exp. with the unit of Angstrom (Å) in crystal structures.

d This table is divided into 7 sections based on the halogen bond type. The shade in gray color is used solely for distinguishing different COD/CSD structures within each section.

e The symbol § before a X-I donor means there exists a close contact (i.e., a halogen bond) with a X···O bond stretching can be calculated by 

\[ \omega_n = \sqrt{k_n / \mu_{\text{d}}} \]

where \[ \mu_{\text{d}} \] is in the unit of mdyn/Å and the resulting vibrational frequency \( \omega_n \) is in the unit of cm\(^{-1}\). The local mode frequencies \( \omega_n \) for Cl-I bond stretching can be calculated by 

\[ \omega_n = \sqrt{k_n / \mu_{\text{d}}} \]

where \( k_n \) is in the unit of mdyn/Å and the resulting vibrational frequency \( \omega_n \) is in the unit of cm\(^{-1}\).

f The symbol § before a X-I donor means there exists a close contact (i.e., a halogen bond) with a distance of less than 3.0 Å with regard to the atom X.

3.2. Comparison of Experimental and Calculated Structures

In contrast to analytical tools qualitatively describing chemical bonding from properties directly extracted from experimentally resolved crystal structures (e.g., the Hirshfeld charge/surface analysis [127]), our local vibrational mode analysis requires the system of interest to be optimized into a local minimum point on the PES. Therefore, it becomes necessary to check the deviation of the
optimized structure from the single-crystal X-ray structure. In Figure 2, the bond lengths for both the X-I donor bonds and I···O halogen bonds of the X-I···O halogen bonding in 8 crystal structures (as listed in Table 1) are examined.

![Figure 2](image_url)

**Figure 2.** Comparison between calculated versus the experimental bond lengths in 8 crystal structures. The diagonal dotted line is $y = x$, i.e., reflecting 100% agreement between experimental and calculated values. Purple and green dots represent I-I and I-Cl covalent bonds, respectively. The diiodine (I$_2$) and iodine monochloride (ICl) molecules in gas phase are also included for comparison. The solid black line shows a linear fit ($R^2 = 0.993$) for I-I and I-Cl covalent bonds altogether. The red dots representing I···O halogen bonds are best fitted with a quadratic curve ($R^2 = 0.996$) shown as the dashed black line.

The calculated covalent bond lengths of I$_2$ and ICl molecules in these 8 systems are slightly longer than those extracted from the experimental structures. We also considered the isolated I$_2$ and ICl molecules in gas phase calculated with the same level of theory and their calculated bond lengths are also marginally longer than experimental values [115]. We combined the I-I and I-Cl distances together and observed a strong linear correlation with a coefficient of determination ($R^2$) as 0.993 between the calculated and experimental bond lengths. On the other hand, most of the bond lengths for I···O interactions are slightly underestimated but they seem to follow a quadratic function with $R^2 = 0.996$ between the calculated and measured values, which is surprising. To the best of our knowledge, it is the first time that one has observed a linear correlation for donor bonds along with a quadratic correlation for donor-acceptor bonds for a certain type of non-covalent interaction between calculated and experimentally measured bond lengths. We postulate that the quadratic correlation for I···O interactions is the result of a second-order perturbation while the covalent X-I bonds follow a first-order perturbation. Further validation of such relationships for other types of donor-acceptor interactions is currently under investigation. This interesting bond length-bond strength relationship can be potentially useful to predict key parameters in the experimental crystal structure based on a calculated model or as a novel metric to benchmark the quality of density functionals [128].
3.3. Intrinsic Strength of Donor Bonds and Halogen Bonds in X-I···OA Halogen Bonding

3.3.1. General Trends

Figures 3 and 4 show the relationship between local stretching force constant $k_a^n$ and bond length $r$ for 34 halogen bonding scenarios in terms of X-I donor bond and I···O halogen bond, respectively. More detailed information for these halogen bonds is collected in Table 1. As revealed in Figure 3, we find a strong correlation between force constant and bond length for I-I donor bonds ($R^2 = 0.971$) and an even stronger correlation ($R^2 = 0.997$) for the Cl-I donor bonds. The correlation between force constant and bond length for the I···O bonds shown in Figure 4 is somewhat weaker with $R^2 = 0.923$ and $R^2 = 0.974$ by excluding two outliers. This reveals that in essence the Badger’s rule [98–100] still holds, namely shorter bonds have larger harmonic frequencies (or force constants) and are thus stronger. Given that the Badger’s rule was generalized to polyatomic molecules based on local stretching force constant by one of the authors of this work [61] and was recently extended to O-H bonds in liquid water [129], this work demonstrates for the first time that the Badger’s rule even holds for crystals.

![Figure 3](image)

**Figure 3.** Relationship between local stretching force constant $k_a^n$ and bond length $r$ for Cl-I (green) and I-I (purple) covalent bonds, respectively. Data points are shown with at least 4 shapes based on the acceptor molecule type, see Figure 1. Power functions in the form of $k_a^n = a \cdot r^b + c$ were employed in fitting the data points for Cl-I and I-I bonds separately.

Figure 3 shows that all Cl-I covalent bonds are shorter than I-I bonds by at least 0.15 Å and the Cl-I bonds in crystals have generally larger local stretching force constant values than I-I bonds although they overlap in the range of 0.95–1.35 mdyn/Å. All X-I covalent bonds in crystal models are longer and weaker than their molecular counterparts in gas phase, which is a result of delocalization of lone pair electrons into the $\sigma^*(X-I)$ anti-bonding orbital [97] upon halogen bonding in crystals. Furthermore, we found that when the halogen bond acceptor is an ether group (G, H, J, K) the X-I bond is stronger than those when halogen bonds have N-oxides (L, M, N, O, P, Q) as acceptors.

It is worth noting that two distinct local stretching force constant-bond length relationships were observed for Cl-I and I-I bonds separately in Figure 3. This result is in line with our previous work where the essential difference between Badger-type rules for diatomic molecules and those for...
polyatomic compounds lies in that in the case of polyatomic molecules an individual curve can be expected for each bond type [61]. So, the important finding of this work is that rules worked out for the description of molecules in the gas phase seem to smoothly transition from bonding in molecules towards bonding in crystals.

Figure 4. Relationship between local stretching force constant $k_a$ and bond length $r$ for Cl-I···O (green) and I-I···O (purple) halogen bonds, respectively. Data points are shown with 4 shapes based on the acceptor molecule type, see Figure 1. An exponential function in the form of $k_a = a \cdot \exp(b \cdot r)$ was used to fit all 34 data points with $R^2 = 0.923$. Two data points are identified as outliers (encircled by blue hexagons) and $R^2 = 0.974$ is obtained for fitting the remaining 32 data points. The blue dashed curve shows the updated fitting function excluding two outliers.

Figure 4 collects two types of halogen bonds as Cl-I···O and I-I···O. Although the X atom in X-I donor bond is different, two halogen bonding atoms (i.e., I and O) are consistent and therefore the local stretching force constants for these two types of I···O bonds can be directly compared to provide detailed chemical insights. Such legitimacy for a direct comparison was also reflected by the uniformly strong correlation between force constants and bond lengths for these two types of halogen bonds.

We find that the Cl-I···O halogen bonds are generally shorter and stronger than I-I···O bonds, which is in line with the $\sigma$-hole theory [23–25]. Within either type of halogen bond, the intrinsic strength is mostly larger when the acceptor is N-oxide (L, M, N, O, P, Q) than those halogen bonds with ether group (G, H, J, K) as acceptors. This can be explained in terms of the charge transfer character (i.e., orbital interaction) [97] that the oxygen atom in N-oxide group could provide more lone pair electrons to be delocalized into the $\sigma^*$ (X-I) anti-bonding orbital for halogen bonding than the oxygen in an ether group.

3.3.2. Acceptor A–F

The acceptors A through F all provide a carbonyl C=O double bond to accept halogen bonding where A is acetone, B has two equivalent carboxamide groups, C and D are urea derivatives while E and F have carbonate ester groups.

Figure 3 shows that the X-I donor bonds in halogen bonding E1-1, E1-2 and E2 with acceptor E are the strongest for the crystal models studied in this work, although they are weaker than their...
molecular counterparts. This scenario can be explained in the following way. The intrinsic strength of a donor X-I bond is largely influenced by the extent to which the lone pair electrons of acceptor oxygen delocalize into the σ∗ anti-bonding X-I orbital. In addition, acceptor E has a delocalized π electron cloud throughout the whole plane and two strongly electronegative oxygen atoms distant from the C=O double bond can “pull” the oxygen lone pair charge density around the double bond towards the carbon atom. This polarization leads to a stronger C=O bond, as shown in Table 2 making the C=O bond of acceptor E the strongest among acceptors A through F, obviously suppressing the charge transfer character to the neighboring halogen bond. This is in line with Figure 4 showing that halogen bonds E1-1, E1-2 and E2 are relatively weak for either X-I···O type.

Table 2. Local stretching force constants of the C=O bonds for acceptor molecules A–F in descending order.

<table>
<thead>
<tr>
<th>No.</th>
<th>I-I···O=C</th>
<th>$k_n^a$ (mdyn/Å)</th>
<th>Cl-I···O=C</th>
<th>$k_n^a$ (mdyn/Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>E1-1</td>
<td>10.327</td>
<td>E1-2</td>
<td>9.932</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>9.755</td>
<td>E2</td>
<td>9.905</td>
</tr>
<tr>
<td>3</td>
<td>F-1</td>
<td>9.000</td>
<td>F-2</td>
<td>8.807</td>
</tr>
<tr>
<td>4</td>
<td>C1-1</td>
<td>8.945</td>
<td>C1-2</td>
<td>8.579</td>
</tr>
<tr>
<td>5</td>
<td>C2-1</td>
<td>8.937</td>
<td>C2-2</td>
<td>8.575</td>
</tr>
<tr>
<td>6</td>
<td>B2</td>
<td>8.431</td>
<td>D</td>
<td>8.561</td>
</tr>
<tr>
<td>7</td>
<td>B1-1</td>
<td>7.917</td>
<td>B1-2</td>
<td>7.651</td>
</tr>
</tbody>
</table>

The local mode frequencies $\omega_n^a$ for C=O bond stretching can be calculated by $\omega_n^a = \sqrt{k_n^a \cdot 497.6}$, where $k_n^a$ is in the unit of mdyn/Å and the resulting vibrational frequency $\omega_n^a$ is in the unit of cm$^{-1}$.

The I···O bond of halogen bonding A is the weakest among all halogen bonds studied in this work, whereas its donor I-I bond is relatively strong. By checking the corresponding crystal structure of A (see Figure 5), we found that both sides of each I$_2$ molecule form I···O halogen bonds and the oxygen atom of each acetone molecule accepts two I···O halogen bonds at the same time. In this situation, the lone pair electrons from one oxygen atom have to be shared by two donor I$_2$ molecules and the charge transfer into either I$_2$ donor is greatly reduced compared with the unshared situation (based on the physical picture of NBOs and their interaction), thus leading to weak halogen bonds and strong donor bonds. As the overall charge transfer from the acetone molecule is weak, the charge density of the C=O double bond region is less polarized towards oxygen atom and therefore leads to strong C=O bonds as shown in Table 2. This clearly shows that the crystal environment plays an important role which can be sensitively reflected by the local stretching force constants.

![Figure 5.](image-url)

The halogen bond B2 has the second lowest intrinsic strength as shown in Figure 4. The optimized crystal model for B2 reveals that each I$_2$ molecule is connected to two carboxamide groups, acting
as a linear bridge connecting two C=O groups with equal I···O distances. This unique structural arrangement weakens the I···O halogen bond because of the following. First, the two iodine atoms of I$_2$ are symmetrically equivalent in terms of charge density, which reduces the area of the σ-holes on both sides. Second, the lone pair electrons from the oxygen atoms on both sides of I$_2$ compete with each other in delocalizing into the σ$^*$ anti-bonding orbital of I$_2$, leading to reduced charge transfer for both I···O bonds. This explanation is rationalized by the fact that the C=O double bond of acceptor B is strongest in the case of B2 halogen bonding compared with B1-1 and B1-2 (see Table 2) as the I$_2$/ICl molecules in B1-1/B1-2 form only one I···O halogen bond.

3.3.3. Acceptor G–K

The molecules G through K possess ether groups as halogen bond acceptors. For the associated halogen bonds of either X-I···O type, we found a consistent strength ordering according to the acceptor type as G>J>H>(K) although their overall I···O bond strength is weak.

Such a consistent ranking can be explained by a through-bond interaction between oxygen atoms in these four acceptor molecules. The tetrahydrofuran acceptor G has only one oxygen atom, thus the charge transfer from this oxygen is not affected. In acceptor J, two acceptor oxygen atoms start to suppress the charge transfer of each other via their high electronegativity although they are relatively distant. The 1,4-dioxane acceptor H molecule has also two oxygen atoms to accept halogen bonding, but they are on para positions of a six-membered ring and much closer to each other than in molecule J. Therefore, the charge transfer from the oxygens in H is more suppressed than in J. K has four oxygen atoms with only two of them participating in halogen bonding in crystals. Due to the large number of oxygen atoms, the charge transfer from oxygen to halogen bonding is most strongly suppressed in K.

3.3.4. Acceptor L–Q

The acceptor molecules L through Q are heteroaromatic N-oxides, where the N$^+$-O$^-$ bond is linked to an aromatic ring [97].

We find that the donor X-I bonds associated with acceptor molecule O (O-1 and O-2) are the weakest, while the corresponding I···O bonds are the strongest among either X-I···O group. These findings can be rationalized in the following way. The nitrogen and oxygen atoms in an N$^+$-O$^-$ bond have comparable electronegativity, but oxygen attracts slightly more bonding electron density. One also needs to consider the substituent effects of the -N(CH$_3$)$_2$ amine group linked to an aromatic ring. As it is well-known that amine group has an ortho- para-directing effect for electrophilic aromatic substitution reactions [130–133], so the nitrogen atom of the N$^+$-O$^-$ bond in the para position with regard to the amine group in acceptor O can attract more electron density from the aromatic ring while loosening the attraction of the NO bonding electrons. Therefore, the oxygen atom of the N$^+$-O$^-$ bond can accumulate electron density via its large electronegativity. This leads to the weakest N$^+$-O$^-$ bond for acceptor O as shown in Table 3. The charge transfer from the oxygen lone pair electrons into the donor I$_2$/ICl molecule is greatly enhanced resulting in strong halogen bonds with weak donor bonds.

**Table 3.** Local stretching force constants of the N$^+$-O$^-$ bonds within acceptor molecules L–Q in ascending order.

<table>
<thead>
<tr>
<th></th>
<th>I-I···O$^-$·N$^+$</th>
<th>k$^d_n$ (mdyn/Å)</th>
<th>Cl-I···O$^-$·N$^+$</th>
<th>k$^d_n$ (mdyn/Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>O-1</td>
<td>4.866</td>
<td>O-2</td>
<td>4.738</td>
</tr>
<tr>
<td>2</td>
<td>Q</td>
<td>4.916</td>
<td>P-2</td>
<td>4.785</td>
</tr>
<tr>
<td>3</td>
<td>F-1</td>
<td>4.944</td>
<td>M</td>
<td>4.867</td>
</tr>
<tr>
<td>4</td>
<td>L-1</td>
<td>5.061</td>
<td>L-2</td>
<td>4.896</td>
</tr>
<tr>
<td>5</td>
<td>N-1</td>
<td>6.829</td>
<td>N-2</td>
<td>5.297</td>
</tr>
</tbody>
</table>

The local mode frequencies $\omega_n$ for N$^+$-O$^-$ bond stretching can be calculated by $\omega_n^2 = k_n^d / 476.8$, where $k_n^d$ is in the unit of mdyn/Å and the resulting vibrational frequency $\omega_n$ is in the unit of cm$^{-1}$. 

Molecules 2020, 25, 1589 11 of 22
Another interesting observation is that N-1 and N-2 halogen bonds are the weakest for either X-I···O type when acceptors are N-oxides. This is mainly caused by the electronic structure of acceptor molecule N. First, molecule N has a pyridine-like structure as shown in Figure 1. The nitrogen atom in a pyridine molecule acts like a meta-directing group for electrophilic aromatic substitution reaction [132], similarly the nitrogen atom (position 4) within acceptor molecule N makes the other nitrogen atom (position 1) of N⁺−O⁻ bond slightly electron-deficient. This helps pulling the bonding electrons of the N⁺−O⁻ bond towards the nitrogen atom making this N⁺−O⁻ bond stronger. This is in line with the fact that the N⁺−O⁻ bonds within the acceptor molecule N are the strongest in Table 3. Then one can infer that the charge transfer from the oxygen atom will be reduced and therefore the corresponding halogen bond is weakened.

The substituent effect in acceptor molecules L, M and P on their N⁺−O⁻ bonds is more similar to that in acceptor molecule O than in acceptor molecule N, as revealed by corresponding local stretching force constants being larger but close to the force constant values for acceptor molecule O (see Table 3). This explains why the local stretching force constants of I···O bonds associated with these three acceptor molecules are between those of I···O bonds associated with acceptor molecule O and acceptor molecule N within either X-I···O type.

This work includes five I-I···O halogen bonds overlapping the recent work by Rosokha and co-workers [96,97] including P-1(iQnO), L-1(2MePyO), O-1(Me2NPyO), N-1(PyrazO) and Q(ClQnO) (abbreviations in parentheses are taken from Reference [97]). The ordering for intrinsic strength of the first four halogen bonds (see Figure 4) is consistent with the ordering for binding energies of dimer complexes, except for P-1 and L-1. However, these two halogen bonds have a binding energy difference of only 0.2 kcal/mol [97] in line with our local mode analysis also showing marginal difference in local stretching force constant of 0.033 mdyn/Å. This means the binding energies calculated from dimer models with different acceptors qualitatively predict the strength of the halogen bonds in these four crystals. However, the I···O bond of halogen bonding Q is much weaker than expected although corresponding binding energy between diiodine and acceptor Q in a dimer complex is comparable to those for acceptors L and P according to Rosokha’s calculations [97]. Moreover, the local stretching force constant of the N⁺−O⁻ bond in acceptor Q indicates that this halogen bond is likely to have comparable strength as P-1, as shown in Table 3.

3.3.5. Outliers

Figure 3 showing the relationship between local stretching force constant and bond length for donor X-I bonds has no obvious outliers deviated from the fitted curves. However, there are a few outliers around the fitted curve for I···O bonds in Figure 4. Besides Q and N-1 identified as two outliers by the deviation criterion of 1.5σ (σ is the standard deviation of the residuals after fitting), we found halogen bonds A, H1 and F-1 also have relatively large deviation from the best fitted curve.

In the case of halogen bonding A, the I···O bond force constant is lower than what the fitted curve predicts by 0.05 mdyn/Å. As pointed out in Section 3.3.2, in its crystal structure with acetone as the acceptor molecule each C=O double bond accepts two halogen bonds instead of one (see Figure 5). This is the major reason halogen bond A becomes an outlier given its peculiar bifurcated halogen bonding geometry.

In the optimized crystal structure of halogen bond F-1, one iodine atom of I₂ forms the F-1 halogen bond, while the other side of the I₂ molecule is pointing towards an oxygen atom on the six-membered ring of another acceptor molecule F with the I···O distance of 3.82 Å. As the I-I bond is nearly perpendicular to the plane of this distant carbonate ester group which is a π conjugated system, the charge transfer into I₂ might be non-trivial given the large radius of the valence shell of iodine atom. Moreover, the electron-rich carbonate ester group with three oxygen atoms could be an ideal halogen bond acceptor to form the second halogen bonding with I₂. This explains why halogen bond F-1 is weaker than what can be expected from the fitted curve, because the other side of the I₂ molecule is affected by another seemingly distant acceptor molecule. One might doubt why the halogen bond
F-2 is not affected to be an outlier, we argue that the chlorine atom of the donor Cl-I bond in F-2 is not a good halogen bond donor according to the σ-hole theory and the charge transfer from oxygen towards chlorine is more difficult given the Cl···O distance as 3.92 Å. The halogen bond F-2 experiences little influence from the other side of ICl donor and therefore its I···O bond is not an outlier.

For halogen bond N-1 identified as an outstanding outlier in Figure 4, the other side of each I₂ molecule forms a rather short (maybe stronger) I···N halogen bonding with another acceptor molecule N (see Figure 6). This I···N interaction strongly outweighs the I···O bonding and makes N-1 an outlier in Figure 4. We notice that the local stretching force constant of the N-1 halogen bond is larger than what the fitted curve predicts by more than 0.1 mdyn/Å. This strengthening is caused by the strong polarization of the push-pull effect [71] (which enhances the charge transfer from oxygen to I₂ molecule) arising from the infinite alternating chain of I₂ and acceptor molecule N [97].

Figure 6. Optimized crystal model structure showing the I···O halogen bonding N-1 (blue dashed lines) between diiodine and acceptor molecule N within the infinite alternating chain. The I···N halogen bonding is shown with orange dotted lines.

In the optimized crystal structure for halogen bond H₁, we also found an infinite alternating chain consisting of I₂ and 1,4-dioxane acceptor H. As shown in Figure 7, each oxygen atom of the dioxane ring forms a halogen bond with an identical I···O distance and an I₂ molecule donates two equivalent halogen bonds connecting two dioxane rings. We believe this infinite chain structure might render H₁ as an outlier above the fitting curve. However, there is no push-pull effect [71] in this case, because both I₂ and acceptor molecule H have no dipole moments within the infinite chain structure. The unusual strengthening for H₁ halogen bond is explained as follows. We first consider a ternary segment of H···I₂···H within the infinite chain. Two oxygen atoms within molecules H on both sides of this ternary complex tend to attract electrons caused by their strong electronegativity. This through-bond interaction suppresses the charge transfer from two inner oxygens into the central I₂ molecule. After we put this ternary segment back into the infinite chain structure, those two flanking oxygens in that ternary complex also form halogen bonding and their suppressive effect on the said charge transfer is reduced. Therefore, the I···O bond in H₁ is strengthened by the infinite alternating chain.

Figure 7. Optimized crystal model structure showing the halogen bonding H₁ between diiodine and acceptor molecule H within the infinite alternating chain.

One might raise the question why the halogen bond B₂ is not an outlier although it has similar infinite alternating chain structure as H₁ without push-pull effect. Based on the above analysis for H₁, this might be related to the fact that two oxygen atoms of carboxamide groups within acceptor B are more distant than in acceptor H. In this way, the halogen bonding within a ternary model B···I₂···B is similar to halogen bonding B₂ in crystals (with infinite alternating chain) due to minor influences from the two flanking carboxamide groups in the ternary complex.

The I···O bond of halogen bonding Q is the most significant outlier in Figure 4. Unlike other outliers discussed above, it has a much shorter bond length of ca. 2.5 Å, indicating that there might be a different mechanism making it an outlier. By checking the crystal structure containing halogen
bond Q, we found there exists a highly ordered I2 network between two rows of acceptor molecules Q (see Figure 8). For each halogen bonding Q, the other side of donor I2 forms two I⋯I halogen bonds of equal distance, and the resulting zigzag pattern is repeated as an infinitely long one-dimensional network [96]. These peculiar interactions between diiodine molecules are the major factor which greatly weakens the I⋯O bonds. Similar interaction networks between diiodines were also found in recent work on metal-organic framework (MOF) structures designed for I2 adsorption [134]. A more detailed and systematic investigation into this problem is currently in progress.

Figure 8. Optimized crystal model structure showing the I⋯O halogen bonding Q (blue dashed lines) between diiodine and acceptor molecule Q. The orange dotted lines show the I2⋯I2 interaction network.

3.3.6. Crystal Packing Effect

By taking a bird’s eye view of Figure 4, we found that when the donor X-I bond forms the second halogen bond with a distance less than 3.0 Å as in the case of the I⋯O bonds A, B2, H1 and N-1 (as listed in Table 1), these halogen bonds have relatively small local stretching force constants equal or less than 0.2 mdyn/Å. Furthermore, these are the only halogen bonds with a bond length larger than 2.75 Å. Moreover, halogen bonds Q and F-1 are also characterized by small local stretching force constants less than 0.2 mdyn/Å. The above six halogen bonds have one thing in common, their donor I2 molecule forms a second halogen bond. The local stretching force constant less than 0.2 mdyn/Å for the halogen bond E1-1 is caused by the strong substituent effect in acceptor molecule E (as discussed in Section 3.3.2).

Removing all data points with local stretching force constants less than 0.2 mdyn/Å from Figure 4 except E1-1, we found all remaining data points stay quite close to the fitted curve between bond length and local stretching force constant, rigorously following the Badger’s rule. In addition, we found all Cl-I⋯O halogen bonds stay close to the fitted curve because the chlorine atom is not a good donor to form a second halogen bond. This means that the impact from crystal packing on the I⋯O halogen bonding will be non-trivial if the donor I-I bond forms a second halogen bonding with oxygen, nitrogen or diiodine.

However, the causal relationship between the weak I⋯O bond, and the existence of the second halogen bond arising from crystal packing is complicated and it may vary in different situations:

- In the case of halogen bonding Q, reference DFT calculations on dimer models by Rosokha and co-workers have demonstrated that the acceptor molecule Q and diiodine could form a halogen
bond as strong as the halogen bond between acceptor molecule P and diiodine [97]. However, in the crystal structure, the I₂ ··· I₂ interaction network dominates the crystal packing and it directly leads to weaker I ··· O halogen bonds than expected;

- In the case of halogen bonding N-1, the I ··· O halogen bond is weak because the molecule N is a poor acceptor. In this case, the crystal packing must enforce the structure to form a second halogen bonding for compensation to stabilize the whole system.

These subtle effects cannot be considered by calculations on dimer models, showing the necessity of in situ quantification of chemical bond strength in crystal structures.

4. Conclusions

In this work, we accurately quantified and systematically ordered the intrinsic strength of 34 X-I ··· OA halogen bonds (where X=Cl or I) from 20 different crystal structures by employing our periodic local vibrational mode theory [85]. The most important findings include the following:

1. We observed strong correlations between bond length and local stretching force constant for both X-I donor bonds (i.e., I-I and Cl-I) and I ··· O halogen bonds, which suggests that the generalized Badger’s rule (based on local stretching force constants [61]) originally derived from molecules is also valid for both covalent bonds and non-covalent interactions in crystals;

2. The local stretching force constants for I ··· O halogen bonds (Figure 4) span a wide range of 0.1–0.8 mdyn/Å, demonstrating the impressive tunability in bond strength even within a specific type of halogen bonding;

3. Our results for some I-I ··· O halogen-bonded crystals previously investigated experimentally and computationally by Rosokha and co-workers [96,97] clearly show the potential of the periodic local mode analysis leading to new deeper insights:

   (a) Rosokha’s statement that “besides electrostatic, molecular orbital interactions play a substantial role in XB between diiodine and N-oxides” can be expanded to the I ··· O halogen bonding between dihalogen X-I and any acceptor oxygen atom with lone pair electrons. This generalization is based on the strong correlations between bond length and force constant for both donor bonds and halogen bonds in Figures 3 and 4;

   (b) In comparison to the dimer model systems used for DFT calculations by Rosokha and co-workers, our first-principle calculations on crystal models could include the crystal packing effects. On one hand, the overall lattice structure (including donor/halogen bond lengths) of molecular crystals is a direct result of crystal packing. On the other hand, the impact from crystal packing on X-I ··· OA halogen bonding varies in different ways:

      i. In those cases where the X atom of the X-I halogen donor molecule has no close contact to neighboring atoms in the crystal or is only stabilized by an interaction with the π cloud of an adjacent aromatic ring, the I ··· O halogen bonds behave like covalent bonds and rigorously follow a local stretching force constant-bond length relationship;

      ii. When both sides of X-I donor are involved in halogen bonding (only observed for I₂ not for Cl-I), the I ··· O halogen bond is weak. If the acceptor oxygen atom has to accept two halogen bonds simultaneously (e.g., halogen bonding A), the halogen bond strength is even lower. If the X atom forms a non-trivial halogen bond which obscures the target I ··· O halogen bond (e.g., N-1 and Q), the target halogen bond largely deviates from the ideal force constant-bond length relationship.

      iii. However, independent of the engagement of the X atom in additional halogen bonding associated with crystal packing, the donor bonds rigorously follow an ideal force constant-bond length relationship (Figure 3) due to their covalent bond nature.
(c) Via delicate analysis in terms of substituent effect and other electronic structure factors in acceptor molecules, we are able to explain the majority of the variations in the intrinsic strength of both donor bonds and halogen bonds in X-I⋯O bonding. Furthermore, the local stretching force constants of the adjacent O-A bonds in acceptor molecules could complement our findings;

(d) In determining the strength of I⋯O halogen bonds, halogen atom X within X-I donor plays a decisive role as the weakest Cl-I⋯O halogen bonds are comparable to the strongest I-I⋯O halogen bonds. The acceptor molecules with different capabilities for $n \rightarrow \sigma^*(X-I)$ charge transfer are the second important factor for determining the I⋯O bond strength. Last but not least, the existence of the second halogen bonding via the X atom of the donor X-I bond can substantially weaken the target I⋯O halogen bond in crystals.

4. We discovered for the first time a linear correlation for X-I donor bonds along with a quadratic correlation for I⋯O halogen bonds between experimental and calculated bond lengths. One application based on this is to estimate the local stretching force constant of either the X-I donor bond or I⋯O halogen bond of X-I⋯OA halogen bonding directly for a newly resolved crystal structure via the correlations in Figures 2–4 given that no second halogen bond exists with atom X of the X-I donor molecule. Such relationship may also hold for other types of non-covalent interactions;

5. All calculations in this work were based on projector-augmented wave (PAW) basis. The resulting chemically sound results of local stretching force constants demonstrate that our periodic local mode analysis is not limited to atomic orbital (AO)-based computational results [85], it is generally applicable and independent of how the wavefunction is obtained. Using periodic local vibrational mode theory as a unique tool to investigate the intrinsic strength of other types of halogen bonding (and non-covalent interactions) in crystals/materials will be one of our current and future directions.

Supplementary Materials: The following are available online at http://www.mdpi.com/1420-3049/25/7/1589/s1, Cartesian coordinates of all optimized primitive cell structures in this work.

Author Contributions: Conceptualization, Y.T. and E.K.; methodology, Y.T. and W.Z.; software, Y.T. and W.Z.; validation, Y.T. and Y.Q.; formal analysis, Y.T. and Y.Q.; investigation, Y.T.; writing—original draft preparation, Y.T. and Y.Q.; writing—review and editing, Y.T., Y.Q., S.N., S.Y. and E.K.; visualization, Y.T.; supervision, E.K.; funding acquisition, E.K., W.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially supported by the National Science Foundation (Grant CHE 1464906). W. Z. was supported by National Natural Science Foundation of China (Grant No. 21673175) and the Double First-Class University Construction Project of Northwest University.

Acknowledgments: We thank SMU for providing generous computational resources. Y.T. thanks Vytor Oliveira for thought-provoking discussions by sharing his expertise on halogen bonding. We thank Daniel Sethio for conducting a workshop on first-principle calculations for the CATCO group in December 2019.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations
The following abbreviations are used in this manuscript:

**X** Halogen

**ESP** Electrostatic Potential

**SAPT** Symmetry-Adapted Perturbation Theory

**CSD** Cambridge Structural Database

**COD** Crystallography Open Database

**VASP** Vienna Ab initio Simulation Package

**PES** Potential Energy Surface

**NBO** Natural Bond Orbital

**DFT** Density Functional Theory
References


7. Meyer, F.; Dubois, P. Halogen Bonding at Work: Recent Applications in Synthetic Chemistry and Materials Science. *CrystEngComm* 2013, 15, 3058–3071. [CrossRef]


53. Kraka, E.; Cremer, D. Dieter Cremer’s Contribution to the Field of Theoretical Chemistry. *Int. J. Quantum Chem.* 2019, 119, e25849. [CrossRef]


58. Zou, W.; Cremer, D. $C_2$ in a Box: Determining its Intrinsic Bond Strength for the $X^1\Sigma^+_g$ Ground State. *Chem. Eur. J.* 2016, 22, 4087–4097. [CrossRef] [PubMed]


83. Oliveira, V.; Cremer, D. Transition from Metal-Ligand Bonding to Halogen Bonding Involving a Metal as Halogen Acceptor: A Study of Cu, Ag, Au, Pt, and Hg Complexes. Chem. Phys. Lett. 2017, 681, 56–63. [CrossRef]


© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).