

UC Riverside

UC Riverside Electronic Theses and Dissertations

Title

Advanced Computational Methods for Ground and Excited States to Study Complex Systems

Permalink

<https://escholarship.org/uc/item/1hb873bk>

Author

Ali, Zulfikhar Ahmad

Publication Date

2022

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Advanced Computational Methods for Ground and Excited States to Study
Complex Systems

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

Doctor of Philosophy

in

Physics

by

Zulfikhar A. Ali

September 2022

Dissertation Committee:

Dr. Bryan M. Wong, Chairperson
Dr. Hai-Bo Yu
Dr. Shan-Wen Tsai

Copyright by
Zulfikhar A. Ali
2022

The Dissertation of Zulfikhar A. Ali is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

Thank you to my advisor, Dr. Bryan M. Wong, for your patience, guidance, and support. His candid approach to research and science is a source of inspiration. I have significantly benefited from your wealth of knowledge and meticulous editing. I am incredibly grateful that you took me on as a student and continued to have faith in me over the years. Thanks for advancing my writing skills.

Thank you to my committee members, Dr. Hai-bo Yu and Dr. Shan-wen Tsai. Your encouraging words and thoughtful feedback have been significant to me.

Thank you to my family for constantly listening to me rant and talking things out and for the sacrifices they've made for me to pursue a Ph.D.

I gratefully recognize the help of the B.M.W. lab postdocs, Chao Lian, Sharma S.R.K.C. Yamijala, Fredy W. Aquino, Kota Hanasaki, Ravindra Shinde, Min Choi, and Mahmut Okyay, for their extensive knowledge and expertise in developing and implementing density functional based theories, specifically, real-time time-dependent density functional theory. Without the help of these postdocs, who graciously showed me the ropes of development and implementation, I doubt I would have gotten as far as I did in my program.

Special thanks to my undergraduate professors, Dr. Björn Birnir and Dr. Joshua M. Deutsch from UCSB and UCSC, respectively, for the essential roles played during my college career.

To all that has worked with me during my Ph.D. career, I'll forever be grateful for the time you set aside to help.

I do not immensely thank the peer reviewers who gave me a hard time with their biased reviews, but I am thankful to them for taking the time to review my manuscripts. I still got my papers published!

Finally, I acknowledge the generous financial support from the United States government and the University of California.

DEDICATED TO

My cousin, **Zahid M. Arshad**.

My trusted friend, Adderall, for the daily support.

My family because they'll annoy me if I don't say something.

My grandfather, **Sheikh K. Maqbool**, who'll never read this work.

Finally, to myself for putting in the time and effort required for a piece of paper.

ABSTRACT OF THE DISSERTATION

Advanced Computational Methods for Ground and Excited States to Study Complex Systems

by

Zulfikhar A. Ali

Doctor of Philosophy, Graduate Program in Physics
University of California, Riverside, September 2022
Dr. Bryan M. Wong, Chairperson

The real-time time-dependent density functional theory (rt-TDDFT) approach, which is complementary to the more traditional linear-response TDDFT (lr-TDDFT), propagates the electron density in real time for studying the ground and excited states. Researchers use rt-TDDFT to study electron dynamics in real time. Moreover, rt-TDDFT can treat non-linear effects, which is especially useful for experimentalists studying external field effects in complex systems. This dissertation presents computational methods for studying a complex system's ground and excited states. We start with ground state calculations, via density functional theory (DFT), for studies on Cyclodextrins as a catalyst, binary compound convex hull, and transition states. Then we go into ground state studies with nuclear motion using the Born-Oppenheimer Molecular Dynamics. An application of lr-TDDFT on TiSe_2 follows this. We transition to excited state studies by introducing our rt-TDDFT formalism. We validate the implementation with benchmark tests for essential elements. Finally, we demonstrate our implementation capabilities and apply them to practical systems. The first application simulates the attosecond transient absorption spectroscopy for charge transfer

and polarization switching in BaTiO₃. The second application involves molecular dynamics, a nonlinear process, for photo-induced degradation mechanisms of perfluorooctanoic acid (PFOA). By explicitly accounting for non-adiabatic excited-state interactions in solvated PFOA, we show that these photo-induced excitations enable a charge-transfer process that polarizes the C–F bond, resulting in a dynamic dissociation on a femtosecond time scale. Ultimately, this dissertation emphasizes the importance of quantum simulations for studying excited states of molecular and extended systems.

Contents

List of Figures	xiii
List of Tables	xxiii
1 Introduction	1
1.1 Outline	3
2 A QM Study of Biomimetic Catalysis of Diels-Alder Reactions Using Cyclodextrins	6
2.1 Introduction	7
2.2 Results and Discussion	10
2.2.1 Explanation of Enhanced Catalysis by QM Calculations	10
2.3 Computational Methods	16
2.3.1 Quantum Mechanics Calculations with the DFT Method	17
2.4 Conclusion	19
3 The Diamine Cation is not a Chemical Example where DFT fails	20
3.1 Introduction	21
3.2 Results and Discussion	22
3.3 Conclusion	28
4 Harnessing Semi-Supervised Machine Learning to Predict Per- and Polyfluoroalkyl Substances (PFASs) Bioactivities	30
4.1 Introduction	31
4.2 Methodology	33
4.3 Results and Discussion	35
4.3.1 Unsupervised vs. semi-supervised machine learning	35
4.3.2 Semi-supervised metric learning	36
4.3.3 Interactions between PFASs and targets	38
4.3.4 Bioactivity predictions on OECD dataset	39
4.4 Conclusion	40

5	Efficient Predictions of Formation Energies and Convex Hulls from Density Functional Tight Binding Calculations	42
5.1	Introduction	43
5.2	Theory and Methodology	45
5.2.1	DFTB	45
5.2.2	DFT	47
5.2.3	Structure Generation with CASM	49
5.2.4	Formation Energy	50
5.3	Computational Details	51
5.3.1	CASM	52
5.3.2	DFT Calculations	54
5.3.3	DFTB Calculations	55
5.3.4	Formation Energy	56
5.4	Results and Discussion	56
5.4.1	SiC	57
5.4.2	ZnO	60
5.4.3	Efficiency Analyses for DFT and DFTB	61
5.5	Conclusion	65
6	Acceleration vs Accuracy: Influence of Basis Set Quality on Dynamics Predicted by Ab Initio MD	67
6.1	Introduction	68
6.2	Computational Details	71
6.3	Results and Discussion	73
6.3.1	Low-Energy (0.6 eV) Collision	75
6.3.2	High-Energy (6 eV) Collision	77
6.3.3	Basis Set Effect on the Adsorption Energies	82
6.3.4	Electron Density Difference Maps	83
6.3.5	Results with Other Software Packages	86
6.4	Conclusion & Outlook	88
7	Charge Density Wave Hampers Exciton Condensation in 1T-TiSe₂	92
7.1	Introduction	93
7.2	Methodology	96
7.3	Results and Discussion	101
7.4	Conclusion	106
8	Implementation of Real-Time TDDFT for Periodic Systems in the Open-Source PySCF Software Package	107
8.1	Introduction	108
8.2	Methodology	110
8.2.1	Real-time propagation scheme	110
8.2.2	Velocity gauge formulation	114
8.2.3	Optical absorption spectrum	115
8.3	Results	117

8.3.1	Optical Absorption Spectra	117
8.3.2	Real-time dynamics	121
8.4	Discussion and Conclusion	124
9	Indirect but Efficient: Laser-Excited Electrons can drive Ultrafast Polarization Switching in Ferroelectric Materials	127
9.1	Introduction	128
9.2	Computational Details	130
9.3	Results and Discussion	131
9.4	Conclusions	139
10	Photo-induced Degradation of PFASs: Excited-State Mechanism from RT-TDDFT	141
10.1	Introduction	142
10.2	Computational Details	144
10.3	Results and Discussion	146
10.4	Conclusion	151
11	Conclusions	153
A	The Diamine Cation is not a Chemical Example where DFT fails	205
A.1	DMP Coords	206
B	Harnessing Semi-Supervised Machine Learning to Automatically Predict Bioactivities of Per- and Polyfluoroalkyl Substances (PFASs)	212
B.1	Data Collection	213
B.2	Discussion of Unsupervised Learning Methods	213
B.3	Model Selection	215
B.4	Brief Discussion of Semi-Supervised Metric Learning	215
B.5	Molecular Docking Calculations	216
B.6	Unsupervised Machine Learning Results	217
C	Acceleration vs Accuracy: Influence of Basis Set	228
C.1	Additional Computational Details	229
C.1.1	Generation of the ZnO Slab Along the (10 $\bar{1}$ 0) Direction	229
C.1.2	Various Convergence Tests	230
C.1.3	Generation of Adsorbate Configurations	233
C.1.4	Additional Tests for Adsorption Energy Calculations	235
C.2	Additional Tests with CP2K	238
C.2.1	Basis Sets	238
C.2.2	Dispersion Effect	240
C.2.3	Planewave Cutoff Effect on Adsorption Energy (eV)	241
C.2.4	Structural Relaxation Effect on Basis-Set Induced Adsorption Energy Changes	241
C.3	NBO Analysis	244
C.4	Mulliken Charge Analysis of the 6 eV Collision Trajectories	245

C.5	Variation in the Adsorption Energy Across Software Packages	248
C.6	Effect of Exchange-Correlation Functionals	249
C.6.1	BOMD Simulation Results with PBE Functional (without Grimme's Dispersion)	249
C.6.2	BOMD Simulation Results with LDA Functional	251
C.7	BOMD Simulations with Changes in the Impact Position	256
D	Charge Density Wave Hampers Exciton Condensation in 1<i>T</i>-TiSe₂	259
E	Implementation of Real-Time TDDFT for Periodic Systems in the Open- Source PySCF Software Package	264
E.1	Nuclear Dynamics	265

List of Figures

2.1	Illustration of the hydrogen bonds between the carbonyl groups on maleimides and the secondary hydroxyl groups on cyclodextrins. N-cyclohexylmaleimide is used as an example. The blue trapezoid represents the cyclodextrin, and the dotted lines represent the hydrogen bonds.	8
2.2	The Diels-Alder reaction of 9-anthracene-methanol (1) with N-cyclohexyl maleimide (2a), with β -cyclodextrin or methyl- β -cyclodextrin as a catalyst.	9
2.3	Reaction paths obtained with DFT. In all plots, the curves are normalized to set the reactant energies at zero. (a) : reaction paths for the reaction catalyzed by β -CD, with and without hydrogen bonds. (b) : reaction paths for the reaction catalyzed by dimethyl- β -CD, with and without hydrogen bonds. (c) : reaction paths for uncatalyzed reaction between compounds 1 and 2a , reaction catalyzed by β -CD, and reaction catalyzed by dimethyl- β -CD. No hydrogen bonds exist between compound 2a and the cyclodextrin.	10
2.4	The transition states of five reactions optimized at the ω B97XD/6-31G(d) level of theory. (1) Transition state structure in the non-catalyzed reaction between compounds 1 and 2a . Atom names used in Table 2.1 are labeled in this structure. (2) Transition state structure in the reaction catalyzed by β -CD, with hydrogen bonds between the two carbonyl groups of 2a and β -CD. (3) Transition state structure in the reaction catalyzed by β -CD, without the hydrogen bonds. (4) Transition state structure in the reaction catalyzed by dimethyl- β -CD, with hydrogen bonds between the two carbonyl groups of 2a and dimethyl- β -CD. (5) Transition state structure in the reaction catalyzed by dimethyl- β -CD, without the hydrogen bonds.	12
2.5	Molecular structure of β -cyclodextrin. The shape of the cyclodextrin macrocycle can be described as a truncated cone, with a narrow rim presenting primary hydroxyl groups and a wide rim presenting secondary hydroxyl groups on the glucose residues.	17

3.1	Calculated potential energy curve between the localized and delocalized state of the dimethylpiperazine cation. The PZ-SIC and M06-HF potential energy curves were obtained from Ref. [57], and the CCSD_CCSD(T)-SP and MP2_CCSD(T)-SP legend labels denote single-point energy calculations that were carried out with the CCSD(T) method using geometry-optimized structures obtained with CCSD and MP2, respectively. All three wavefunction-based approaches (CCSD, CCSD_CCSD(T)-SP, and MP2_CCSD(T)-SP) are in agreement by producing an extremely small energy barrier (< 0.01 eV), with the CCSD_CCSD(T)-SP and MP2_CCSD(T)-SP curves in close agreement with the M06-HF DFT calculations.	22
3.2	Comparison of transition-state geometries for the diamine cation optimized with the MP2, CCSD, and PZ-SIC computational methods.	26
3.3	Axial and side views of transition-state geometries for the diamine cation optimized with the CCSD and PZ-SIC computational methods.	28
4.1	Machine-learning-based workflow for QSAR construction to predict bioactivity of PFASs.	33
4.2	CF dataset molecule distribution using semi-supervised metric learning. Each point represents either a bioactive (red circular edges) or an inactive (light blue circular edges) molecule towards (a) CYP2C9, (b) CYP3A4, (c) CYP2D6, and (d) ATXN. The olive green-filled circles represent molecules having the substructure depicted in the plot; i.e., (a, b) ester groups, (c) phenylprimidyl groups, and (d) 4-benzyl-2-(4-fluorophenyl)-1,2-thiazole. The pink-filled circles in (c) represent molecules with phenylethanone. The percentage value represents the ratio of the number of bioactive molecules within the identified substructure. Table B.3 lists the predicted substructures for specific targets.	37
4.3	Clustering of molecules predicted with unsupervised learning (dimension reduction) on CF datasets containing (a) chemical structures and (b) chemical structures and binding affinities with CYP2C9. Each point represents a molecule that is either bioactive (red) or inactive (blue) towards CYP2C9.	39
4.4	(a) OECD dataset classified by PC t-SNE and clustered based on the k-means clustering method. The orange and yellow dots represent ester-containing molecules. The colors closer to red (yellow) represent a higher (lower) concentration of bioactive molecules. (b) PFAS molecules included in the OECD list are grouped into 40 clusters. Each point represents a molecule, and clusters 13, 25, and 39 denote a high ratio of ester-containing groups	41
5.1	Flowchart of the structure generation algorithm used in the CASM software package to enumerate structures.	49
5.2	General workflow in our DFTB+CASM implementation for calculating formation energies and the convex hull.	53

5.3	Primitive unit cells used to generate the various supercells and configurations. Using the CASM software package, we generate all symmetrically distinct configurations with the stoichiometry Si_2C_2 . Panels (a), (b), and (c) show the primitive cell in different orientations. Light green and grey atoms represent Si and C, respectively.	54
5.4	Formation energy convex hull for the Si-C binary system computed with DFT and DFTB. Panel (a) shows the DFT formation energy. Panel (b) shows the DFTB formation energy calculated using the SKfIV SK files. Each point corresponds to a different crystal structure.	57
5.5	Comparison of SiC structural parameters (located on the convex hull) after optimization with (a) DFT and (b) DFTB (SKfIV), visualized along lattice vector \vec{c} . Bond lengths are in Angstroms, and light green and grey atoms represent Si and C, respectively.	58
5.6	Formation energy of the ZnO binary system obtained from DFT and DFTB. Panel (a) shows the DFT formation energy, and panel (b) shows the DFTB formation energy calculated using the znorg-0-1 SK files. Each point corresponds to a different crystal structure.	60
5.7	Comparison of DFT and DFTB wall times per SCF iteration step as a function of the number of electrons in various ZnO configurations. Each vertical bar value was calculated by averaging multiple configurations having the same number of electrons.	62
5.8	Comparison of wall times for geometry optimization of various $\text{Si}_2\text{C}_{2(1-x)}$ compositions calculated via DFT and DFTB. Panel (a) shows the wall time for optimizing each configuration using DFT. Panel (b) shows the ratio between the DFT and DFTB wall times for a geometry optimization of each configuration. Each point corresponds to a different crystal structure. . . .	63
5.9	Comparison of wall times for geometry optimization of various $\text{Zn}_2\text{O}_{2(1-x)}$ compositions calculated via DFT and DFTB. Panel (a) shows the wall time for optimizing each configuration using DFT. Panel (b) shows the ratio between the DFT and DFTB wall times for a geometry optimization of each configuration. Each point corresponds to a different crystal structure. . . .	64
5.10	Comparison of DFT and DFTB total wall times for geometry optimization of all SiC and ZnO configurations. The DFTB calculations used the SKfIV and znorg-0-1 SK files for SiC and ZnO, respectively.	65

6.1	(a) Initial configuration used to study the interaction of an H ₂ O molecule colliding with a ZnO slab preadsorbed with four CO molecules. The four CO molecules are numbered to distinguish the 4th CO molecule (closest one to the impact region) from the others. The blue, red, gray, and white colors are used to represent the Zn, O, C, and H atoms, respectively. Oxygen atoms of the CO and H ₂ O are colored in orange to distinguish them from the oxygen atoms of the ZnO lattice. Panels (b) and (c) show the side and top views of how the CO molecules in the configuration (a) have changed during a 4.4 ps NVT run performed with the TZV2P basis sets at 300 K. In both (b) and (c), the cyan, lime, and blue trajectories show the change in the positions of the 1st, 2nd, and 3rd CO molecules, respectively, for the entire trajectory.	71
6.2	Panels (a) and (b) show how the configuration in Figure 6.1a has changed for a H ₂ O molecule with 0.6 eV of incident energy after a 3 ps NVE simulation performed with the DZVP and TZV2P basis sets, respectively. (c) Relative change in the positions of the 4th CO and H ₂ O molecules along the a-direction of the lattice.	72
6.3	Panel (a) shows the changes in total (E), kinetic (T), and potential (V) energies of the entire system for the 0.6 eV collision simulated with the TZ basis. Panel (b) shows the changes in the kinetic energy of each of the fragments during the same NVE run. The inset in (b) shows the changes in kinetic energies of the fragments during the first 100 fs of the simulation.	73
6.4	Panels (a) and (b) show how the configuration in Figure 6.1a has changed for the H ₂ O molecule with 6 eV of incident energy, after a 5 ps NVE simulation performed with the DZVP and TZV2P basis sets, respectively. (c) Relative change in the positions of the 4th CO and H ₂ O molecules along the c-direction of the lattice.	74
6.5	Panel (a) shows the changes in total (E), kinetic (T), and potential (V) energies of the entire system for the 6 eV collision simulated with the TZ basis. Panel (b) shows the changes in the kinetic energy of each of the fragments during the same NVE run. The inset in (b) shows the changes in kinetic energies of the fragments during the first 100 fs of the simulation.	75
6.6	Similar results in the rise of the KE of the 4th CO using both the DZ and TZ basis during the first 100 fs of the simulation.	77
6.7	The first two columns show electron density differences obtained for a single CO molecule adsorbed on a ZnO slab with the DZVP (DZ) and TZV2P (TZ) basis sets, respectively. Here, each of these columns is obtained by subtracting the total electron density of the entire system from its individual components (i.e., $\rho[\text{diff}] = \rho[\text{ZnO} + \text{mol}] - \rho[\text{ZnO}] - \rho[\text{mol}]$). The last column shows the difference in the densities between the first two columns (i.e., $\rho^{basis}[\text{diff}] = \rho^{TZ}[\text{diff}] - \rho^{DZ}[\text{diff}]$). We have used an isovalue of 0.001 eÅ ⁻³ for the first two columns and 0.0003 eÅ ⁻³ for the last column.	81

6.8	Both panels show how the CO molecules evolve during the 3.5 ps NVT simulation with DZVP basis sets at 300 K (in the absence of any collision). In both panels, the cyan, lime, and blue trajectories show the change in the positions of the 1st, 2nd, and 3rd CO molecules, respectively, for the entire trajectory (3.5 ps). The changes in the 4th CO position are shown with the stick model without changing its colors.	83
6.9	Panels (a) and (b) show how the CO molecules evolve during the 3 ps NVE simulation with the DZVP and TZV2P basis sets, respectively, for the 0.6 eV collision. In both panels, the cyan, lime, and blue sticks show the change in the positions of 1st, 2nd, and 3rd CO molecules, respectively, for the entire trajectory (3 ps). The changes in the 4th CO position are shown with the stick model without changing its colors, and the changes in the water molecule are shown with the glass ball-and-stick model. Larger fluctuations in the position of 4th CO molecule can be seen in both cases. This large fluctuation is due to its closeness to the impact area. However, a hopping mechanism is only observed with the TZ basis.	85
6.10	Panel (a) shows the changes in total (E), kinetic (T), and potential (V) energies of a ZnO slab covered with four CO molecules impinged by an H ₂ O molecule having 0.6 eV of kinetic energy during an NVE run (at PBE/DZVP). Panel (b) shows the changes in the kinetic energy of each of the fragments during the NVE run corresponding to the left panel. The inset in the right panel shows the changes in kinetic energies of the fragments during the first 100 fs of the simulation.	86
6.11	Both panels show how the CO molecules evolve during the 5 ps NVE simulation with the DZVP and TZV2P basis sets, respectively, for the 6 eV collision. In both panels, the cyan, lime, and blue trajectories show the change in the positions of 1st, 2nd, and 3rd CO molecules, respectively. The changes in the 4th CO position are shown without changing its color scheme, and the changes in the water molecule are shown with the glass ball-and-stick model. With the TZ basis, the 4th CO trajectory does not seem continuous because it crossed the periodic boundaries at the points of visual discontinuity.	87
6.12	Snapshots of an NVE run of H ₂ O colliding with a ZnO slab at 6 eV of incident energy. (a) [(b)] panel shows the results obtained using the DZVP (TZV2P) basis set. The 4th CO is identified with the number 4 in all the frames.	88
6.13	The left panel shows the changes in total (E), kinetic (T), and potential (V) energies of a ZnO slab covered with four CO molecules impinged by an H ₂ O molecule having 6 eV of kinetic energy, during an NVE run (at the PBE/DZVP level of theory). The right panel shows the changes in the kinetic energy of each of the fragments during the NVE run corresponding to the left panel. The inset in the right panel shows the changes in kinetic energies of the fragments during the first 100 fs of the simulation.	89

6.14	The top panel shows the electron density differences obtained for a single CO ₂ molecule adsorbed on a ZnO slab obtained using the DZVP (DZ) and TZV2P (TZ) basis sets. Each column in the top panel is obtained by subtracting the total electron density of the entire system from its individual components (i.e. $\rho[\text{diff}] = \rho[\text{ZnO}+\text{mol}] - \rho[\text{ZnO}] - \rho[\text{mol}]$). The bottom panel shows the difference in the densities between the columns of the top panel (i.e., $\rho^{\text{basis}}[\text{diff}] = \rho^{\text{TZ}}[\text{diff}] - \rho^{\text{DZ}}[\text{diff}]$) plotted at two different iso values. For the top panel, we have used an iso-value of 0.001 eÅ ⁻³ , and in the bottom panel, both 0.0003 and 0.001 eÅ ⁻³ iso-values are used. Clearly, even at a 0.0003 eÅ ⁻³ iso-value, the changes in the density with different basis sets are only observed in the CO ₂ molecule but not in the interaction region between the CO ₂ and ZnO. For the case of CO, there are huge changes in the interaction region, as shown in Figure 6.5.	90
7.1	(a-d) Schematic diagram of the exciton condensation. E_G and E_B are the band gap and excitation binding energy, respectively, as shown in (h). E_i and E_f are the energies of the incident and scattered beam, respectively. w is the momentum between the CBM and the VBM as shown in (g). (e) Atomic structure of 1T-TiSe ₂ . The blue and orange circles denote the Ti atoms and the Se atoms, respectively. The arrows denote the periodic lattice distortion (PLD) displacements $\{\mathbf{d}_i\}$. (f) The Brillouin zone (BZ) of TiSe ₂ . The solid orange and dashed blue lines denote the BZ of the 1×1 and 2×2 TiSe ₂ cell, respectively. $\{\Gamma, M, K\}$ and $\{\bar{\Gamma}, \bar{M}, \bar{K}\}$ denote the special k points in the 1×1 and 2×2 BZ, respectively. (g)(h) Band structures of 1T-TiSe ₂ at (g) its normal state and (h) charge density wave state. (i) Experimental [259] (dots) and simulated (solid lines) plasmon dispersions at different temperatures. (j) Phase diagram of the plasmon in TiSe ₂ as a function of PLD and temperature. The blue circles and orange squares denote the experimental measurements and simulation results, respectively. The experimental line describes the PLD as a function of temperature, reproduced from [267].	94
7.2	Comparison of the EELS with the random phase approximation (RPA) vs Bootstrap XC kernels for (a) the normal state (b) the CDW state.	98
7.3	(a-f) Effective band structures (EBS) and (g-l) electron energy loss spectra (EELS) at different PLDs $\{\eta\mathbf{d}_i\}$. The colorbars for (a-f) and (g-l) denote the spectral weight [Eq. 7.14] of the EBS and the momentum q of the EELS [Eq. 7.10], respectively. The dashed lines in panels (a-f) and (g-l) denote the parabolic fittings of the band structures $E(k)$ [Eq. 7.16] near the M point and the plasmon dispersions $\epsilon(q)$ [Eq. 7.17] near w , respectively.	102
7.4	(a-f) Effective band structures (EBS) and (g-l) electron energy loss spectra (EELS) at different electronic temperatures T_e . The colorbars for (a-f) and (g-l) denote the electron population of EBS [Eq. 7.14 and 7.15] and the momentum q of the EELS [Eq. 7.10], respectively. The dashed lines in (g)-(l) denote the plasmon dispersions [Eq. 7.17] near w . The dark dots denote the EELS at $q = w$	104

8.1	Optical absorption spectrum of the CH ₄ molecule. Panel (a) shows the oscillator strength distribution of the CH ₄ molecule obtained with the LDA xc potential and aug-cc-pVTZ [397] basis set. The solid green line shows the RT-TDDFT result, whereas the dotted red line shows the LR-TDDFT result. Panel (b) shows the RT-TDDFT spectra calculated using the velocity gauge (VG, solid green line) and the length gauge (LG, dotted blue line) formulations. Panel (c) compares the velocity (solid green line) and dipole (dotted black line) spectra. The inset shows the behavior of the spectra around $\omega = 0.119$	119
8.2	Dynamical conductivity $\sigma(\omega)$ of monolayer <i>h</i> -BN. The solid red line was obtained from our RT-TDDFT PySCF implementation, and the dashed gray line was obtained from Ref. [17].	121
8.3	Geometry of our simulation cell consisting of a (ZnO) ₄ molecular cluster and a 4 × 4 periodic graphene sheet. The purple, red, and dark gray spheres represent zinc, oxygen, and carbon atoms, respectively.	123
8.4	(a) Time evolution of the time-derivative of the total electronic charge Q above the $z = 1.0 \text{ \AA}$ dividing plane. The inset shows the electric field of the applied laser pulse. (b) Mulliken charges as a function of time. The blue, green, and red lines show the total summation of the Mulliken charges on the zinc, carbon, and oxygen atoms, respectively.	124
8.5	Dynamic charge transfer in a system made up of a (ZnO) ₄ molecular cluster and a periodic graphene sheet. The charge density difference was calculated between $t = 0$ and 11.22 fs, and the green and purple isosurfaces are plotted for $\Delta\rho = \pm 0.04$ a.u.	125
9.1	(a) Schematic diagram of a FE array. The bright and dark blocks denote the (b) up-polarized (c) and down-polarized structures of BaTiO ₃ , respectively. (d) Diagram of ultrafast optical polarization switching as a function of time. The red lines denote two sequential identical laser pulses. (e) Diagram of laser-induced modification of the dynamical potential energy surface (PES). The gray (black) line represents the ground- and excited-state PES, respectively.	129
9.2	(a) FE polarization P and (b) Effective temperature \tilde{T} as a function of time. FE displacements as a function of time for (c) $\alpha = z$ (d) $\alpha = x$. Solid lines denote the FE displacements calculated in the 1 × 1 × 1 cell, and dashed lines denote the average FE displacements in the 3 × 3 × 3 cell.	133
9.3	Two-dimensional contour plots of the charge densities at the plane $y = a/2$ for (a) $\rho_{\text{tot}}(x, z, t_0)$, (b) $\rho_{\text{sym}}(x, z, t_0)$, (c) $\rho_{\text{FE}}(x, z, t_0)$, (d) $\rho_{\text{tot}}(x, z, t_f)$, (e) $\rho_{\text{sym}}(x, z, t_f)$, (f) $\rho_{\text{FE}}(x, z, t_f)$, (g) $\Delta\rho_{\text{tot}}(x, z, t_f)$, (h) $\Delta\rho_{\text{sym}}(x, z, t_f)$, and (i) $\Delta\rho_{\text{FE}}(x, z, t_f)$, where t_0 and t_f are the start and end times of the laser pulse, respectively. Panels (j) and (k) depict the bonding and anti-bonding charge, respectively, as a function of time. For ease of comparison, the initial values of ρ_{tot} , ρ_{sym} , and ρ_{FE} in (j) are downshifted by 20.7×10^{-3} , 20.0×10^{-3} , and 5.70×10^{-3} a.u., respectively.	136

10.1	(a) PFOA molecule surrounded with 43 explicit water molecules (depicted as thin sticks for visual clarity). (b) Optical absorption spectra of PFOA + 43H ₂ O molecules. (c) Temporal dependence of the electric field of the applied optical pulse used in our calculations.	146
10.2	Variations in the dissociation dynamics of PFOA + 43H ₂ O molecules as a function of time. The encircled C3-F14 and C1-F10 bonds in PFOA dissociate at the end of 36 fs.	147
10.3	(a) Variation of selected C–F and C–C bond lengths in PFOA as a function of time (see Fig. 10.1(a) for atom numbering). After 12 fs, an irreversible elongation of the C3-F14 and C1-F10 bond in PFOA occurs. (b) Variation of the HOMO-LUMO gap of the system as a function of time. At 12 fs, the HOMO-LUMO gap reaches a minimum at which a photo-excited electron transfer occurs, and the C3–F14 bond starts to irreversibly break.	148
10.4	(left panel) Time-dependent Bader charges for the dissociated C3-F14 bond on PFOA; (right panel) variation of the total RT-TDDFT charge density of PFOA for the C3-F14 bond as a function of time.	149
10.5	Variation of the HOMO for PFOA + 43H ₂ O molecules as a function of time. As time progresses, charge dynamically transfers from the water to an anti-bonding orbital on PFOA. The C-F bond begins to subsequently destabilize and dissociate as it continues to vibrate dynamically. The positively charged hole remaining in the solvent arising from the afore-mentioned charge-transfer excitation becomes stabilized by the surrounding water molecules.	151
B.1	Clustering of molecules in the C3F6 dataset. PFAS molecules are represented on a 2-dimensional latent space using the PC t-SNE algorithm. Each point represents a molecule, and the colors of the points designate clusters classified using k-means clustering.	219
B.2	Clustering of molecules in the CF dataset. Each point represents a molecule that is either bioactive (red) or inactive (blue) towards (a, b) K18 and (c, d) CYP2C9. The molecules are visualized on a 2-dimensional space using (a, c) PC t-SNE (unsupervised) or (b, d) semi-supervised metric learning. . . .	223
B.3	Representative molecules found in the ChEMBL database that are bioactive on CYP2C9. The ChEMBL IDs of the molecules are (a) ChEMBL1411743, (b) ChEMBL1411201, and (c) ChEMBL1332759. The yellow circles represent the functional groups resulting in a structural alert of each molecule. . .	225
B.4	Histogram of binding energies between molecules in the CF dataset and CYP2C9. The histogram bins are organized by binding energies, where each bin represents the number of bioactive (red) and inactive (blue) molecules. . .	225
B.5	Interactions between CYP2C9 and an ester-group-containing molecule (methyl 4-[2-propyl-1-([4-trifluoromethyl]phenyl)sulfonylamino)-2-hexen-1-yl]benzoate), as predicted with Autodock. The ester group is highlighted with yellow spheres. The planar structure at the bottom represents the HEME group in CYP2C9. The C, N, O, F, and Fe atoms are colored grey, blue, red, pink, and brown, respectively.	226

B.6	Interactions between CYP and ester-containing PFAS ligands, (a) ethyl-[(nonafluorobutyl)sulfinyl]iminoacetate, (b) nonafluorohexyl2-chloroacrylate, (c) heptafluoropentyl acetate, and (d) hexafluorooxane-2,6-dione from the OECD list, as predicted with Autodock. The ester groups are highlighted with yellow spheres. The planar structure at the bottom represents the HEME group of CYP2C9. The C, N, O, F, and Fe atoms are colored grey, blue, red, pink, and brown, respectively.	227
C.1	Geometries at the 125, 1000, and 2000 time steps of the NVE simulation with 6 eV of collision energy using the TZ basis set.	236
C.2	Differences in the optimized geometries of CO/CO ₂ on a 4x4 ZnO slab with different basis sets. The differences are minor; however, it can be noticed that the optimized structures at the DZ basis have shorter adsorbate-surface distances compared to the optimized structures at the TZ basis, reflecting the stronger (weaker) adsorption energies with the DZ (TZ) basis.	242
C.3	The geometric configuration of the CO-Zn ₉ O ₉ cluster considered for the NBO analysis. Important bond lengths are shown.	245
C.4	Top- and side-view of the 6 eV collision structure along with the numbering scheme used. The atoms with a maximum change in the Mulliken charge during the NVE simulation were highlighted with green (> 0.2 e ⁻ change) and yellow (> 0.25 e ⁻ change). The top and bottom panels correspond to the DZ and TZ basis set, respectively.	247
C.5	Changes in the Mulliken charges for a few important atoms during the simulation (6 eV collision). The left and right panels show the changes in the Mulliken charges for the simulations performed with the DZ and TZ basis, respectively. The numbering scheme of the atoms is given in Figure C.4. . .	247
C.6	Panels (a) and (b) show how the CO and H ₂ O molecules evolve during the NVE simulation with the DZVP and TZV2P basis sets, respectively, for the 6 eV collision. In both panels, the cyan, lime, and blue trajectories show the change in the positions of the 1st, 2nd, and 3rd CO molecules, respectively, for the entire trajectory (the simulation was stopped after the desorption of any molecule from the ZnO surface). The changes in the 4th CO position are shown with the stick model without changing its colors, and the changes in the water molecule are shown with the glass ball-and-stick model. Once again, due to its proximity to the impact region, the 4th CO molecule exhibited significant changes with both basis sets. A difference in the behavior of the 4th CO with the DZ and TZ basis sets is apparent (results are discussed in the main text). The trajectories of the molecules appear discontinuous since they crossed the periodic boundaries at the points of visual discontinuity. In panel (a), the diffusion of the 4th CO and scattering of H ₂ O were observed (DZ basis); in panel (b), the desorption of the 4th CO and the dissociation of H ₂ O were observed (TZ basis); (c) Relative change in the positions of the 4th CO and H ₂ O molecules along the c-direction of the lattice.	250

C.7	Panels (a) and (b) show how the CO and H ₂ O molecules evolve during the NVE simulation with the DZVP and TZV2P basis sets, respectively, for the 6 eV collision using LDA. Panels (c) and (d) show the final frames of the simulations. The dissociation of the H ₂ O molecule occurs with the TZ basis while diffusion of the H ₂ O occurs when the DZ basis is used.	252
C.8	Top panels (a), (c), and (e) show the changes in the KE of the 4th CO using both the DZ and TZ basis during the entire simulation time. Bottom panels (b), (d), and (f) show the relative change in the positions of the 4th CO and H ₂ O molecules along the c-direction of the lattice during the entire simulation time. From left to right , both the top and bottom panels correspond to simulations with the PBE-D3 (0.6 eV collision), LDA (6 eV collision), and PBE (6 eV collision) functionals, respectively.	254
C.9	Relative change in the positions of the 2nd CO, 4th CO, and H ₂ O molecules along the c-direction of the lattice. The results are presented for NVE simulations performed with the PBE (no dispersion correction) and PBE+D3 (dispersion-corrected) functionals and with the DZ and TZ basis sets. Here, the impinging H ₂ O molecule collides with the ZnO substrate in the vicinity of the 2nd CO molecule.	257
C.10	Relative change in the positions of the 2nd CO, 4th CO, and H ₂ O molecules along the c-direction of the lattice. The results are presented for NVE simulations performed with the PBE (no dispersion correction) and PBE+D3 (dispersion-corrected) functionals and with the DZ and TZ basis sets. Here, the impinging H ₂ O molecule collides with the ZnO substrate between the 2nd and 4th CO molecules.	258
D.1	The peak position $\epsilon(q = w)$ [Eq. 7.17] (a) and the peak height at $\epsilon(q = w)$ (b) of the soft mode as a combined effect of the PLD and T_e . (c)(d) The lattice energy E_{lat} , exciton energy E_{ex} , electron-phonon energy E_{ph} , and total energy E_{tot} as a function of the PLD, indicated by (c) the analytical model and (d) our DFT and experimental framework.	260

List of Tables

2.1	Key geometric parameters (bond lengths and improper dihedral angles) of the transition states optimized with DFT. Atom labels are depicted in Figure 2.4. Bond lengths are in Å and improper dihedral angles are in degrees.	13
2.2	Activation free energies and its enthalpy and entropy components obtained by DFT for the Diels-Alder reaction with and without cyclodextrin catalysts. Units are in kcal/mol.	15
3.1	Relative energies of the DMP-L ⁺ and DPM-D ⁺ states obtained from various computational methods.	24
3.2	Comparison of various geometric parameters for the diamine cation transition state obtained with the MP2, CCSD, and PZ-SIC computational methods. The atom labels and numbers refer to the labeling scheme used in Figure 3.2. The shaded entries denote geometric parameters with the largest variation between the MP2/CCSD and PZ-SIC approaches.	29
5.1	Initial lattice parameters of SiC (2H) and ZnO (B4).	53
5.2	Comparison of optimized lattice parameters of the minima structure calculated with DFT and DFTB for SiC.	60
6.1	Adsorption Energies of the Molecules Calculated Using CP2K with Different Basis Sets and Differences in the Adsorption Energies between Basis Sets.	79
6.2	Adsorption Energies of the Molecules Calculated Using FHI-aims with Two Different Basis Sets and Differences in the Adsorption Energies between Basis Sets.	82
9.1	Symmetric positions α_I^{sym} ($\alpha = x, y, z$), FE displacements Δz_I and Born effective charge of the I -th atom. We categorize these three O atoms into two types: one O atom that is parallel to the polarization and two O _⊥ atoms that are perpendicular to the polarization.	131
A.1	Reference Cartesian coordinates (in Å) for the MP2-optimized DMP-L ⁺ molecular cation.	206

A.2	Reference Cartesian coordinates (in Å) for the MP2-optimized DMP-D ⁺ molecular cation.	207
A.3	Reference Cartesian coordinates (in Å) for the MP2-optimized transition state of the diamine molecular cation.	208
A.4	Reference Cartesian coordinates (in Å) for the CCSD-optimized DMP-L ⁺ molecular cation.	209
A.5	Reference Cartesian coordinates (in Å) for the CCSD-optimized DMP-D ⁺ molecular cation.	210
A.6	Reference Cartesian coordinates (in Å) for the CCSD-optimized transition state of the diamine molecular cation.	211
B.1	Selected targets (ion channels, enzymes, and receptors).	213
B.2	Optimized hyperparameters and Silhouette scores for various unsupervised machine learning models.	217
B.3	Cluster number, accuracy, and maximum common structure most likely to be found in bioactive molecules toward each target. Most molecules with 1,3-bis(trifluoromethyl) benzene are bioactive toward the CYP enzymes, while molecules with 1-ethyl-3,5-bis(trifluoromethyl) benzene are likely to be bioactive toward ATXN.	220
B.4	Predicted substructures that induce bioactivity toward each target, based on semi-supervised metric learning.	224
C.1	Optimized cell parameters of bulk ZnO obtained with each software package.	229
C.2	Layer convergence test for a 2x1 slab with a 6x6x1 k-mesh (single-point calculations).	230
C.3	K-point convergence test for a 2x1 slab with 4 layers (single-point calculations).	231
C.4	Vacuum convergence tests for a 2x1 slab with a 6x6x1 k-mesh and with 4 layers (single-point calculations).	232
C.5	Single-point adsorption energy of CO on the ZnO surface calculated with three different basis set tiers. In all of these adsorption energy calculations, the geometries of ZnO, CO, and ZnO+CO were fixed at the optimized geometry of ZnO+CO obtained using the Tier-2 basis.	233
C.6	Script for generating a 4-layer-thick slab and all the different adsorbate configurations for a CO molecule on a 2x1 slab.	234
C.7	Difference in the adsorption energy (eV) of a CO molecule on a 4x4 ZnO slab with a change in the k-mesh size (single-point calculations).	235
C.8	Difference in the interaction energy (eV) of the 4th CO molecule with the rest of the system (three CO molecules, one H ₂ O molecule, and the 6x3 ZnO slab) with a change in the k-mesh size (single-point energy calculations) at three different time steps of the NVE simulation (with 6 eV of collision energy).	237
C.9	Adsorption energies of molecules calculated using CP2K with 3 different basis sets and the differences in the adsorption energies between basis sets. Here, DZVP, DZVP-SR, and TZV2P correspond to the basis set used for the elements O, H, C, and N. For Zn, only short-ranged basis sets are used in all of the cases.	240

C.10 Adsorption energies of molecules calculated using CP2K with 2 different basis sets and the differences in the adsorption energies between basis sets. The effects of dispersion corrections are also presented.	240
C.11 Adsorption energies of CO and CO ₂ molecules on a ZnO slab with various plane-wave cutoff energies. The plane-wave cutoff is converged to 1 meV at the 1000 Ry cutoff.	241
C.12 Differences in the adsorption energies for CO and CO ₂ molecules with different basis sets for a fixed nuclear geometry. Here, we used the geometries of a single CO/CO ₂ on a 4x4 ZnO slab optimized at the PBE/TZV2P level of theory. Note that the adsorption energy differences in the last column are very close to the adsorption energy differences reported for the completely optimized structures, as shown in Table C.9.	244
C.13 Calculated Wiberg bond indices for the Zn-C bond and natural charges for the Zn and C atoms involved in the bonding with the DZ and TZ basis sets.	245
C.14 Adsorption energies of different molecules on a ZnO slab as predicted by the three different software packages. In the last 3 columns, the difference in the adsorption energies predicted by any two software packages is given for comparison.	248
C.15 Adsorption energy of the CO molecule calculated using CP2K with different basis sets and exchange-correlation functionals.	250
C.16 Adsorption energy of the CO molecule calculated using FHI-aims with different basis sets and exchange-correlation functionals.	255
C.17 Adsorption energy of the CO molecule on a Zn ₉ O ₉ cluster calculated using the Gaussian16 software package with different basis sets and exchange-correlation functionals. All results were corrected for the basis set superposition error (BSSE) using the counterpoise method as implemented in Gaussian16.	255

Chapter 1

Introduction

A deep understanding of the electron dynamics in atoms, molecules, and periodic systems is necessary to model their responses to an applied external field. The broad range of field-induced real-time dynamics, via an external field, in complex systems has many potential applications, including field-induced transitions [1, 2], surface reactions [3], photocatalysis [4, 5], and strong-field dynamics [6–8]. These applications can help pave the way for real-time dynamics to quantitatively investigate charge carrier dynamics in both the ground and excited states. Theoretical reproduction and understanding of field-induced dynamics are also critical for the design and analysis of nanoscale devices [9]. Thanks to the recent development of laser technologies, experimental and theoretical researchers can study laser-induced electron dynamics in materials [10–12].

To fully understand laser-matter interactions such as laser-induced electron dynamics, attosecond transient absorption spectroscopy (ATAS) [13, 14] is one experimental technique that can resolve the dynamics with natural time resolution. It can naturally cap-

ture electron dynamics in matter due to its attosecond time resolution being shorter than a typical time-scale of electron dynamics, which is on the order of femtosecond. While ATAS for solid-state materials provides a wealth of electron dynamics information, the resulting experimental data are often complicated due to dynamically complex electronic structure and highly-nonlinear effects. Thus, accurate and affordable computational simulation of electron dynamics is crucial for progress in this area.

Dynamical process analyses in solids require a size-scalable excited-state calculation tool. The real-time time-dependent density functional theory (rt-TDDFT) [15] is one of the most promising numerical approaches for real-time simulations, receiving intensive research interest in laser-matter interactions [15–17]. The method propagates the electron density in real-time through numerical integration of the time-dependent Kohn–Sham (TDKS) equations. Furthermore, the rt-TDDFT simulations have been applied to ATAS experiments for solid-state systems, playing a significant role in constructing the data interpretation [18–21], which exemplify the rt-tddft’s size-scalable excited-state calculation possibilities.

Since the rt-TDDFT is an extension of the Density Functional Theory (DFT), which calculates ground-state properties, the first part of this dissertation is related to DFT and its application in ground-state systems. These chapters consist of DFT research during the initial stages of my graduate program, which was necessary to understand DFT before proceeding to the real-time case. Finally, the latter half introduces the rt-TDDFT formalism, implementation for the electron dynamics and transient absorption spectra simulation, and concludes with an application of rt-TDDFT. Below is each chapter’s general

outline.

1.1 Outline

The first part of this dissertation, Chapters 2 to 5, describes work on analyzing a system's ground state via DFT [22]. The goal of these chapters is to convey to the reader DFT's methodologies and demonstrate its utility for studying ground state properties. Specific topics in these chapters are as follows:

- Chapter 2 begins with an introduction to the Diels-Alder reaction and Cyclodextrins. It then transitions to the results and discussion of the DFT calculations, describing the effect of Cyclodextrins being used as a catalyst and why methyl- β -cyclodextrin outperforms β -cyclodextrin to promote the Diels-Alder reaction. Finally, the chapter finishes with the computational details and conclusions of the results.
- Chapter 3 explores the charge-localization effect of N/N' -dimethylpiperazine (DMP^+) diamine cation using DFT methods, where a previous group claimed that **all** DFT methods fail on this specific cation. We show the inconsistency in their work, leading to refuting their claim.
- Chapter 4 takes a step away from DFT to explore how machine learning could help with bioactivity analysis of per- and polyfluoroalkyl substances (PFASs). This chapter aims to familiarize the reader with PFAS molecules and their environmental effects, which are further explored in Chapter 10 via rt-TDDFT.
- Chapter 5 returns to DFT for convex hull studies. It introduces the Density Functional

Tight Binding (DFTB) theory and goes over DFT. This chapter shows a limitation of DFT for specific phase space calculations.

For the rest of this dissertation, Chapters 6 to 10 explores DFT's time-dependent (TD) case. More specifically, these chapters start exploring excited states of systems by studying system responses to light-matter interactions. Specific topics in these chapters are as follows:

- Chapter 6 explores the quality of the employed basis set when using **ab initio molecular dynamics** to study collisions between the reactants of a water-gas shift reaction with a ZnO surface. The calculations in this work use the Born-Oppenheimer Molecular Dynamics (BOMD) theory. The chapter follows with the computational details, results, discussion, and concluding remarks. The goal here is to transition into real-time simulations by first performing DFT with molecular dynamics calculations.
- Chapter 7 provides *ab initio* atomic-level framework for rationalizing recent experiments and further manipulating exciton condensates in TiSe₂. The work was only possible with Dr. Chao Lian's extensive help. In this chapter, we accurately reproduce and characterize the experimental spectra using linear-response time-dependent density functional theory (lr-TDDFT). The goal is to display the time-dependent density functional theory (TDDFT) in the linear regime before going into the real-time regime.
- Chapter 8 introduces the reader to the rt-TDDFT's formalism, implementation in the velocity gauge formalism for periodic systems, and provides benchmark results for comparison. The main focus is to display the successful development and performance of the theory described in the chapter.

- Chapter 9 describes a new mechanism that enables the reversible switching of the ferroelectric polarization with optical pulses. With Dr. Chao Lian's implementation of rt-TDDFT, we introduce a new tool that enables the reversible switching of ferroelectric polarization with optical pulses.
- Chapter 10 takes a deeper dive into PFAS molecules via velocity-gauge real-time time-dependent density functional theory with molecular dynamics (VG-RT-TDDFT-MD). This chapter explores photo-induced degradation mechanisms of perfluorooctanoic acid (PFOA) contaminants by applying rt-TDDFT with molecular dynamics within the velocity gauge. This work is an extension of Dr. Sharma S.R.K.C. Yamijala's previous work. The results in this chapter are primarily due to his collaborative work.

Finally, Chapter 11 concludes this dissertation by summarizing the findings. The appendix at the end of the dissertation consists of the computer codes used throughout the dissertation for computing and plotting the various properties, the theory and mathematical formalisms used for specific systems, and supplementary information from my respective collaborative publications.

Chapter 2

A QM Study of Biomimetic Catalysis of Diels-Alder Reactions Using Cyclodextrins

This chapter describes the effect of Cyclodextrins as a catalyst and why methyl- β -cyclodextrin outperforms β -cyclodextrin to promote the Diels-Alder reaction. The work in this chapter resulted from a collaboration with the computational group of Dr. Chia-en A. Chang. It was published as an article in *Journal of Catalysts* [23], where the complete research is provided. Only the density functional theory (DFT) results are presented and analyzed below.

2.1 Introduction

Supra-molecular catalysts offer potential advantages over enzymes, including more significant physical and chemical stability, lower molecular weight, and a more varied selection of chemical versatility for creating structure and functionality. Due to their unique chemical and biological properties, cyclodextrins are useful in pharmaceutical, food, and agricultural industries [24–28]. Their hydrophobic pocket also allows them to catalyze chemical reactions similar to an enzyme [29–35]. In addition, β -cyclodextrins have been used as a catalyst to promote the Diels-Alder reaction and synthesize various compounds [36–39].

The Diels-Alder reaction is an important carbon–carbon bond formation reaction in organic synthesis. It forms two carbon–carbon bonds and up to four new stereo centers in one step. For typical Diels-Alder reactions between a diene and dienophile, frontier molecular orbital theory states that the interaction of the highest occupied molecular orbital (HOMO) of the diene with the lowest unoccupied molecular orbital (LUMO) of the dienophile is the dominant interaction in the transition state [40]. The rate of the Diels-Alder reaction could be accelerated by narrowing the energy gap between the HOMO and LUMO [41, 42].

Recently, cyclodextrins were reported to promote Diels-Alder reactions of 9-anthracene-methanol with a variety of N-substituted maleimides under mild reaction conditions [38]. In this paper, Chaudhuri et al. proposed a mechanism whereby the cyclodextrins bind the hydrophobic substituents on the maleimides (**2a–2d**) and activate the dienophile by electronic modulation of the maleimide double bond via hydrogen bonds (Hbonds) with cyclodextrins (Figure 2.1). Methyl- β -cyclodextrin was found to bind maleimides with large

N-substituents, such as **2a**, more strongly than β -cyclodextrin, as indicated by the greater changes in the ^1H NMR chemical shifts. Methyl- β -cyclodextrin is also significantly more efficient than β -cyclodextrin at promoting the reaction of N-cyclohexylmaleimide. The authors further suggested that this situation is likely because methyl- β -cyclodextrin is both more flexible and has a more non-polar cavity than β -cyclodextrin [38]. However, if the reaction is promoted by the Hbonds between the maleimides and cyclodextrins, then β -cyclodextrin should perform at least as well as methyl- β -cyclodextrin since it has more hydroxyl groups to participate in the hydrogen bonding. The mechanism of how the flexibility and hydrophobicity of methyl- β -cyclodextrin aids the catalytic process is also unknown.

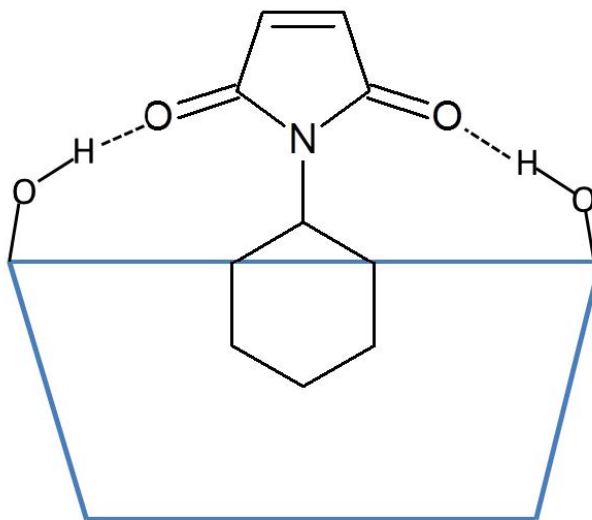


Figure 2.1: Illustration of the hydrogen bonds between the carbonyl groups on maleimides and the secondary hydroxyl groups on cyclodextrins. N-cyclohexylmaleimide is used as an example. The blue trapezoid represents the cyclodextrin, and the dotted lines represent the hydrogen bonds.

A number of studies utilized computational methods to look at the catalytic mechanism of β -cyclodextrin as a catalyst in organic reactions. Luzhov and Venanzi used the Austin Model 1 (AM1) method to study the reaction of phenyl acetate with

β -cyclodextrin [43] and found large energy differences between different reaction sites. Eto et al. used the Parameterized Model number 3 (PM3) method to study the pericyclic reaction of cinnamyl xanthates in β -cyclodextrin cavities [44] and found that the transition state structure was stabilized by a hydrogen bond with one of the hydroxyl groups of β -cyclodextrin, which explained the rate acceleration. Castro et al. studied the effect of β -cyclodextrin on the hydrolysis of N-phenylphthalamide (Ph) and N-adamantylphthalamide (Ad) by the PM3 method [45]. Their result showed that all the β -cyclodextrin/Ad complexes are more stable than the β -cyclodextrin/Ph complexes. Recently Giuseppe Floresta et al. studied β -cyclodextrin as a catalyst for the synthesis of 2-methyl-3, 5diarylisoxazolidines with the PM3 method [46] and found that the transition states (TSS) coming from the nitron in an E-configuration are about 3 kcal/mol more stable than that derived from the Z-one. In this paper, we performed computational work using quantum mechanical (QM) calculations to investigate Diels-Alder reactions with cyclodextrins as catalysts. We studied the reaction of 9-anthracene-methanol (compound **1**) with N-cyclohexylmaleimide (compound **2a**) (Figure 2.2) catalyzed by β -cyclodextrins. The study describes a new mechanism and explains why methyl- β -cyclodextrin outperforms β -cyclodextrin to promote the reaction.

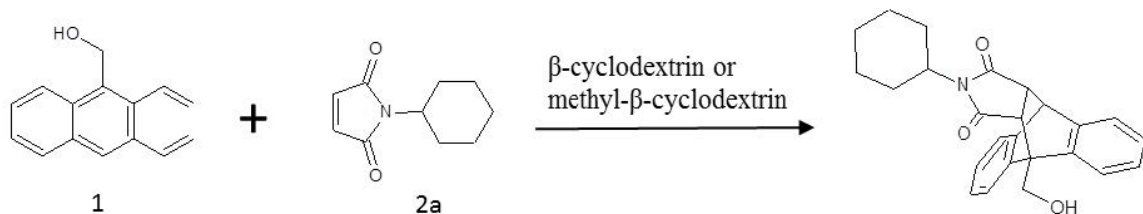


Figure 2.2: The Diels-Alder reaction of 9-anthracene-methanol (**1**) with N-cyclohexyl maleimide (**2a**), with β -cyclodextrin or methyl- β -cyclodextrin as a catalyst.

2.2 Results and Discussion

2.2.1 Explanation of Enhanced Catalysis by QM Calculations

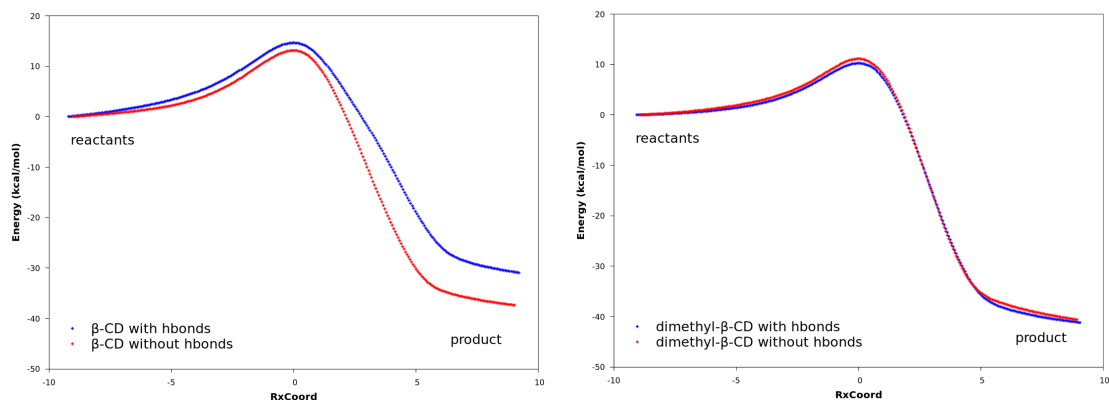


Figure 2.3(a)

Figure 2.3(b)

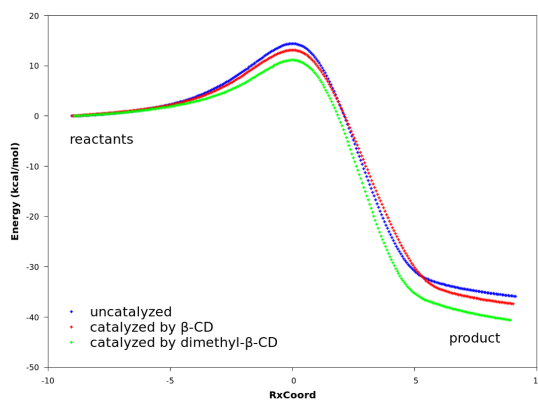


Figure 2.3(c)

Figure 2.3: Reaction paths obtained with DFT. In all plots, the curves are normalized to set the reactant energies at zero. **(a)**: reaction paths for the reaction catalyzed by β -CD, with and without hydrogen bonds. **(b)**: reaction paths for the reaction catalyzed by dimethyl- β -CD, with and without hydrogen bonds. **(c)**: reaction paths for uncatalyzed reaction between compounds **1** and **2a**, reaction catalyzed by β -CD, and reaction catalyzed by dimethyl- β -CD. No hydrogen bonds exist between compound **2a** and the cyclodextrin.

Molecular mechanics calculations suggest that cyclodextrins help pre-organize the two reactants in the correct geometry arrangement [23]. To gain a better understanding of whether such a pre-organization can actually lower the activation barrier and promote

a Diels-Alder reaction, we used large-scale QM calculations for five reactions: the Diels-Alder reaction of **1** with **2a** without cyclodextrins, and **1** and **2** plus dimethyl- β -CD or β -CD with or without the Hbond formation. Our DFT method was used to locate the TSs and the corresponding Intrinsic Reaction Coordinates (IRC) with and without β -CD or dimethyl- β -CD. In addition, our QM calculations were also carried out at the ω B97XD/6-31G(d) level of theory to compute the optimized TS, reactant, and product geometries for all five reactions. We have specifically chosen to use the ω B97XD functional for this work since this QM method incorporates dispersion effects in conjunction with a large amount of exact exchange—both effects (particularly the latter) are required for accurately calculating reaction barriers in Diels-Alder reactions [47].

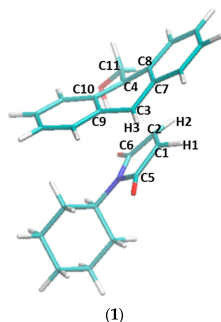


Figure 2.4(a) Non-catalyzed Transition State.

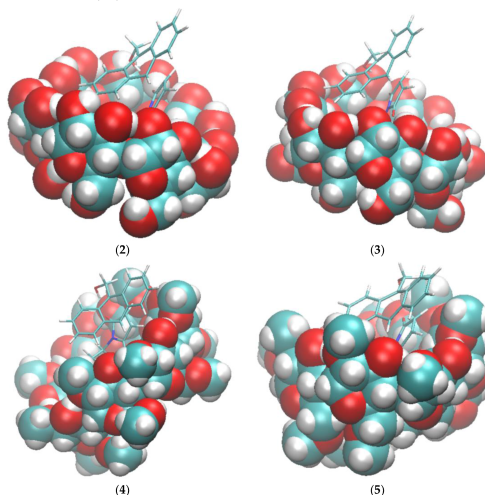


Figure 2.4(b) Catalyzed Transition States.

Figure 2.4: The transition states of five reactions optimized at the ω B97XD/6-31G(d) level of theory. (1) Transition state structure in the non-catalyzed reaction between compounds **1** and **2a**. Atom names used in Table 2.1 are labeled in this structure. (2) Transition state structure in the reaction catalyzed by β -CD, with hydrogen bonds between the two carbonyl groups of **2a** and β -CD. (3) Transition state structure in the reaction catalyzed by β -CD, without the hydrogen bonds. (4) Transition state structure in the reaction catalyzed by dimethyl- β -CD, with hydrogen bonds between the two carbonyl groups of **2a** and dimethyl- β -CD. (5) Transition state structure in the reaction catalyzed by dimethyl- β -CD, without the hydrogen bonds.

The QM calculations from the DFT method suggest that Hbonds play a nearly negligible role in lowering the activation energy (Figure 2.3). Structures with the lowest energies were optimized with the geometry optimization method using an energy-represented direct inversion in the iterative subspace (GEDIIS) algorithm which was subsequently used

for searching the TS geometries of the cyclodextrin-catalyzed reactions with the density functional theory (DFT) at the ω B97XD/6-31G(d) level of theory. The optimized structures of the TSs and their key geometric parameters are depicted in Figure 2.4 and Table 2.1 (DFT), which show that the non-catalyzed TS configuration is largely maintained in the catalyzed reactions. In all of the TSs, the lengths of the two new “covalent bonds” between **1** and **2a** are similar, as are the improper dihedral angles at the four carbon atoms where these two new bonds are to be formed. Therefore, the catalytic effects of cyclodextrins are not from the direct participation of bond formation at TS. In addition, the IRCs with and without Hbonds with both catalysts (Figure 2.3, (a) and (b)) are similar. The reactions have slightly lower activation energies by no more than 2.0 kcal/mol when Hbonds exist between catalysts and **2a**. This finding suggests that hydrogen bonding does not play a significant role in catalyzing the reaction. Interestingly, when the IRC of the uncatalyzed reaction is plotted against that of the catalyzed reactions without Hbonds (Figure 2.3(c)), they look quite similar. Because IRC calculations report only the potential energies for the reaction systems, the high similarity in potential energies suggests that the catalytic power of cyclodextrins may be related to the entropy.

Table 2.1: Key geometric parameters (bond lengths and improper dihedral angles) of the transition states optimized with DFT. Atom labels are depicted in Figure 2.4. Bond lengths are in Å and improper dihedral angles are in degrees.

No.	Bond C1-C3	Bond C2-C4	Improper C1-C2-H1-C5	Improper C2-H2-C1-C6	Improper C3-H3-C7-C9	Improper C4-C8-C11-C10
1	2.203	2.391	21.498	18.172	16.042	11.803
2	2.112	2.603	25.175	13.678	18.587	9.557
3	2.171	2.473	22.346	17.424	16.720	11.938
4	2.230	2.398	20.148	18.362	14.907	11.434
5	2.264	2.342	19.919	21.709	13.786	11.243

As such, we calculated the activation free energy and its enthalpy and entropy

components for all five reactions (Table 2.2 for DFT). Surprisingly, the activation free energy for the non-catalytic reaction was 16.89 kcal/mol from our DFT calculations, which is lower than those of the catalyzed reactions. This result is not consistent with the experimental observation [38]. The calculated activation free energies can be decomposed into enthalpy and entropy terms. Most importantly, the QM method show that the entropy term $-T\Delta S$ is the dominant factor in differentiating the activation energies. For the second sets, entropy drops from 16.03 kcal/mol in the non-catalyzed reaction to 2.98 kcal/mol in the β -CD catalyzed reaction, and to 0.5 kcal/mol in the dimethyl- β -CD catalyzed reaction according to the DFT results. The energies from the DFT calculations (Table 2.2) show no difference with or without Hbonds in the β -CD catalyzed reaction. In the dimethyl- β -CD catalyzed reaction, the system with Hbonds has an enthalpy that is 10 kcal/mol lower than the one without Hbonds. However, the enthalpy computed from DFT for the non-catalyzed forward reaction may not be highly accurate because the non-catalytic reaction has an unreasonably low enthalpy of 0.86 kcal/mol, compared with all other sets of molecules. This low enthalpy is not consistent with the energies on the IRC curve either (Figure 2.3). Regardless of the presence of Hbonds, the TSs in the dimethyl- β -CD catalyzed reaction have a lower enthalpy than those in the β -CD catalyzed reaction as obtained from the QM method. Therefore, the binding of compounds **1** and **2a** to cyclodextrins acts as a pre-organization and remarkably reduces the entropy penalty to the activation free energy for the transition state of **1** and **2a**.

The above results show that instead of directly participating in chemical bonding at the transition state, the pre-organization provided by complexation between compounds

Table 2.2: Activation free energies and its enthalpy and entropy components obtained by DFT for the Diels-Alder reaction with and without cyclodextrin catalysts. Units are in kcal/mol.

Reaction	ΔG	ΔH	$-T\Delta S$
non-catalyzed forward reaction	16.89	0.86	16.03
non-catalyzed reverse reaction	48.06	48.66	-0.60
forward reaction catalyzed by β -CD, set 1 ^a	22.75	21.29	1.47
reverse reaction catalyzed by β -CD, set 1 ^a	47.83	49.44	-1.61
forward reaction catalyzed by β -CD, set 2 ^b	24.12	21.14	2.98
reverse reaction catalyzed by β -CD, set 2 ^b	48.67	49.92	-1.25
forward reaction catalyzed by dimethyl- β -CD, set 1 ^a	13.73	10.05	3.68
reverse reaction catalyzed by dimethyl- β -CD, set 1 ^a	48.51	51.50	-2.99
forward reaction catalyzed by dimethyl- β -CD, set 2 ^b	19.68	20.18	-0.50
reverse reaction catalyzed by dimethyl- β -CD, set 2 ^b	49.48	51.25	-1.77

^a Transition state structure has hydrogen bonds between the two carbonyl groups of **2a** and the cyclodextrin

^b Transition state structure does not have hydrogen bonds.

1 and **2a** and CDs lowers the activation energy of the Diels-Alder reaction. The catalytic effect is achieved by the reduction of ΔS when going from the reactant state to the reaction barrier. The limitations of this QM study also deserve a few comments. Because of the existence of multiple conformers of cyclodextrins, there may be multiple reaction pathways and TS structures. Because cyclodextrins do not participate in the chemical bonding at the barrier, its configurations are not expected to change significantly during the reaction. Therefore, the configuration effect is largely canceled in the activation energy calculations. Furthermore, the initial configurations used in the QM calculations are those with the lowest energies, so they may represent the configuration with the largest statistical weight. Nevertheless, we could further investigate the configurational contributions of cyclodextrin to the activation barrier. Since the thermodynamic functions by the QM method do not take

into account configurational effects over multiple conformations, the calculated entropies are also underestimated as compared to reality. Solvent effects were not considered in this QM study, and they are expected to be similar for both catalyzed and un-catalyzed reactions because the reaction coordinates of the TSs (i.e., eigenvectors) are similar for both catalyzed and un-catalyzed reactions. Omitting solvent effects should not affect the conclusions of this paper; however, our QM activation energies for the reaction barrier were in the range of the experiment and correctly ranked the CD catalysts. These observations suggest that the QM results are reasonable.

In summary, the major findings from this QM study are as follows: **(1)** entropy plays a significant role in lowering the activation free energy barrier in the Diels-Alder reactions catalyzed by cyclodextrins; **(2)** hydrogen bonding, although lowering the barrier to some extent, is less important than the entropy contribution; and **(3)** dimethyl- β -CD is more efficient than β -CD in catalyzing the reaction.

Interestingly, our findings are echoed by a very recent paper [48]. Henrik Daver et al. studied the cycloaddition reaction between phenyl acetylene and phenyl azide inside a synthetic, self-assembled capsule with DFT, and concluded that the reduction of the entropic cost of bringing together the reactants are the major contributors to the rate acceleration compared to the background reaction.

2.3 Computational Methods

Each α -D-glucopyranoside on the cyclodextrin has one primary hydroxyl group and 2 secondary hydroxyl groups (Figure 2.5). All the commercially available methy-

lated β -cyclodextrins are mixtures of various isomers and homologues except trimethyl β -cyclodextrin. Since the study by Chaudhuri et al., showed an average of 1.8 methyl groups in methyl- β -cyclodextrin, and the primary hydroxyl group was more active than secondary hydroxyl groups in methylation, we assume that the major component of methyl- β -cyclodextrin is 2,6-O-dimethyl- β -cyclodextrin. For these reasons, we investigated the interactions of the two reactants (**1** and **2a**) in the Diels-Alder reaction with the catalysts β -cyclodextrin (β -CD) and 2,6-O-dimethyl- β -cyclodextrin (dimethyl- β -CD) by the computational methods briefly described below. The figures of the molecule conformations in this paper were generated using Visual Molecular Dynamics (VMD) [49].

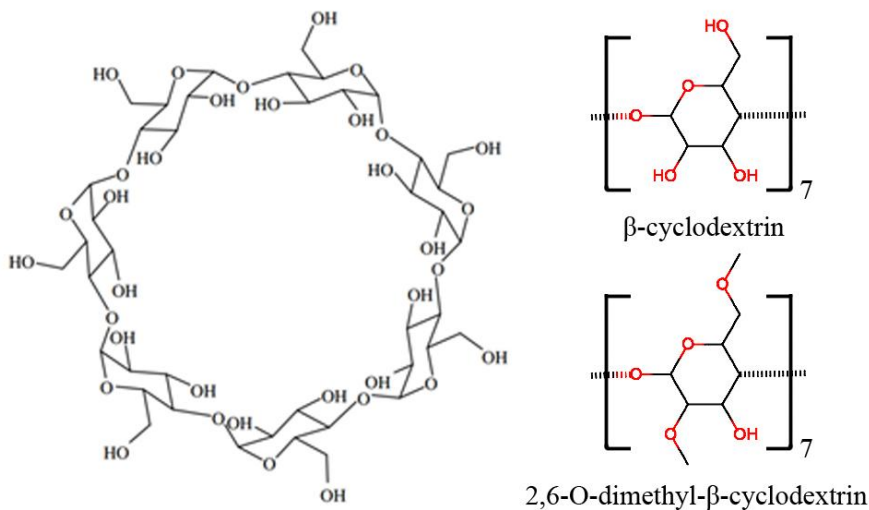


Figure 2.5: Molecular structure of β -cyclodextrin. The shape of the cyclodextrin macrocycle can be described as a truncated cone, with a narrow rim presenting primary hydroxyl groups and a wide rim presenting secondary hydroxyl groups on the glucose residues.

2.3.1 Quantum Mechanics Calculations with the DFT Method

Reaction paths and transition states (TSS) with and without β -CD and dimethyl- β -CD catalysts were calculated with the DFT method at the ω B97XD/6-31G(d) level of

theory in the G09 package [50, 51]. Aqueous solvent was not included in the present calculations. The Gibbs free energies as well as the enthalpy and entropy components of reactants and TS were calculated with standard statistical thermodynamics equations [52] at 298K using the optimized molecular structures. For example, entropy was calculated with Equation 2.3.1, where the partition function, q , is obtained by including contributions from translational, rotational, vibrational, and electronic degrees of freedom.

$$S = Nk_B + Nk_B \ln \left(\frac{q(V, T)}{N} \right) + Nk_B T \left(\frac{\partial \ln q}{\partial T} \right)_V \quad (2.1)$$

The activation free energies were obtained by computing the difference between the free energies of the reactants and the TS. Molecular structures were optimized by GEDIIS method [53] in G09.

We performed two sets of QM calculations for compounds **1** and **2a** with β -CD or dimethyl- β -CD complexes with two β -CD conformations. The first set involved a conformation with Hbonds between the two carbonyl groups of **2a** and the catalyst, and the other set did not have Hbonds between **2a** and β -CD or dimethyl- β -CD. We also performed calculations for compounds **1** and **2a** without β -CDs. Complex structures with the lowest energies were selected as initial structures for searching the TS of a cyclodextrin-catalyzed reaction and geometry optimization. The Berny algorithm [54, 55] was used for locating the TSs. After the TSs were identified, the energy profile along the reaction coordinate connecting reactants, TS, and products was calculated using the IRC. We used a step size of 0.02 amu^{1/2}-Bohr and 150 steps were run in both forward and reverse directions along the reaction path.

2.4 Conclusion

Dimethyl- β -CD has the potential to catalyze Diels-Alder reactions of 9-anthracene-methanol with a variety of N-substituted maleimides under mild reaction conditions. To understand the mechanism, we performed QM calculations. The driving force of the catalysis is not the Hbonds between cyclodextrins and **2a** because of the low occurrence of the Hbonds suggested by DFT results; furthermore, IRC calculations yielded small differences with or without the Hbonds. However, cyclodextrins, especially dimethyl- β -CD, do show a strong binding affinity with a pre-organized **1-2a** reactant complex and increase the probability of the Diels-Alder reaction. Moreover, QM calculations showed that cyclodextrins remarkably reduced the activation entropy. These findings depict a possible two-step catalysis mechanism by cyclodextrins: the first step is the complex formation between cyclodextrins and reactants, and the second step is the reaction of reactants inside the cavity of cyclodextrins to form the product. The formation of the complex is the key to the catalysis because it not only increases the rate that reactant molecules would meet each other but also lowers the entropy and thus the activation free energy of this reaction. Dimethyl- β -CD outperforms β -CD because of stronger van der Waals interactions with the pre-organized reactants and better performance in reducing the activation energy. Our QM results are in qualitative agreement with experiments on activation free energies of the Diels-Alder reactions and the relative catalytic activities for the systems under study.

Chapter 3

The Diamine Cation is not a Chemical Example where DFT fails

This chapter explores the charge-localization effect of N/N' -dimethylpiperazine (DMP⁺) diamine cation using DFT methods. The majority of the work in this chapter resulted from a collaboration with Dr. Fredy W. Aquino and was published as an article in *Nature Communications* [56]. The full paper is presented below, with the supplementary information in Appendix [A](#).

3.1 Introduction

In a recent communication, Weber et al. [57] presented a surprising study on charge-localization effects in the N/N' -dimethylpiperazine (DMP^+) diamine cation to provide a stringent test of density functional theory (DFT) methods. Within their study, the authors examined various DFT methods and concluded that “all DFT functionals commonly used today, including hybrid functionals with exact exchange, fail to predict a stable charge-localized state.” [57] This surprising conclusion is based on the authors’ use of a self-interaction correction (namely, complex-valued Perdew–Zunger Self-Interaction Correction (PZ-SIC)) [58, 59] to DFT, which appears to give excellent agreement with experiment and other wavefunction-based benchmarks. Since the publication of this recent communication, the same DMP^+ molecule has been cited in numerous subsequent studies [60–69] as a prototypical example of the importance of self-interaction corrections for accurately calculating other chemical systems. In this correspondence, we have carried out new high-level CCSD(T) analyses on the DMP^+ cation to show that DFT actually performs quite well for this system (in contrast to their conclusion that all DFT functionals fail), whereas the PZ-SIC approach used by Weber et al. is the outlier that is inconsistent with the high-level CCSD(T) (coupled-cluster with single and double excitations and perturbative triples) calculations. Our new findings and analysis for this system are briefly discussed in this correspondence.

3.2 Results and Discussion

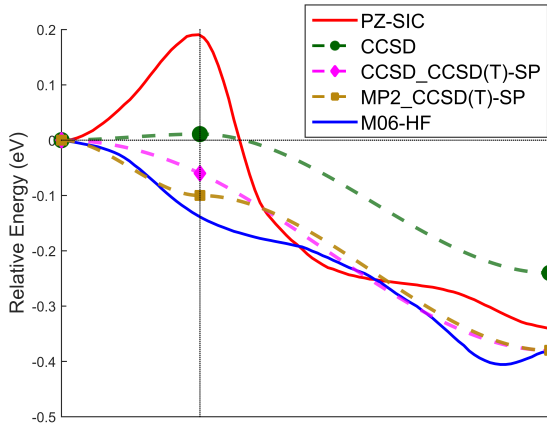


Figure 3.1: Calculated potential energy curve between the localized and delocalized state of the dimethylpiperazine cation. The PZ-SIC and M06-HF potential energy curves were obtained from Ref. [57], and the CCSD_CCSD(T)-SP and MP2_CCSD(T)-SP legend labels denote single-point energy calculations that were carried out with the CCSD(T) method using geometry-optimized structures obtained with CCSD and MP2, respectively. All three wavefunction-based approaches (CCSD, CCSD_CCSD(T)-SP, and MP2_CCSD(T)-SP) are in agreement by producing an extremely small energy barrier (< 0.01 eV), with the CCSD_CCSD(T)-SP and MP2_CCSD(T)-SP curves in close agreement with the M06-HF DFT calculations.

In Fig. 3.1, we re-plot the PZ-SIC and M06-HF (Minnesota 06 with Hartree Fock exchange) potential energy curves (reproduced from Ref. [57] by Weber et al.) overlaid on top of our new MP2 (Møller-Plesset 2nd order perturbation theory), CCSD (coupled-cluster with single and double excitations), and CCSD(T) calculations. Using the same nomenclature as Ref. [57], the charge-delocalized dimethylpiperazine (DMP-D⁺) structure occupies the global minimum on the potential energy curve and is characterized by a positive charge that is delocalized over the two equivalent nitrogen atoms. In contrast, the charge-localized dimethylpiperazine (DMP-L⁺) structure occupies a local minimum on the potential energy curve and has a positive charge that is localized on only one of the nitrogen atoms. The

CCSD_CCSD(T)-SP and MP2_CCSD(T)-SP legend labels in Fig. 3.1 denote single-point (SP) energy calculations that were carried out with the CCSD(T) method using geometry-optimized structures obtained with CCSD and MP2, respectively. To maintain a consistent comparison with the previous study by Weber et al., the same basis sets from Ref. [57] were used throughout this work (i.e., all optimizations were carried out with the aug-cc-pVDZ (augmented correlation-consistent polarized valence double-zeta) basis set, and single-point energy CCSD(T) calculations utilized the cc-pVTZ (correlation-consistent polarized valence triple-zeta) basis). It is worth noting that Weber et al. did not examine any details of the transition state structure using high-level wavefunction-based calculations, which we provide for the first time in Figures 3.1, 3.2 and 3.3. Upon examination of Fig. 3.1, we observe several clear trends. First, all three wavefunction-based approaches (CCSD, CCSD_CCSD(T)-SP, and MP2_CCSD(T)-SP) are in agreement by producing a potential energy curve with an extremely small energy barrier (<0.01 eV), which is in stark contrast to the much larger 0.2 eV barrier obtained from the PZ-SIC approach. Most interestingly, a single-point CCSD(T) energy calculation on top of the CCSD- and MP2-optimized geometries further lowers the barrier to the point where it more closely resembles the M06-HF potential energy curve. While we take the CCSD_CCSD(T)-SP curve in Fig. 3.1 to be the most accurate calculation among all the methods studied, it is interesting to note that the MP2_CCSD(T)-SP curve still closely resembles both the CCSD_CCSD(T)-SP and M06-HF curves. In addition to the barrier height, the CCSD(T) single-point calculations alter the relative energy difference between the DMP-D⁺ and DMP-L⁺ structures such that CCSD_CCSD(T)-SP and MP2_CCSD(T)-SP curves are even closer in agreement with the M06-HF DFT calcu-

Table 3.1: Relative energies of the DMP-L⁺ and DPM-D⁺ states obtained from various computational methods.

Method	Relative Energy (eV)	
	Barrier Height	Energy (DMP-L ⁺)-Energy (DMP-D ⁺)
PZ-SIC	0.20	0.34
M06-HF	0.00	0.38
CCSD	0.01	0.24
MP2_CCSD(T)-SP	0.00	0.38
CCSD_CCSD(T)-SP	0.00	0.38
Experiment	- ^a	0.33 (0.04) ^b

The PZ-SIC and M06-HF energies were obtained from Ref. [57], and CCSD_CCSD(T)-SP and MP2_CCSD(T)-SP denote single-point energy calculations that were carried out with the CCSD(T) method using geometry-optimized structures obtained with CCSD and MP2, respectively.

^a No value is shown since the experimental barrier height was not provided by Ref. [57].

^b The experimental error in the relative energy difference between DMP-L⁺ and DPM-D⁺ is 0.04 eV.

lations. Table 3.1 summarizes the barrier heights and relative energy differences obtained from PZ-SIC, M06-HF, and the various wavefunction-based methods examined. Taken together, both the small barrier heights and the DMP-D⁺/DMP-L⁺ relative energy differences obtained from the high-level CCSD(T) calculations show good agreement with the DFT methods examined in Ref. [57], and it is actually the PZ-SIC calculation that is the outlier and inconsistent with the highly accurate CCSD(T) benchmarks.

Before proceeding to a final discussion on the PZ-SIC transition-state geometry, we briefly discuss the accuracy of our CCSD(T) calculations, which we used as high-level calculations to benchmark both the PZ-SIC and DFT methods discussed above. First, to check for possible non-dynamical correlation effects in our CCSD(T) calculations, we computed the T1 diagnostic [59, 70] for the DMP-D⁺, DMP-L⁺, and transition-state structures, which resulted in T1 values <0.031 (T1 values greater than 0.044 for open-shell systems in-

dicating that a multi-reference electron correlation method is necessary [70]). Next, to address any possible basis set convergence issues, we also carried out large-scale CCSD(T)-F12/cc-pVTZ (coupled-cluster with single and double excitations and perturbative triples with explicitly correlated F12 corrections) [71] calculations—these methods exhibit dramatic improvements in basis set convergence since they are constructed from a wavefunction that depends explicitly on the interelectronic coordinates (i.e., results of quintuple-zeta quality have been obtained with CCSD(T)-F12 methods, even when triple-zeta basis sets were used [71, 72]). Nevertheless, our explicitly correlated CCSD(T)-F12 calculations are in full agreement with our CCSD(T) calculations by producing an extremely small energy barrier (~ 0 eV) and a relative energy difference of 0.41 eV between the DMP-D⁺ and DMP-L⁺ structures. Finally, the CCSD(T) method is often referred to as the gold standard of quantum chemistry (whereas the performance of the PZ-SIC functional is much less known), and both the CCSD(T) and CCSD(T)-F12 barrier heights are extremely small (essentially barrierless and in agreement with DFT), in stark contrast to the significantly larger 0.2 eV barrier obtained from the PZ-SIC approach.

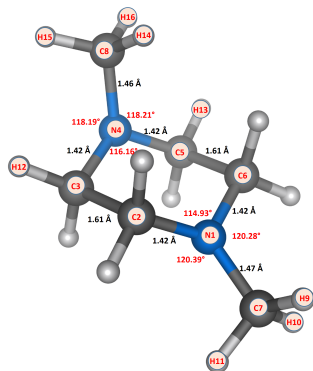


Figure 3.2(a) MP2

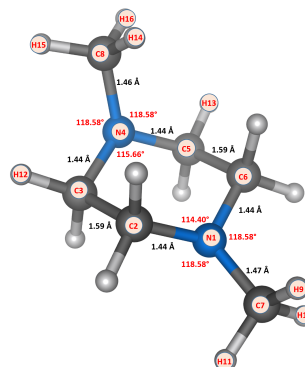


Figure 3.2(b) CCSD

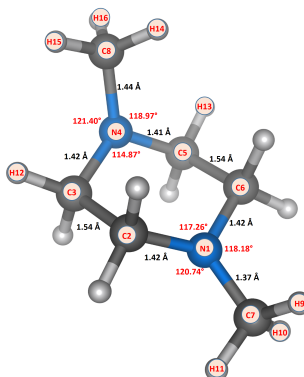


Figure 3.2(c) PZ-SIC

Figure 3.2: Comparison of transition-state geometries for the diamine cation optimized with the MP2, CCSD, and PZ-SIC computational methods.

Finally, we discuss a few discrepancies regarding the transition-state geometries obtained from the PZ-SIC vs. the wavefunction-based approaches. We obtained our CCSD- and MP2-optimized transition-state geometries using the Synchronous Transit-Guided Quasi-Newton (STQN) [71] method which uses a linear/quadratic synchronous transit approach to converge towards a transition-state geometry. Upon convergence, we obtained CCSD- and MP2-optimized transition-states (both exhibiting C_s point-group symmetries) that closely resembled each other. However, due to the extreme computational cost of an open-shell CCSD vibrational frequency analysis, we carried out a vibrational

frequency analysis on the MP2-optimized geometry (which, again, closely resembled the CCSD-optimized transition-state geometry) and obtained a single imaginary harmonic frequency of $301.18i \text{ cm}^{-1}$ that connected the DMP-D⁺ and DMP-L⁺ structures along the potential energy curve depicted in Fig. 3.1 (Weber et al. did not carry out a vibrational frequency analysis in their study). Both the CCSD- and MP2-optimized Cartesian coordinates for DMP-D⁺, DMP-L⁺, and the transition state are provided in the Appendix Tables A.1– A.6. In contrast to the Cs point-group symmetries of the transition-states obtained from MP2/CCSD, the PZ-SIC transition state is somewhat distorted and possesses a lower C1 symmetry. The carbon-nitrogen bond lengths in the PZ-SIC transition-state structure are 0.1 Å smaller than those in the CCSD-optimized transition state; however, the most significant difference between the PZ-SIC and CCSD geometries were the dihedral angles of the methyl hydrogens relative to the DMP⁺ molecule ring, which differed by as much as 33° between the two methods.

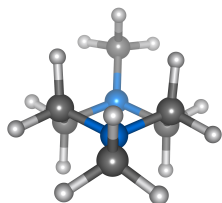


Figure 3.3(a) CCSD (axial view)

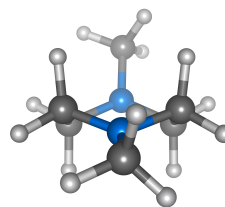


Figure 3.3(b) PZ-SIC (axial view)

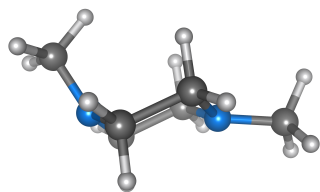


Figure 3.3(c) CCSD (side view)

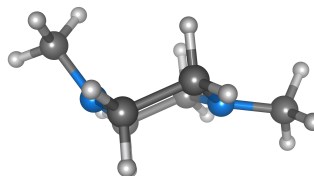


Figure 3.3(d) PZ-SIC (side view)

Figure 3.3: Axial and side views of transition-state geometries for the diamine cation optimized with the CCSD and PZ-SIC computational methods.

3.3 Conclusion

In conclusion, we have carried out new high-level CCSD(T) analyses on the DMP^+ cation to investigate the surprising claim that “all DFT functionals commonly used today, including hybrid functionals with exact exchange, fail to predict a stable charge-localized state” for this relatively simple system. Our new high-level CCSD(T) analyses on the DMP^+ cation show that DFT actually performs quite well for this system, whereas the PZ-SIC approach used by Weber et al. is the outlier that is inconsistent (and predicts a significantly larger barrier height), compared to the highly accurate CCSD(T) benchmarks. Although the experiments by Weber et al. appear to give excellent agreement with their PZ-SIC approach, it should also be noted that their rationale for a charge-localized state of

Table 3.2: Comparison of various geometric parameters for the diamine cation transition state obtained with the MP2, CCSD, and PZ-SIC computational methods. The atom labels and numbers refer to the labeling scheme used in Figure 3.2. The shaded entries denote geometric parameters with the largest variation between the MP2/CCSD and PZ-SIC approaches.

	Bond Length (Å)				Dihedral Angle (degrees)		
	MP2	CCSD	PZ-SIC		MP2	CCSD	PZ-SIC
C7-N1	1.47	1.47	1.37	C3-C2-N1-C7	-141.54	-150.91	-148.89
N1-C2	1.42	1.44	1.42	C2-C3-N4-C8	-86.80	-89.82	-94.33
C2-C3	1.61	1.549	1.54	C6-C5-N4-C8	86.79	89.82	94.75
C3-N4	1.42	1.44	1.42	C5-C6-N1-C7	141.53	150.91	148.51
N4-C5	1.42	1.44	1.41	C2-N1-C7-H9	-73.83	-73.17	-90.63
C5-C6	1.61	1.59	1.54	C2-N1-C7-H10	161.31	166.14	151.76
C6-N1	1.42	1.44	1.42	C2-N1-C7-H11	43.03	47.52	28.90
N4-C8	1.45	1.46	1.44	C6-N1-C7-H9	77.73	73.18	66.55
				C6-N1-C7-H10	-43.13	-47.50	-51.06
				C6-N1-C7-H11	-161.42	-166.12	-173.91
				C3-N4-C8-H14	74.26	74.53	107.13
				C3-N4-C8-H15	-45.07	-44.92	-12.45
				C3-N4-C8-H16	-166.40	-166.02	-133.79
				C5-N4-C8-H14	-74.51	-74.61	-47.18
				C5-N4-C8-H15	166.16.94	-165.94	166.77
				C5-N4-C8-H16	44.84	44.84	71.89
				N1-C2-C3-H12	-175.99	-175.69	-173.57
				N1-C6-C5-H13	176.703	175.70	172.39

	Bond Angle (degrees)		
	MP2	CCSD	PZ-SIC
C7-N1-C2	120.39	118.59	120.74
C7-N1-C6	120.28	118.58	118.11
C8-N4-C3	118.19	118.58	121.40
C8-N4-C5	118.21	118.58	118.97
C3-N4-C5	116.16	115.66	114.87
C2-N1-C6	114.40	117.26	114.93

the DMP-L⁺ cation was inferred from time-resolved measurements of the Rydberg states of the cation rather than the ground-state potential energy surface of the cation itself. While there are certainly cases where self-interaction corrections are essential for obtaining correct results in pathological chemical systems, the potential energy surface of the diamine cation, unfortunately, is not one of them.

Chapter 4

Harnessing Semi-Supervised Machine Learning to Predict Per- and Polyfluoroalkyl Substances (PFASs) Bioactivities

This chapter investigates machine learning techniques to explore PFASs bioactivity, which resulted from collaborating with a former lab colleague, Hyuna Kwon. She applied machine learning and auto docking methods to the datasets. We helped build and analyze the quantitative-structure activity relationship (QSAR) model and published it in *the Journal of Environmental Science & Technology Letters* [73]. The supplementary information is available in Appendix B, which contains the unsupervised results.

4.1 Introduction

Since the 1930s [74], per- and polyfluoroalkyl substances (PFASs) have been used in several consumer products (including fire-fighting foams) due to their outstanding stability and water/oil repellent properties [75]. However, these compounds pose significant risks to the environment and biosystems. The presence of PFASs in surface water and groundwater can result in exposure to organisms, subsequently leading to accumulation in the body, with adverse effects on the liver, kidneys, blood, and immune system [75, 76]. Because of these deleterious effects, there is a pressing need to identify and understand the bioactivity of PFAS-based compounds that can adversely affect human health.

For these reasons, several international groups including the Organization for Economic Cooperation and Development (OECD), United States Environmental Protection Agency, Food and Drug Administration, European Chemicals Agency, European Food Safety Authority, and Ministry of Ecology and Environment (China) continue to monitor PFASs that are produced in the global market [77, 78]. According to a 2018 OECD report, more than 4,700 PFASs currently exist as manufacturers bring new forms of PFASs into industrial and consumer products (it is worth pointing out, however, that not all 4,700 structures exist in commerce). Nevertheless, among the wide varieties of PFAS molecules, the potential hazards of these new forms remain largely unknown.

Due to the sheer number of PFAS species, *in vivo* and *in vitro* biological experiments are both time-consuming and costly. As such, the construction of predictive and reliable quantitative-structure activity relationship (QSAR) models [79–81] is essential for assessing the bioactivities of these contaminants (even for PFAS species that are yet to

be made). Specifically, a QSAR model that can accurately predict the bioactivities of PFASs can be harnessed to screen several of these contaminants, saving immense time and experimental resources. While there have been prior machine learning studies on PFAS molecules [82, 83], most of these approaches used supervised learning techniques to suggest *general* structure-bioactivity trends after post-processing of the data (i.e., the focus was on aggregate data for all targets as opposed to analyzing chemical trends specific to each target).

In this work, we present a new QSAR model using semi-supervised metric learning techniques to assess which functional groups affect bioactivities toward specific biological targets. Semi-supervised learning is a different machine learning approach that has the advantages of both supervised and unsupervised learning. It can be used on a dataset with primarily unlabeled data and only a few labeled data. Like unsupervised learning, it can also automatically cluster unlabeled data. Our approach is integrated with molecular docking calculations to predict possible bioactivities of PFAS molecules based on their chemical functional groups and specific biological targets (e.g., genes, proteins, or cell lines). Our approach first combines dimension reduction methods with clustering methods to classify PFASs based on their molecular structures. We then apply a semi-supervised metric learning method to improve classification accuracy. Finally, we use a molecular docking approach to shed light on the physicochemical reasons for their bioactivity. Our study provides the first unsupervised/semi-supervised learning approach for screening potentially bioactive PFAS molecules beyond conventional supervised learning or QSAR approaches.

4.2 Methodology

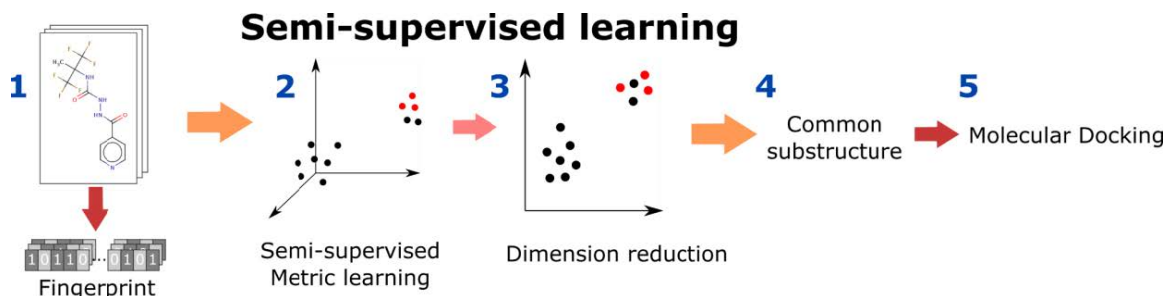


Figure 4.1: Machine-learning-based workflow for QSAR construction to predict bioactivity of PFASs.

Our QSAR machine-learning framework, shown in Figure 4.1, utilizes four sequential steps followed by a reasoning/validation step: (1) collecting a training dataset from verified open-source databases, (2) encoding those compounds into molecular fingerprints, (3) clustering the data to predict chemical properties based on the molecular fingerprints and assessing the performance of the models, (4) evaluating the clustering by choosing the optimal model and predicting molecular groups responsible for bioactivity based on the clustering, and (5) molecular docking simulations to rationalize the role of the chemical functional groups. All of our machine learning algorithms are publicly available (see Appendix B).

In our first step, we obtained datasets from comprehensive open-source databases, including PubChem’s BioAssay [84], Maximum Unbiased Validation [85], Toxicology in the 21st Century [86], beta-secretase 1 [87], and blood-brain barrier penetration datasets [88], which are available from the Supporting Information of Ref. [83]. We used two different datasets without further modification from Ref. [83]: (1) the CF dataset, which includes substances containing at least one $-\text{CF}-$ moiety (62,043 molecules), and (2) the C3F6

dataset, which includes substances containing a perfluoroalkyl moiety with three or more carbons (1,012 molecules). For both datasets, we used bioactivity data against 26 biological targets.

Encoding the compounds to molecular fingerprints followed next in our framework. We used the extended connectivity fingerprint (ECFP) featurization [89] with a default diameter of 4 (i.e., ECFP4), which considers a maximum of four neighbors. ECFPs are topological molecular representations developed for substructure and similarity searching. By encoding molecular structures into fingerprints, we obtained a binary array with a constant length of 2,048, making it a convenient input for the unsupervised/semi-supervised learning models. Furthermore, since the simplified molecular-input line-entry system (SMILES) sequences for all PFAS molecules are readily available, they can be easily converted into fingerprint-based representations using the RDKit software package [90].

We then applied semi-supervised metric learning to the generated fingerprints by training machine learning models to predict the bioactivities of PFAS molecules by first **(a)** *reducing the dimension of the fingerprint datasets* and then **(b)** *classifying/clustering them* (see Figure 4.1). Our QSAR model used a semi-supervised metric learning algorithm to automatically group/classify molecules with similar bioactivities. Metric learning has two main advantages: **(1)** its predictions are more efficient/accurate since the model distinctly separates new molecular representations according to their bioactivities (by reducing the distance metric between the same-labeled pair of data and increasing the distance between opposite-labeled pair of data), and **(2)** it automatically generates a vector-shaped representation from the molecular fingerprint and can be directly integrated with conventional

dimension reduction methods. The final clusters were selected based on the best Silhouette score, which analyzes the distances of each data point to its cluster and neighboring clusters [91]. In short, a higher Silhouette score indicates more distinct and separated clusters. We then identified which substructures or molecular functional groups played essential roles in determining the bioactivity of the molecules.

Lastly, we conducted several molecular docking calculations using Autodock [92] to elucidate the physicochemical reasons for the bioactivity trends obtained from our QSAR model (i.e., using ligand-protein binding conformations to rationalize the role of chemical substructures that induces bioactivity on biological targets.)

4.3 Results and Discussion

4.3.1 Unsupervised vs. semi-supervised machine learning

To systematically evaluate the performance of our semi-supervised metric approach, we first performed traditional unsupervised machine learning and compared the performance of the two models. To maintain a concise discussion of our results, Appendix B contains a detailed analysis and comparison of our unsupervised vs. semi-supervised machine learning results. Figure B.1 shows our clustering results using unsupervised machine learning on the C3F6 dataset, and Figure B.2 shows a comparison between the unsupervised and semi-supervised results using the CF dataset on two different targets. Table B.3 summarizes the substructures that induce bioactivity as predicted from our unsupervised learning calculations. In summary, our extensive analyses in the Supporting Information showed that semi-supervised metric learning performed significantly better than unsupervised machine

learning; as such, we only focus on the results of the former in this manuscript.

4.3.2 Semi-supervised metric learning

Figure 4.2 displays true-positive ratios and classifications between bioactive/inactive molecules on four representative targets that show the best performance in the CF dataset using semi-supervised metric learning (for example, in Fig. 4.2a, we obtain a true-positive ratio of 97.3% by computing $\frac{\# \text{ of molecules containing esters and are also bioactive}}{\# \text{ of ester-containing molecules in the cluster}}$. Using the Maximum Common Structure (MCS) module in the RDKit software package on bioactive molecules, we found that the ester functional group is the critical substructure that causes bioactivity on Cyps (Figures 4.2a, b, and c) and ATXN (Figure 4.2d). Table B.4 summarizes the substructures predicted to play a vital role in bioactivity toward nine different targets. The other 17 targets did not demonstrate as distinct clustering as the nine targets in Table B.4 due to a relatively weak correlation between molecular structure and bioactivity.

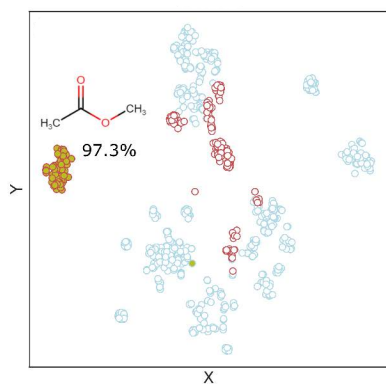


Figure 4.2(a) CYP2C9

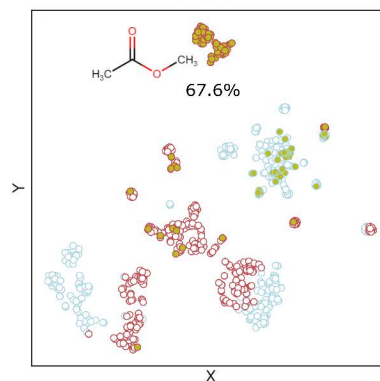


Figure 4.2(b) CYP3A4

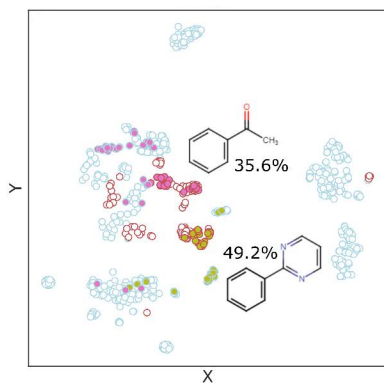


Figure 4.2(c) CYP2D6

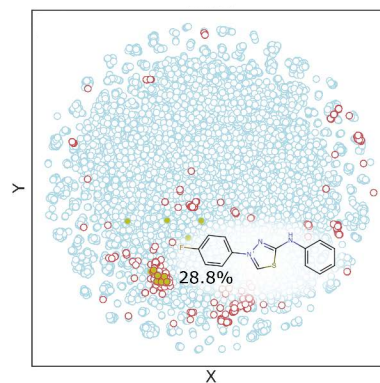


Figure 4.2(d) ATXN

Figure 4.2: CF dataset molecule distribution using semi-supervised metric learning. Each point represents either a bioactive (red circular edges) or an inactive (light blue circular edges) molecule towards (a) CYP2C9, (b) CYP3A4, (c) CYP2D6, and (d) ATXN. The olive green-filled circles represent molecules having the substructure depicted in the plot; i.e., (a, b) ester groups, (c) phenylprimidyl groups, and (d) 4-benzyl-2-(4-fluorophenyl)-1,2-thiazole. The pink-filled circles in (c) represent molecules with phenylethanone. The percentage value represents the ratio of the number of bioactive molecules within the identified substructure. Table B.3 lists the predicted substructures for specific targets.

We used structural alerts to cross-check the validity of the predicted substructures that play a crucial role in bioactivity. Within the bioinformatics community, structural alerts are molecular functional groups associated with a particularly adverse outcome, in our case, bioactivity [93, 94]. We cross-referenced the ChEMBL dataset to our machine

learning results since it contains structural alert information for some PFAS molecules [95]. Figure B.3 shows structural alerts of the molecules that are bioactive on CYP2C9, and, as mentioned previously, the ester group was found to be the critical structure that induces interaction with Cyps [96, 97].

4.3.3 Interactions between PFASs and targets

We carried out molecular docking calculations with Autodock [94] to rationalize the underlying molecular causes of bioactivities in PFAS and predict their interaction with target enzymes. Appendix B gives additional details of our molecular docking calculations. We successfully docked all PFASs into the active sites of the targets and binned the binding affinity results based on their bioactivity with the target. Figure B.5 displays one of the bioactive structures with the ester group of the CYP2C9-PFAS complex, methyl 4-[2-propyl-1-([4-trifluoromethyl]phenyl)sulfonylamino)-2-hexen-1-yl]benzoate.

To verify the correlation between the Autodock binding affinities and their bioactivity, we performed a dimension reduction procedure using unsupervised learning on the CF dataset, which consists of molecular structures with binding affinity data (see Figure 4.3). We used unsupervised learning here to make the point that unsupervised learning underperforms when only structural data is provided. Specifically, if the classification accuracy is improved with additional feature inputs, those features must contain some information to discriminate among the population [98, 99]. In other words, if the inclusion of binding affinity data enhances the clustering accuracy, it provides another co-descriptor for bioactivity. Indeed, Figures 4.3b and 4.3a show that descriptors consisting of chemical structures *and* binding affinity data give a better separation/distinction between active and inactive

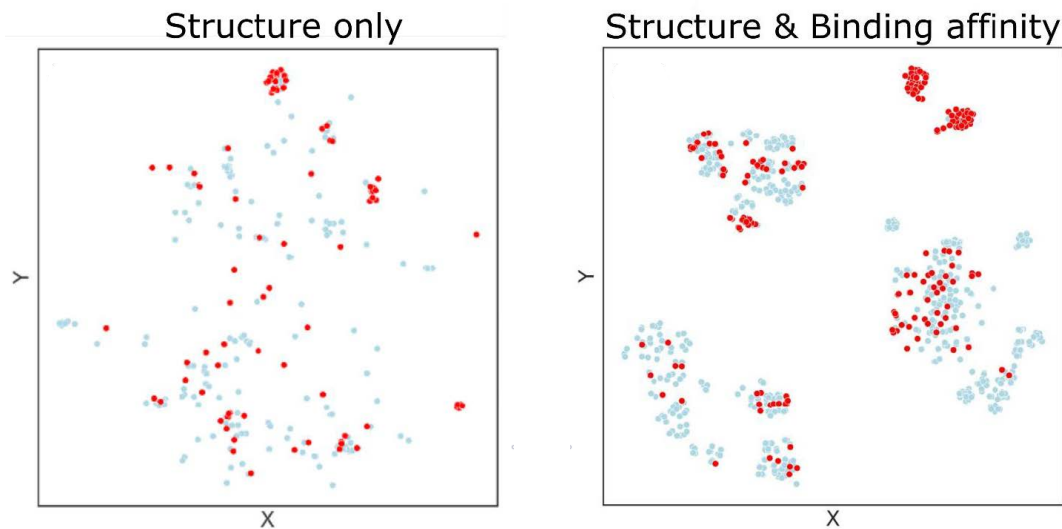


Figure 4.3(a)

Figure 4.3(b)

Figure 4.3: Clustering of molecules predicted with unsupervised learning (dimension reduction) on CF datasets containing (a) chemical structures and (b) chemical structures and binding affinities with CYP2C9. Each point represents a molecule that is either bioactive (red) or inactive (blue) towards CYP2C9.

molecules compared to the unsupervised learning results based only on chemical structures.

4.3.4 Bioactivity predictions on OECD dataset

In 2018, the Global Perfluorinated Chemicals Group [100] within the OECD published a list of 4,730 PFASs to develop regulatory approaches for reducing the use of perfluorinated substances in products. However, researchers have yet to discover the bioactivities of the molecules in the list. Using the QSAR model developed in this work, we give predictions and a rationale for the bioactivities of molecules in the OECD list.

We performed molecular docking calculations on molecules containing the ester group among the OECD list to verify similar binding conformations. Of the 4,730 PFASs in the OECD list, 414 have an ester functional group. Figure B.6 shows four different repre-

sentative ester-containing molecules bound to CYP2C9. In particular, the ester-containing molecules in the OECD list bind strongly with Fe^{2+} of the HEME group (an active site of Cyp enzyme), which is similar to the binding interactions that we observed in the CF dataset. Therefore, we expect a large portion of the 414 ester-containing molecules among the OECD list to form strong bonds with Fe^{2+} of the HEME group with a similar conformation, leading to bioactivity toward Cyp enzymes. Furthermore, based on our docking calculations, 87.7% of these 414 molecules have a stronger binding affinity than -5 kcal/mol (the average binding affinity is -5.77 kcal/mol), which falls in the range of the mean binding affinity of the bioactive molecules from the CF dataset.

We then clustered the OECD dataset into 40 clusters using the k-means clustering method. Using both the clustered results (Figure 4.4b) and the distribution of ester-group-containing molecules (Figure 4.4a), we found that clusters 13, 25, and 39 contain ester functional groups. Analyzing the CF dataset, we found that the ester group plays a possible role in bioactivity toward Cyp enzymes; that is, molecules in these clusters have a high probability of being bioactive against CYP2C9 and CYP3A4.

4.4 Conclusion

In summary, we have developed a new QSAR model validated with ChemMLB structural alerts and molecular docking calculations, which constitutes the first application of semi-supervised metric learning for predicting/rationalizing bioactivities in PFASs. Using a semi-supervised metric learning algorithm, our machine-learning-based QSAR model accurately identified specific substructures, such as ester-containing groups, that play a possible

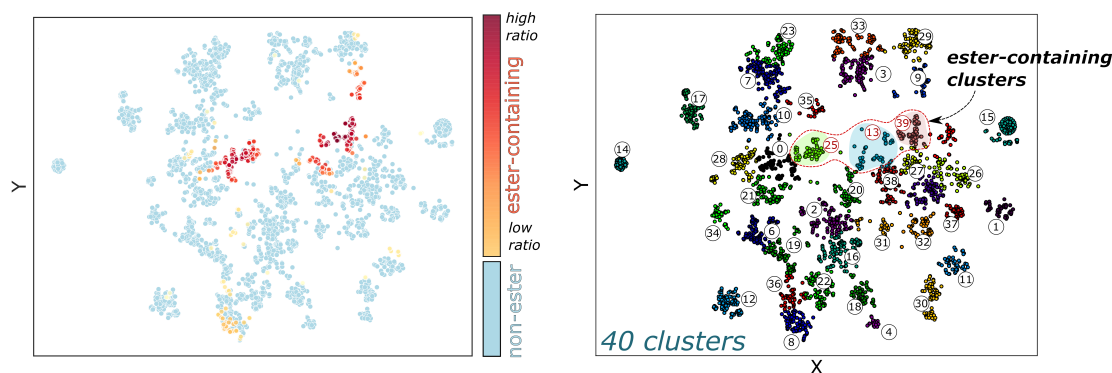


Figure 4.4(a)

Figure 4.4(b)

Figure 4.4: (a) OECD dataset classified by PC t-SNE and clustered based on the k-means clustering method. The orange and yellow dots represent ester-containing molecules. The colors closer to red (yellow) represent a higher (lower) concentration of bioactive molecules. (b) PFAS molecules included in the OECD list are grouped into 40 clusters. Each point represents a molecule, and clusters 13, 25, and 39 denote a high ratio of ester-containing groups

role in determining bioactivities. With our semi-supervised learning approach, we obtained a distinct classification between bioactive and inactive molecules, resulting in an accuracy of up to 97.3% in the CF dataset. We also used semi-supervised metric learning to automatically classify/cluster and predict functional groups that could possibly play a role in bioactivity.

In addition, our machine learning model proposed a few significant substructures that could induce bioactivity, which were subsequently examined with molecular docking calculations. Most importantly, our machine learning predictions on bioactivities can provide a more efficient screening of potentially bioactive PFASs that can be used to complement *in vitro* assessments. All of our machine learning algorithms are publicly available (see Appendix B), and we anticipate that researchers can further extend our methodology to screen other contaminants or analyze the potential bioactivity of PFAS molecules.

Chapter 5

Efficient Predictions of Formation Energies and Convex Hulls from Density Functional Tight Binding Calculations

This chapter's work resulted from collaborating with a former lab colleague, Anshuman Kumar. Calculations aside, he exclusively performed the integration and application of CASM with DFTB. We contributed equally to analyzing the convex hull results and writing the manuscript. When writing this chapter, the manuscript is in *review* [\[101\]](#).

5.1 Introduction

Point defects, such as vacancies, play a vital role in the electronic and structural properties of semiconductor materials. These properties ultimately dictate the performance of electronic devices, structural alloys, and functional materials; therefore, a deep understanding of vacancies at the atomistic level can provide a rational path towards their improvement [102–105]. One of the most important quantities to characterize a defect is its formation energy, which can, in principle, be computed with Density Functional Theory (DFT) [106, 107]. Although accurate and widely transferable, DFT can be computationally prohibitive for the routine exploration of the numerous types of defects in material systems. Alternatively, semi-empirical methods, such as Density Functional Tight Binding (DFTB) [108–111], have recently emerged as efficient approaches for addressing these computational bottlenecks. In particular, the DFTB formalism has already been used to calculate a variety of large systems such as metallic nanoparticles [112, 113], explicitly-solvated chromophores [114], massive biological structures [115], molecules/clusters with numerous conformations [116, 117], and immense nanostructures [118].

To enable fast and accurate calculations of formation energies, we combined DFTB with the Clusters Approach to Statistical Mechanics (CASM) [119, 120] software package to predict thermodynamically stable phases of materials for the first time. Specifically, our new capability allows the rapid and accurate calculation of formation energies and the convex hull (if favorable). In short, the convex hull provides a global view of the relative stabilities of structures after the formation energies are calculated. Although computationally demanding, the calculation of formation energies and convex hulls has enabled the discovery

of new materials, including superconducting hydrides [121–125], metal nitrides [126], and metal carbides [127]. Predicting the convex hull for general materials is time-consuming since it requires the classification of a vast number of energy minima on the lattice energy surface. Software programs, such as USPEX [128, 129] and CALYPSO [130, 131] have been used to explore the vast compositional phase space of these materials. Compared to other crystal structure prediction codes, the main advantage of CASM is its ability to evaluate the kinetic and thermodynamic properties of multi-component crystalline solids using group theoretical techniques.

Using this new capability, we calculate the formation energies of various silicon carbide (SiC) and zinc oxide (ZnO) configurations to highlight the efficiency of our DFTB+CASM implementation. SiC is one of the most promising materials for high-temperature, radiation-resistant, power and high-speed electronics [132–135]. ZnO is an affordable, earth-abundant, wide band gap transparent conducting oxide with applications in electronics, optoelectronics, pharmaceuticals, sensors, and catalysis [136]. ZnO crystallizes in many forms, with hexagonal wurtzite (B4), zinc blende (B3), and cubic rocksalt (B1) being the most common. Using SiC and ZnO as representative examples, we compare the accuracy and efficiency of DFTB and DFT for predicting formation energies and convex hulls of these binary compounds.

Our paper is organized as follows: Section 5.2 gives a brief overview of DFTB, DFT, structure generation algorithm in CASM, formation energies, and convex hulls. Section 5.3 provides computational details and Section 5.4 presents our results and discussion. Finally, we conclude with closing remarks and a summary in Section 5.5.

5.2 Theory and Methodology

5.2.1 DFTB

Before proceeding with a discussion of our approach for evaluating the formation energy/convex hull, we briefly review the DFTB formalism. DFTB is based on a Taylor series expansion of the DFT Kohn-Sham (KS) total energy, E_{KS} , with respect to electron density fluctuations $\rho(r) = \rho_0(r) + \delta\rho(r)$, where $\rho_0(r)$ is a reference density of neutral atomic species. In our work, we use the second-order expansion of the KS energy, abbreviated as DFTB2 [137]. The unmodified KS total energy is given by

$$E_{\text{KS}} = \sum_i^{\text{occ}} \langle \psi_i | -\frac{1}{2}\nabla^2 + V_{\text{ext}} | \psi_i \rangle + E_{\text{H}} + E_{\text{xc}} + E_{\text{II}}, \quad (5.1)$$

where ψ_i are the KS orbitals, V_{ext} is the external potential, E_{H} is the Hartree energy, E_{XC} is the exchange-correlation (XC) energy, and E_{II} is the ion-ion interaction energy. Rewriting Eq. 5.1 in terms of $\rho(r)$ and expanding up to second order, we obtain the DFTB energy:

$$\begin{aligned} E_{\text{DFTB}} &= \sum_i^{\text{occ}} \langle \psi_i | \hat{H}_0 | \psi_i \rangle + \frac{1}{2} \sum_{\text{AB}}^{\text{M}} \gamma_{\text{AB}} \Delta q_{\text{A}} \Delta q_{\text{B}} + E_{\text{rep}}^{\text{AB}} \\ &= E_{\text{BS}} + E_{\gamma} + E_{\text{rep}}, \end{aligned} \quad (5.2)$$

where the second term sums over the number of atoms, M, in the system.

The first term in Eq. 5.2, E_{BS} , corresponds to the band structure energy (i.e., the sum over the occupied orbital energies) obtained from the diagonalization of the non-self-consistent DFTB Hamiltonian, \hat{H}_0 :

$$\hat{H}_0 = \langle \phi_{\mu} | \hat{T} + \nu_{\text{eff}}[\rho_0^{\text{A}} + \rho_0^{\text{B}}] | \phi_{\nu} \rangle, \quad \mu \in \text{A}, \nu \in \text{B}, \quad (5.3)$$

where ϕ_{μ} , ϕ_{ν} forms a minimal Slater type atomic basis (μ and ν are the indices of the valence atomic basis function centered on atoms A and B, respectively), \hat{T} is the kinetic

energy operator, ρ_0^I is the reference density of neutral atom I, and ν_{eff} is an effective Kohn-Sham potential. As shown in Eq. 5.3, only two-center elements are treated within the DFTB framework, which are explicitly calculated using analytical functions within the linear combination of atomic orbitals (LCAO) formalism. The Hamiltonian and overlap matrix elements are pre-computed and stored in Slater-Koster (SK) files for all pairs of chemical elements as a function of the distance between atomic pairs. Thus, no explicit integral evaluation occurs during the simulation, which significantly improves the computational efficiency of the DFTB approach. The second term in Eq. 5.2, E_γ , is the energy due to charge fluctuations, where γ_{AB} is an analytical function of the interatomic distance, and the Hubbard parameter U and takes into account the electron-electron interaction. The Hubbard parameter U is related to the hardness of the atoms and controls how the electron density is distributed between the atoms. The $\Delta q_{A/B}$ term ($= q_{A/B} - q_{A/B}^0$) is the net charge of atom A/B. The summation in the E_γ term in Eq. 5.2 is performed over the number of atoms, M, in the system. The last term, E_{rep} , is the distance-dependent diatomic repulsive potential, which includes core-electron effects, ion-ion repulsion, and a portion of exchange-correlation effects. The pairwise repulsive functions are obtained by fitting to DFT calculations using a suitable reference structure and, like the matrix elements, are pre-tabulated. By applying the variational principle, we obtain the DFTB Kohn-Sham equations:

$$\sum_B \sum_{\nu \in B}^M c_{\nu i} (H_{\mu\nu} - \epsilon_i S_{\mu\nu}) = 0, \quad \forall A, \mu \in A, i, \quad (5.4)$$

where the DFTB Hamiltonian is given by

$$H_{\mu\nu} = \langle \phi_\mu | \hat{H}_0 | \phi_\nu \rangle + S_{\mu\nu} \sum_\xi^M \Delta q_\xi \left(\frac{1}{2} (\gamma_{A\xi} + \gamma_{B\xi}) \right), \quad (5.5)$$

with $\mu \in A$, $\nu \in B$, and $S_{\mu\nu}$ is the overlap matrix of the atomic orbitals. Because the atomic charges are dependent on the one-particle wave functions, ψ_i , Eq. 5.5 must be solved iteratively until self-consistency is reached.

DFTB is a tight-binding-based method and provides a reasonable accuracy at a much lower computational cost compared to DFT calculations. As mentioned earlier, the Hamiltonian and the overlap matrix elements $H_{\mu\nu}$ and $S_{\mu\nu}$ are pre-computed and tabulated; i.e., they are not computed during the execution of the program. This and the use of a minimal valence basis set can typically lead to huge computational savings (2-3 orders of magnitude) compared to full DFT.

5.2.2 DFT

In Kohn-Sham DFT (KS-DFT), the electronic energy of a system is given by

$$E[\rho] = T_{\text{KS}}[\rho] + E_{\text{ext}}[\rho] + E_{\text{H}}[\rho] + E_{\text{xc}}[\rho]. \quad (5.6)$$

All of the terms in Eq. 5.6 are a functional of the density ρ , where $T_{\text{KS}}[\rho]$ is the kinetic energy of the electrons, $E_{\text{ext}}[\rho]$ is the attractive interaction between the electrons and the nuclei (also known as the external potential energy), $E_{\text{H}}[\rho]$ is the classical (Coulomb) electrostatic energy of the electronic density charge distribution (also known as the Hartree energy), and $E_{\text{xc}}[\rho]$ is the exchange-correlation energy which describes exchange and correlation quantum interactions.

DFT calculations are performed in an iterative way, with the electron density, $\rho(\vec{r})$, expressed as the sum of one-electron wave functions, ψ_i , known as molecular orbitals

(MOs):

$$\rho(\vec{r}) = \sum_{i=1}^{N_e} |\psi_i(\vec{r})|^2. \quad (5.7)$$

These MOs are obtained by solving the Kohn-Sham (KS) eigenvalue equation:

$$\left[-\frac{1}{2}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + V_{\text{H}}[\rho](\mathbf{r}) + V_{\text{xc}}[\rho](\mathbf{r}) \right] \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}), \quad (5.8)$$

where $V_{\text{H}}[\rho]$ and $V_{\text{xc}}[\rho]$ are themselves functionals of the density (making them functionals of the MOs), which are obtained after solving the KS eigenvalue problem in Eq. 5.8. As such, the KS eigenvalue problem in Eq. 5.8 cannot be solved directly and must be calculated in an iterative fashion within a self-consistent field (SCF) method. In this iterative process, a set of guess wavefunctions, $\{\psi_i\}$, are used to compute the terms in Eq. 5.8. This then allows the calculation of a new set of $\{\psi_i\}$, and the process repeats until $\{\psi_i\}$ and the energy in Eq. 5.6 are converged.

One of the key computational bottlenecks in KS-DFT is the numerous three-dimensional integrals that are evaluated in each SCF step, which is extremely time consuming. Since DFTB uses pre-computed integrals in each SCF cycle, its performance can be significantly faster (up to 2-3 orders of magnitude) than full DFT calculations.

5.2.3 Structure Generation with CASM

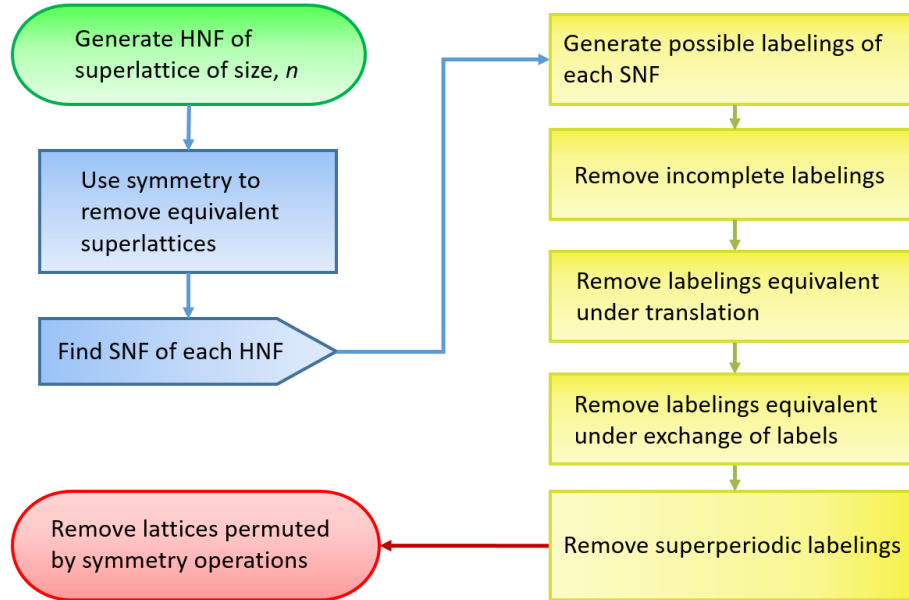


Figure 5.1: Flowchart of the structure generation algorithm used in the CASM software package to enumerate structures.

The algorithms in the CASM software package that enumerate symmetrically distinct configurations utilize an approach based on Hermite Normal Forms of integer matrices [138, 139]. The following is a brief outline of the algorithm [140]:

1. All Hermite normal form (HNF) matrices are generated for each superlattice of size n .
2. Symmetry of the parent lattice is used to remove rotationally equivalent superlattices, thus shrinking the list of HNF matrices.
3. For each index n of the superlattice, the Smith normal form (SNF) is determined for each HNF in the list.
 - (a) A list of possible labelings (atomic configurations) is generated for each SNF,

which is a list of all k^n numbers in a base k , n -digit system. For the labels, the first k letters of the alphabet, (a, b, \dots) are used.

- (b) Incomplete labelings, where each of the k labels (a, b, \dots) does not appear at least once, are removed.
 - (c) Labelings that are equivalent under the translation of the parent lattice vectors are removed. This reduces the list labelings by a factor of $\sim n$.
 - (d) Labelings that are equivalent under an exchange of labels, i.e., $a \rightleftharpoons b$, are removed (for example, the labeling $aabbaa$ is removed from the list because it is equivalent to $bbaabb$).
 - (e) Superperiodic labelings that correspond to a non-primitive superstructure are removed. This can be done without using the geometry of the superlattice.
4. Labelings are removed for each HNF that are permuted by symmetry operations (of the parent lattice) that leave the superlattice fixed.

An important feature of this algorithm is that the list of possible labelings, generated in step (3)(a), forms a minimal hash table with a perfect hash function. All duplicate labelings in a list of N can be eliminated and accomplished in $O(N)$ time. Coupled with the group-theoretical approaches in CASM, this results in a highly efficient algorithm that is orders of magnitude faster than the algorithm of Ferreira, Wei, and Zunger [141].

5.2.4 Formation Energy

The formation energy, e^f , is normalized per primitive unit cell of a particular atomic configuration, σ . For a binary compound AB_x (in this work, atom A is Si/Zn, and

atom B is C/O), the formation energy can be calculated with the expression:

$$e^f(\sigma) = e(\sigma) - e^{\text{ref}}(x), \quad (5.9)$$

where $e^f(\sigma)$ is the formation energy of configuration σ , $e(\sigma)$ is the DFT/DFTB total energy (normalized per primitive unit cell) of configuration σ , and $e^{\text{ref}}(x)$ is the DFT/DFTB total energies (normalized per primitive unit cell) of reference state with composition x . The energy of the reference state, $e^{\text{ref}}(x)$, is calculated from the following expression:

$$e^{\text{ref}}(x) = e^{\text{ref}}(x_1) + (x - x_1) \frac{e^{\text{ref}}(x_2) - e^{\text{ref}}(x_1)}{x_2 - x_1}, \quad (5.10)$$

where $e^{\text{ref}}(x_1)$ and $e^{\text{ref}}(x_2)$ are the DFT/DFTB calculated total energies (normalized per primitive unit cell) of the reference states with composition x_1 and x_2 , respectively. The composition, x , can be calculated from the expression:

$$x = 1 - (B_n/A_n), \quad (5.11)$$

where B_n and A_n denote the number of C (O) and Si (Zn) atoms in the unit cell, respectively. A value of $x = 1$ implies that the unit cell has only Si (Zn) atoms, while $x = 0$ implies that the unit cell consists of an equal number of Si (Zn) and C (O) atoms. In this study, we chose x_1 and x_2 as 0 and 1, respectively. Setting $x_1 = 0$ and $x_2 = 1$ in Eq. 5.10, the reference state energy, $e^{\text{ref}}(x)$, for a composition x simplifies to:

$$e^{\text{ref}}(x) = e^{\text{ref}}(x_1 = 0) + x \left(e^{\text{ref}}(x_2 = 1) - e^{\text{ref}}(x_1 = 0) \right). \quad (5.12)$$

5.3 Computational Details

The general workflow of our calculations is depicted in Fig. 5.1. We explore thermodynamically stable compositions for binary compounds of the form $A_2B_{2(1-x)}$ using the

following four steps: (1) determination of the most stable crystal structure at a fixed composition x using CASM integrated with DFTB, (2) calculation of the formation energy e^f of the compound with respect to its composition at $x = 0$ and $x = 1$, (3) repeating the same calculations by changing x , and (4) plotting the formation energy per fixed composition, e^f vs. x . Details of each step are given below.

5.3.1 CASM

To generate the configuration space for the various material compositions, we start with the primitive unit cell of 2H-SiC and B4-ZnO (the 2H prefix denotes a two-layer hexagonal symmetry stacking periodicity). The 2H-SiC unit cell is shown in Fig. 5.3, where the large green atoms represent silicon, and the small gray atoms are carbon. The crystal structure is an AB-type covalent bond crystal, and each Si atom is surrounded by four C atoms. Each 2H-SiC unit cell contains two Si and two C atoms. The space group of 2H-SiC is $P6_3mc$, and the lattice parameters are shown in Table 5.1. B4-ZnO (Wurtzite) has a similar unit cell as 2H-SiC, but is not shown here for brevity. Table 5.1 also lists the lattice parameters of the 2H-SiC and B4-ZnO unit cells used to generate different configurations. Both 2H-SiC and B4-ZnO follow the AB type stacking sequence. We consider a binary ordering between two C/O atoms and a vacancy at the C/O lattice sites in the unit cell. All symmetrically distinct supercells and derivative configurations (up to a supercell volume that is 4 times the primitive unit cell) were generated using the algorithm developed by Hart et al. [140, 142] in the CASM code. This algorithm enumerates superlattices and atomic configurations in a geometry-independent way using the concept of quotient groups associated with each superlattice to determine all unique

Table 5.1: Initial lattice parameters of SiC (2H) and ZnO (B4).

Structure	Lattice Vectors (Å)			Lattice Angles		
	a	b	c	α	β	γ
SiC (2H)	5.54	6.09	6.43	105.50°	89.91°	117.18°
ZnO (B4)	5.56	6.19	6.45	105.21°	89.94°	116.80°

atomic configurations. We follow the standard procedure for calculating the formation energy/convex hull [143] using DFT/DFTB with the CASM code.

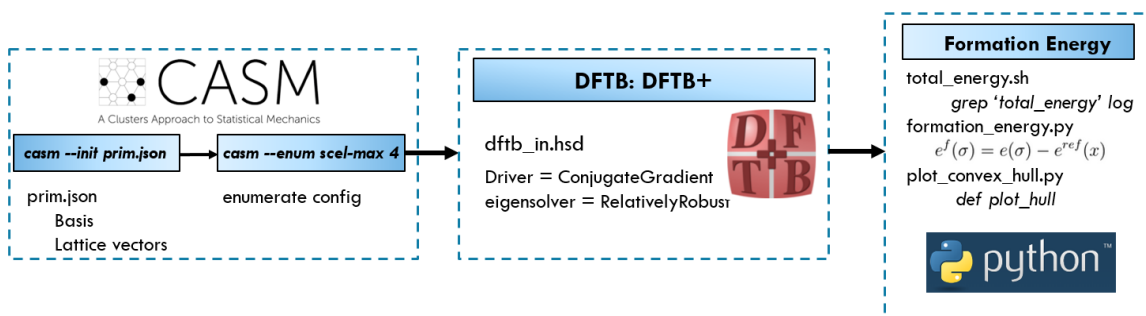


Figure 5.2: General workflow in our DFTB+CASM implementation for calculating formation energies and the convex hull.

After defining the basis and lattice vectors of the system, the symmetrically distinct configurations are generated using the CASM software package. All configurations in symmetrically distinct supercells were generated in 2H and B4 symmetries for SiC and ZnO, respectively. The primitive unit cell, as shown in Fig. 5.3, was used to generate the 401 derivative supercell configurations for each SiC and ZnO system. In the last panel of Fig. 5.2, we calculate the total energies for each configuration via DFT and DFTB.

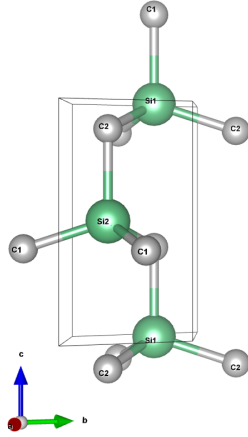


Figure 5.3(a) Lattice Vector \vec{a}

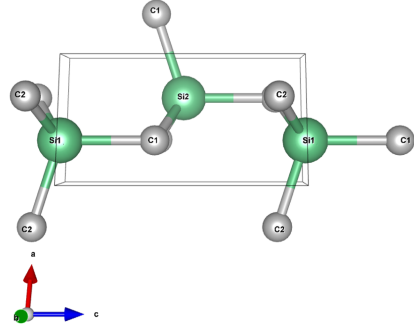


Figure 5.3(b) Lattice Vector \vec{b}

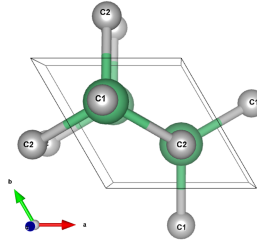


Figure 5.3(c) Lattice Vector \vec{c}

Figure 5.3: Primitive unit cells used to generate the various supercells and configurations. Using the CASM software package, we generate all symmetrically distinct configurations with the stoichiometry Si_2C_2 . Panels (a), (b), and (c) show the primitive cell in different orientations. Light green and grey atoms represent Si and C, respectively.

5.3.2 DFT Calculations

Density functional theory calculations were carried out with the Vienna Ab Initio Software Package (VASP) code [144, 145]. We used projector augmented wave (PAW) pseudopotentials [146, 147], and the generalized gradient approximation (GGA) exchange-correlation functional, as parameterized by Perdew, Burke, and Ernzerhof (PBE) [148]. We calculated the total energies using a plane-wave energy cutoff of 400 and 520 eV for SiC and ZnO, respectively. All of our calculations used the standard VASP pseudopotentials for all of the atoms. We used a Γ -point centered Monkhorst-Pack k-point mesh (approximately

$12 \times 12 \times 6$) for both SiC and ZnO, and the total energies were found to be suitably converged with this k-point sampling. The k-point sampling differs depending on the unit cell of each configuration, and the CASM software package keeps the mesh density constant for all configurations [149]. In all the optimizations, the geometry was relaxed such that all the forces were less than $0.04 \text{ eV}/\text{\AA}$. The energy convergence for the electronic degree of freedom was set to 10^{-5} eV . While performing geometry optimizations, the atomic positions, lattice parameters, and angles were allowed to relax for each structure.

5.3.3 DFTB Calculations

As discussed earlier, DFTB is an approximate tight-binding scheme with a low computational cost due to the use of parameterized integrals and a minimal valence basis set [137, 150–155]. In the present study, we used the self-consistent charge formulation of DFTB (SCC-DFTB) in its second-order scheme (DFTB2) [156], which includes the second-order term in the DFT energy expansion around the reference density [156]. Previous studies have shown DFTB to be particularly well suited for describing both SiC and ZnO materials [157–161]. In all of our SiC DFTB calculations, we used a recent DFTB parameterization that accurately reproduces a large dataset of DFT calculations, which includes potential energy surfaces, energies, and forces [162]. We designate the SiC SK files from Ref. [162] as SKfIV throughout this paper. For our ZnO DFTB calculations, we used standard parameters from the znorg-0-1 SK set [163, 164]. We used a similar k-point mesh as mentioned in the DFT calculations section. In all of our DFTB calculations, the geometry was relaxed with periodic boundary conditions such that all the forces were less than $0.04 \text{ eV}/\text{\AA}$, and the SCC convergence tolerance was set to 10^{-5} a.u. Both the DFT and DFTB calculations were

carried out in a spin-unpolarized formalism. After optimizing the structures, the formation energies and convex hull plots were produced using bash and python scripts.

5.3.4 Formation Energy

As depicted in the last panel of Fig. 5.2, the formation energy for each structure is calculated. The ‘total energy.sh’ script creates a file that contains the optimized DFT/DFTB ground state energy of each of the configurations. In the next step, the ‘formation_energy.py’ python script reads the ground-state energies and uses Eqs. 5.10 and 5.9 to compute the formation energy of each of the configurations. Finally, the ‘plot_convex_hull.py’ python script is used to plot the convex hull from the DFT/DFTB computed formation energies. In the case of SiC, we constructed the convex hull by connecting all the minima of the negative formation energies at various compositions, x , with straight lines.

5.4 Results and Discussion

Using the configuration space described above, we generated distinct compositions up to 4 times the primitive unit cell volume, which produces a total of 401 symmetrically distinct configurations for SiC and ZnO, each. The DFT and DFTB energies of all the enumerated configurations were calculated to obtain the formation energy using Eq. 5.9.

5.4.1 SiC

Convex Hull

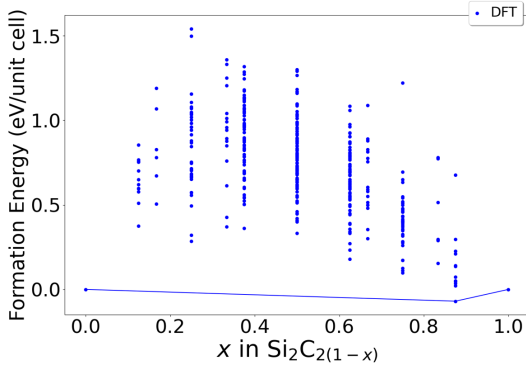


Figure 5.4(a) DFT

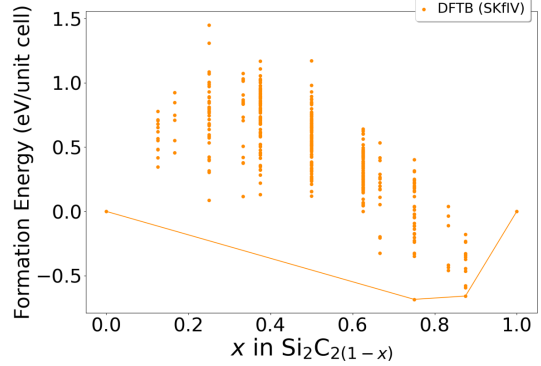


Figure 5.4(b) DFTB (SKfIV)

Figure 5.4: Formation energy convex hull for the Si-C binary system computed with DFT and DFTB. Panel (a) shows the DFT formation energy. Panel (b) shows the DFTB formation energy calculated using the SKfIV SK files. Each point corresponds to a different crystal structure.

Figs. 5.4(a) and (b) show the formation energies of various SiC configurations calculated with DFT and DFTB, respectively. Although the DFT and DFTB calculations find different minima on the convex hull plot, the results match qualitatively. DFTB predicts a minimum at composition $x = 0.75$ while DFT has a minimum nearby at $x = 0.875$. The two minima structures in the DFTB convex hull plot at $x = 0.75$ and $x = 0.875$ have a formation energy difference of 0.025 eV. This small energy difference is due to the approximations in the SK parameters inherent to the DFTB formalism.

The most probable reason for this small discrepancy in the formation energy is the parameterization of the repulsive DFTB potential between the Si and C atoms in the SKfIV SK files. The repulsive potentials in the SKfIV SK files result in a much stronger repulsion between the Si and C atoms, increasing the bond length between the atoms in

the structures, which alters their energetics.

A similar phenomenon was also observed for in a previous study using TiO_2 DFTB SK files [165]. Nevertheless, our results show that the DFT and DFTB formation energies for SiC are similar; if more accurate results are desired, configurations near the convex hull could be first down-selected via DFTB and subsequently refined/re-calculated with DFT to improve their accuracy (which would be more efficient than computing all 401 structures with DFT alone).

Structure Comparison

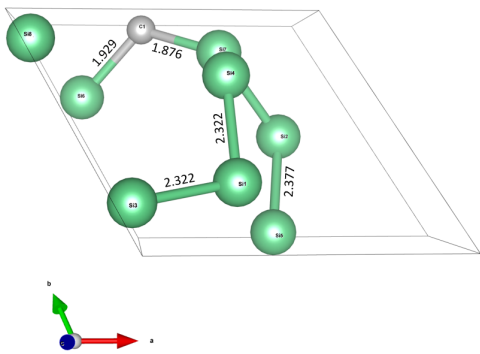


Figure 5.5(a) DFT

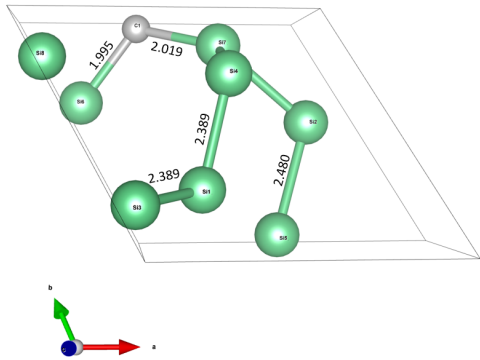


Figure 5.5(b) DFTB (SKfIV)

Figure 5.5: Comparison of SiC structural parameters (located on the convex hull) after optimization with (a) DFT and (b) DFTB (SKfIV), visualized along lattice vector \vec{c} . Bond lengths are in Angstroms, and light green and grey atoms represent Si and C, respectively.

The most stable crystal structure obtained via DFT and DFTB is depicted in Fig. 5.5. DFTB predicts a minimum at $x = 0.75$ while DFT gives a minimum nearby at $x = 0.875$. The two minima structures in the DFTB convex hull plot at $x = 0.75$ and $x = 0.875$ have a formation energy difference of 0.025 eV. As discussed in the previous section, this discrepancy in the minima for the DFTB formation energy results from approximations

in the repulsive potential between the Si and C atoms in the SKfIV SK files. Moreover, the crystal structures at $x = 0.875$ obtained via DFT and DFTB show the same P1 symmetry, which consists of eight Si atoms and one C atom in the unit cell (see Fig. 5.5). Since both the DFT and DFTB calculations correctly predict the same crystal structure and relative ratio of Si/C at $x = 0.875$, our results show that DFTB can be employed as an efficient computational approach tool for calculating and pre-screening formation energies.

Fig. 5.5 compares the DFT and DFTB optimized structural parameters of the most stable configurations at the convex hull minimum located at $x = 0.875$. Two types of Si atoms exist in the unit cell: one that is bonded to only Si atoms and another which is bonded with one C atom. Each Si and C atom has a coordination number of 4. As can be seen in Fig. 5.5(a), the Si-C bond length ranges from 1.93 to 2.38 Å in the DFT-optimized structure. Fig. 5.5(b) shows the bond lengths between various Si and C atoms of the DFTB-optimized structure. DFTB predicts slightly longer bond lengths for almost all the Si-C bonds, which on average are longer by 0.06 Å compared to DFT calculations. Table 5.2 compares the optimized lattice parameters of the minimum structure calculated with DFT and DFTB. DFTB overestimates the optimized lattice parameters and predicts slightly longer lengths for \vec{a} , \vec{b} , and \vec{c} .

As stated earlier, the discrepancy in the structural parameters is due to the parameterization of the repulsive DFTB potential between the Si and C atoms in the SKfIV SK files. The repulsive potentials used in the SKfIV SK files result in a much stronger repulsion between the Si and C atoms which increases the bond length between the atoms in the structures. Previous work has also shown that DFTB predicts longer lattice parameters

Table 5.2: Comparison of optimized lattice parameters of the minima structure calculated with DFT and DFTB for SiC.

Method	Lattice Vectors (\AA)			Lattice Angles		
	a	b	c	α	β	γ
DFT	6.15	6.15	6.16	121.39°	97.24°	110.62°
DFTB (SKfIV)	6.27	6.27	6.40	119.38°	99.60°	109.24°

for systems containing C atoms [166].

5.4.2 ZnO

Formation Energy

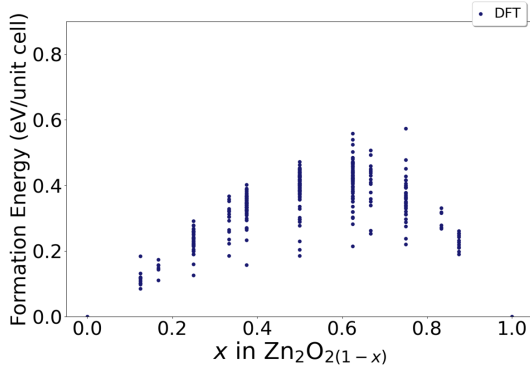


Figure 5.6(a) DFT

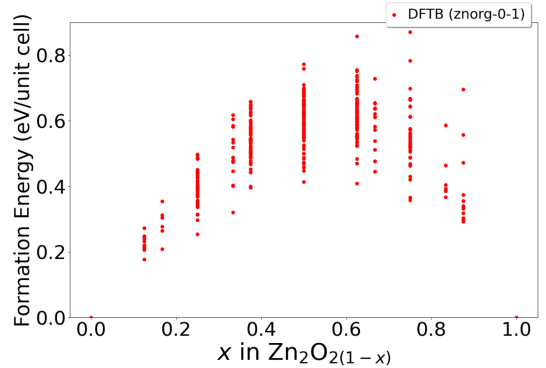


Figure 5.6(b) DFTB (znorg-0-1)

Figure 5.6: Formation energy of the ZnO binary system obtained from DFT and DFTB. Panel (a) shows the DFT formation energy, and panel (b) shows the DFTB formation energy calculated using the znorg-0-1 SK files. Each point corresponds to a different crystal structure.

We now proceed to ZnO, which is an even more complex material but shows more accurate results between DFT/DFTB. The wurtzite (B4) structure is the most stable form of ZnO at ambient conditions in nature [167]. The zinc blende (B3) structure of ZnO has a less stable cohesive energy than the B4 structure, and is, therefore, energetically unfavorable

at zero temperature and pressure [168]. As such, we enumerated various ZnO configurations starting with the B4-ZnO unit cell. We obtain a significant performance improvement for ZnO structures optimized via DFTB (more details on efficiency are discussed in the next section). It is interesting to note that both the DFT and DFTB calculations do not find a convex hull in the formation energy plots of the $\text{Zn}_2\text{O}_{2(1-x)}$ binary compounds. As shown in Fig. 5.6, the calculated formation energies are all positive, showing no stable minima structure predicted at any composition x . A previous study reported similar findings of positive formation energies when oxygen vacancies were introduced in the ZnO lattice [169]. Specifically, these previous studies showed that defects often induce occupied states in the bandgap and increase the formation energy [170–173].

5.4.3 Efficiency Analyses for DFT and DFTB

In this section, we give a detailed analysis of the computational timings and efficiency of DFT and DFTB. Fig. 5.7 compares the wall time per SCF iteration step (WT-SCF) as a function of the number of electrons in various ZnO configurations. As the number of electrons increases, the WT-SCF increases rapidly for both DFT and DFTB. For example, the DFT WT-SCF for a 30-electron ZnO configuration is around 16 seconds, whereas a 120-electron configuration is roughly 134 secs. Fitting the DFT WT-SCF data to a cubic polynomial gives a high R^2 correlation coefficient of 0.92, indicating an $O(N^3)$ scaling, where N is the number of electrons. This scaling can be attributed to matrix diagonalization in KS DFT, which is an $O(N^3)$ process, where N is the size of the matrix. Similarly, fitting the DFTB WT-SCF data gives an $O(N)$ linear scaling with an R^2 of 0.91. For all configurations, the DFTB WT-SCF is less than 6 secs, which is significantly faster than

DFT. As can be seen from Fig. 5.7, the DFTB WT-SCF for a ZnO configuration with 138 electrons is an order of magnitude faster than DFT. It is worth mentioning that although the DFTB WT-SCF is more efficient than DFT, DFTB geometry optimizations may take more SCF cycles compared to DFT, which may result in a small loss in efficiency.

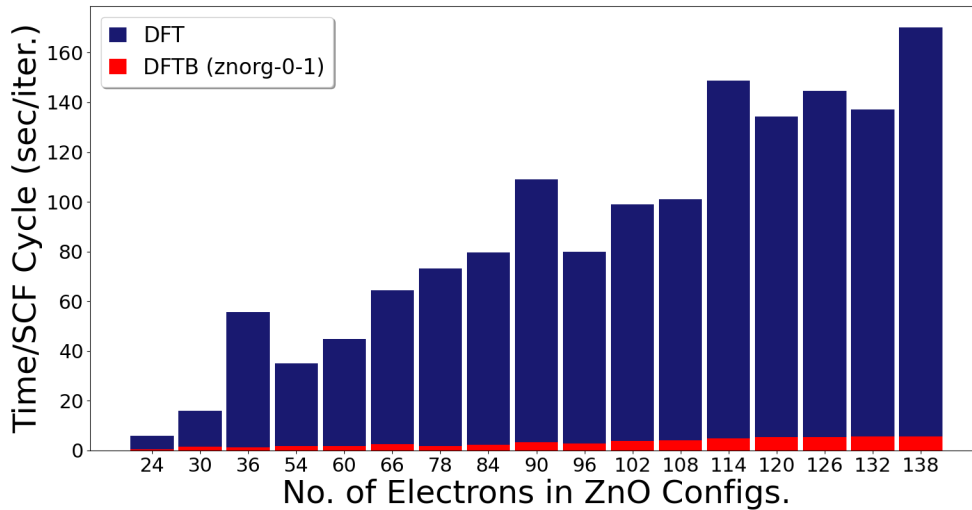


Figure 5.7: Comparison of DFT and DFTB wall times per SCF iteration step as a function of the number of electrons in various ZnO configurations. Each vertical bar value was calculated by averaging multiple configurations having the same number of electrons.

Computational Timings for SiC

Figs. 5.8(a) and (b) compare computational timings for geometry optimizations of various $\text{Si}_2\text{C}_{2(1-x)}$ structures using DFT (VASP) and DFTB (DFTB+). In general, the DFTB calculations take significantly less time compared to DFT, and Fig. 5.8(b) shows that DFTB can be an order of magnitude faster than DFT in some cases. For most configurations, the DFTB calculations required more SCF cycles for geometry optimization compared to DFT; however, the total compute time for DFTB is still significantly smaller.

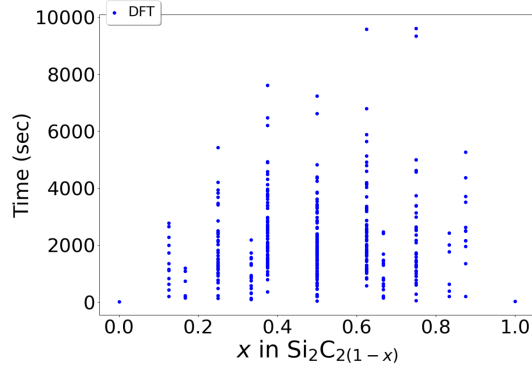


Figure 5.8(a) DFT

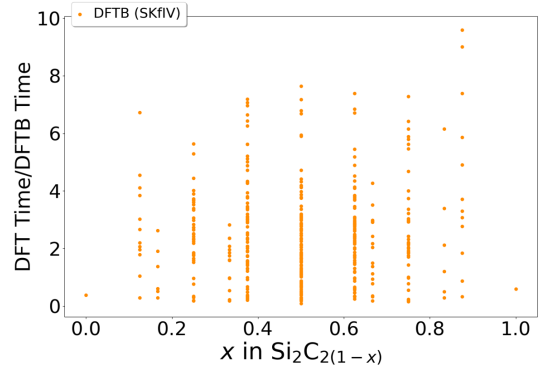


Figure 5.8(b) DFTB (SKfIV)

Figure 5.8: Comparison of wall times for geometry optimization of various $\text{Si}_2\text{C}_{2(1-x)}$ compositions calculated via DFT and DFTB. Panel (a) shows the wall time for optimizing each configuration using DFT. Panel (b) shows the ratio between the DFT and DFTB wall times for a geometry optimization of each configuration. Each point corresponds to a different crystal structure.

Computational Timings for ZnO

Fig. 5.9 compares wall times for geometry optimization for various $\text{Zn}_2\text{O}_{2(1-x)}$ binary compounds calculated via DFT and DFTB. As in the case of SiC, Fig. 5.9(b) shows that DFTB is an order of magnitude faster than DFT in most cases. It is important to note that DFTB is almost 40 times faster than DFT for $\text{Zn}_2\text{O}_{2(1-x)}$ binary compounds (see Fig. 5.9(b)), whereas the maximum performance enhancement of DFTB is only 10 times for $\text{Si}_2\text{C}_{2(1-x)}$. In general, as the number of electrons in the system increases, we show that DFTB exhibits more performance gains than conventional KS DFT calculations.

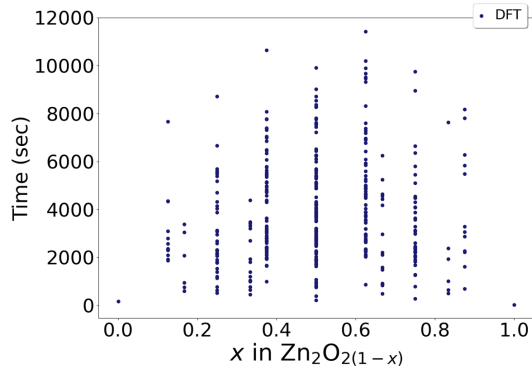


Figure 5.9(a) DFT

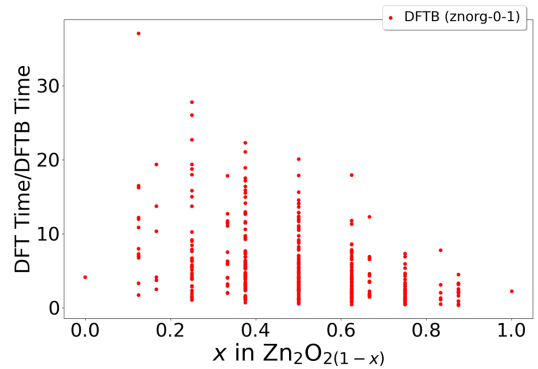


Figure 5.9(b) DFTB (znorg-0-1)

Figure 5.9: Comparison of wall times for geometry optimization of various $\text{Zn}_2\text{O}_2(1-x)$ compositions calculated via DFT and DFTB. Panel (a) shows the wall time for optimizing each configuration using DFT. Panel (b) shows the ratio between the DFT and DFTB wall times for a geometry optimization of each configuration. Each point corresponds to a different crystal structure.

Finally, Fig. 5.10 compares the total time (sum of individual wall times) for the geometry optimization of $\text{Si}_2\text{C}_2(1-x)$ and $\text{Zn}_2\text{O}_2(1-x)$ binary compounds. For $\text{Si}_2\text{C}_2(1-x)$, the total time for geometry optimization of all 401 configurations is around five days for DFTB, whereas the DFT calculations require nearly ten days. The difference in performance between DFTB and DFT for $\text{Zn}_2\text{O}_2(1-x)$ is even more significant - DFTB geometry optimizations take approximately five days, whereas the DFT calculations take nearly 18 days (more than 3 times longer than DFTB). As mentioned previously, larger systems such as ZnO (which contain more electrons), scale more favorably with DFTB, resulting in speedups that can be orders of magnitude faster than conventional DFT methods.

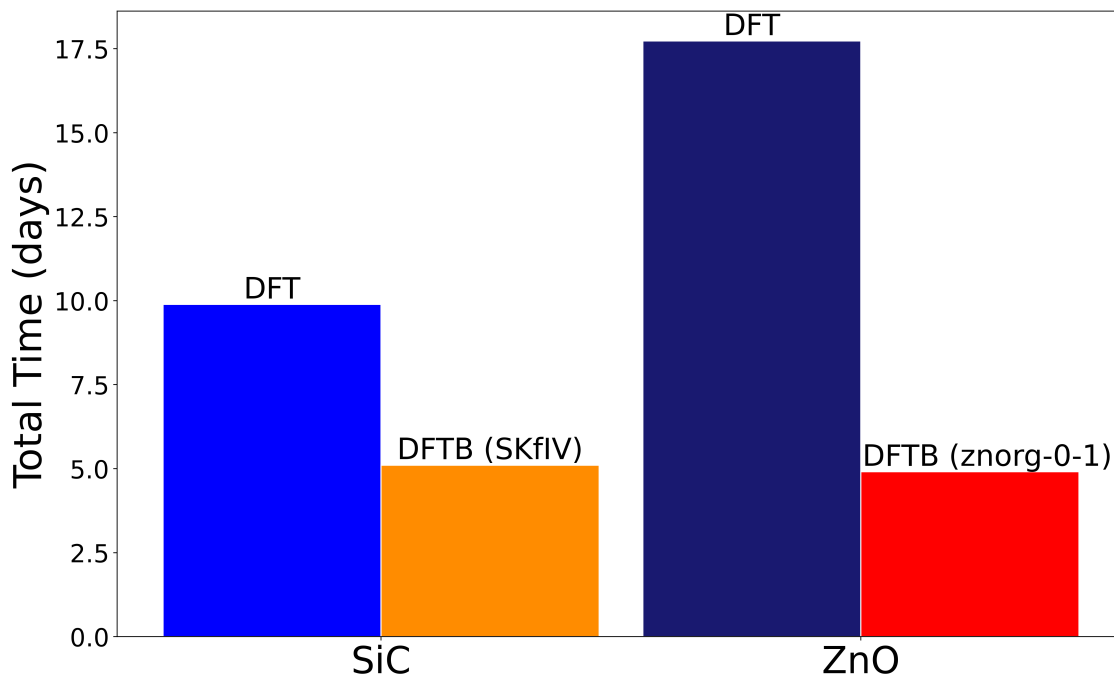


Figure 5.10: Comparison of DFT and DFTB total wall times for geometry optimization of all SiC and ZnO configurations. The DFTB calculations used the SKfIV and znorg-0-1 SK files for SiC and ZnO, respectively.

5.5 Conclusion

In closing, we have interfaced DFTB with the CASM software package for the first time to enable routine and efficient calculations of formation energies and convex hulls. Our extensive calculations show that DFTB can be used as an efficient screening tool to compute the numerous formation energies (and convex hull if it exists) of complex materials. To highlight the efficiency and accuracy of our approach, we calculated and compared the formation energies of SiC and ZnO with both DFT and DFTB. We find that the DFTB approach enables extremely efficient calculations of formation energies in a completely unbiased manner to predict low-lying metastable phases over the entire composition space. By

comparing the convex hull/formation energy from both approaches, we found that DFTB gives similar trends as the DFT calculations. Finally, we performed an extensive benchmark of the computational timings for both DFT and DFTB and found that the DFTB calculations can be an order of magnitude faster (larger systems give even higher computational efficiency). Moreover, we show that DFTB gives accurate results and can be a computationally cheaper alternative to DFT. In summary, our DFTB+CASM implementation allows for an efficient exploration (up to an order of magnitude faster than DFT) of formation energies and convex hulls, which researchers can use to routinely probe other complex systems.

Chapter 6

Acceleration vs Accuracy:

Influence of Basis Set Quality on Dynamics Predicted by Ab Initio MD

This chapter explores the quality of the employed basis set when using ab initio molecular dynamics to study collisions between the reactants of a water–gas shift reaction with a ZnO surface. The majority of the work in this chapter resulted from a collaboration with Dr. Sharma S.R.K.C. Yamijala and was published as an article in *the Journal of Physical Chemistry C* [174].

6.1 Introduction

Energy channeling via “hot” electrons and thermally excited atoms/molecules toward a specific reaction channel has recently emerged as a new area in the catalysis community for controlling and understanding the mechanism and dynamics of catalytic reactions [175–185]. Predicting the reaction mechanisms of these external-energy-mediated reactions using traditional quantum chemical approaches, where the energies of the reactants, products, and transition states are used to predict the reaction mechanisms, is nontrivial [186–189]. This complexity arises due to various reasons, such as difficulties pertaining to the determination of transition states, requirements in considering numerous configurations for the adsorbate, the formation of complex transition states involving more than two molecules, etc [186, 187]. Apart from this complexity, typical computational results are obtained only at zero Kelvin, and extrapolations of these results to realistic reaction temperatures are questionable. An alternative, automatic, and unbiased approach to predict the mechanism of a reaction is to use ab initio molecular dynamics (AIMD). By its nature, an AIMD simulation automatically includes the dynamic and steric effects of a reaction with a simultaneous prediction of possible reaction mechanisms [186, 187].

Born Oppenheimer molecular dynamics (BOMD), one of the various AIMD methods, has been quite successful in predicting the mechanisms and possible outcomes of reactions involving atomic or molecular collisions [186, 187, 190–199]. For example, Schatz and co-workers have recently used BOMD simulations to understand the detailed mechanisms of industrially relevant reactions. These include the reverse water–gas shift reaction and Fischer–Tropsch synthesis on nickel surfaces, [192, 194] where the precoated Ni surface was

bombarded with energetic hydrogen atoms and methylene, respectively. Wang et al. [191] have used BOMD to probe the water dissociation equilibrium on TiO_2 by colliding energetic water molecules toward the TiO_2 surface. BOMD has also been employed in predicting the reaction outcomes of atmospherically relevant molecular collisions. Examples of these include the formation of carbonic acid obtained by the collision between a CO_2 molecule and H_2O clusters [199] and the formation of sulfuric acid [186, 187] via the oxidation of SO_2 to SO_3 , resulting from a collision between SO_2 and $\text{O}_3^- (\text{H}_2\text{O})_n$ clusters.

Considering the previously mentioned successes of BOMD with collision, reactions, it is both essential and necessary to study the specific computational parameters that can alter the mechanism and dynamics predicted by this method. Both the mechanism and dynamics of a chemical reaction are dictated by the forces acting on the nuclei. In BOMD, these forces are obtained in an ab initio manner from the ground-state potential, which, in turn, is obtained with electronic structure methods such as Hartree–Fock, DFT, DFTB, etc. Furthermore, as density functional theory (DFT) is the most widely used approach for calculating the electronic structure during a BOMD simulation [200–202] and since DFT results are sensitive to various parameters such as basis sets and exchange–correlation functionals, the resulting BOMD dynamics is also expected to be affected by these factors. However, the effect of these parameters on the predicted results of a BOMD simulation has not been well studied, and in the present work, we focus our attention on one of these parameters, namely, the importance of the basis set.

It is worth noting that although there are numerous studies on the effect of basis sets on ground-state (i.e., stationary) structures [203–206], there are only a handful of

studies on basis set effects in BOMD simulations [207–213]. Furthermore, most of these studies were devoted to understanding basis set effects in clusters. Thus, to the best of our knowledge, an in-depth study of basis set effects in large-scale Born–Oppenheimer molecular dynamics (BOMD) calculations (with periodic boundary conditions) is less common. Such investigations have not been widely carried out since the conventional assumption in these studies is that double-zeta (DZ) quality basis sets are assumed to work well [186, 191, 197–199], and comparing the BOMD calculations with larger basis sets is quite computationally demanding. Although the above assumption may seem reasonable, we demonstrate in this work that calculations using the widely used double zeta basis can yield completely different results than those obtained with a larger basis (for example, a triple-zeta (TZ) quality basis). As shown in our calculations, the reaction mechanisms obtained from BOMD simulations using DZ basis sets may need to be re-evaluated. For example, the dissociation equilibrium of water predicted by Wang et al. [191] depends on whether the water molecule colliding with the TiO_2 dissociates on the surface or whether it scatters from the surface. However, as shown in this work, the fate of a colliding molecule highly depends on the employed basis set, which can qualitatively change the entire prediction.

In the present work, we have studied the collision of a water (H_2O) molecule with a ZnO surface precovered with carbon monoxide (CO) with both double- and triple-zeta quality basis sets using BOMD simulations. To gain further insight, we have also studied the interaction of CO, CO_2 , H_2 , and H_2O with the ZnO surface with both of these basis sets. Based on the results from both BOMD and DFT, we propose that single-point binding energy calculations can be used as a simple diagnostic tool to estimate the impact of basis-

set quality before carrying out a computationally expensive BOMD simulation. Finally, we note that our simulations are relevant to the industrially relevant water–gas shift reaction, where ZnO is one of the co-catalysts, and all of the molecules mentioned above are either products or reactants of this reaction [214–218].

6.2 Computational Details

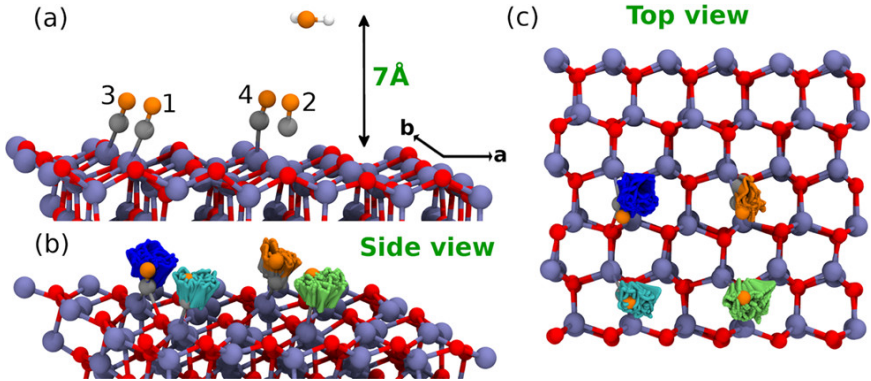


Figure 6.1: (a) Initial configuration used to study the interaction of an H₂O molecule colliding with a ZnO slab preadsorbed with four CO molecules. The four CO molecules are numbered to distinguish the 4th CO molecule (closest one to the impact region) from the others. The blue, red, gray, and white colors are used to represent the Zn, O, C, and H atoms, respectively. Oxygen atoms of the CO and H₂O are colored in orange to distinguish them from the oxygen atoms of the ZnO lattice. Panels (b) and (c) show the side and top views of how the CO molecules in the configuration (a) have changed during a 4.4 ps NVT run performed with the TZV2P basis sets at 300 K. In both (b) and (c), the cyan, lime, and blue trajectories show the change in the positions of the 1st, 2nd, and 3rd CO molecules, respectively, for the entire trajectory.

All calculations were performed with Kohn–Sham density functional theory using the PBE ([219]) exchange–correlation functional as implemented in the CP2K, ([200]) FHI-aims, ([201]) and VASP ([202]) software packages. In CP2K, we have used both the molecularly optimized double-zeta quality (DZVP) and triple-zeta quality (TZV2P) basis-sets, ([220]) which are compatible with the employed Goedecker–Teter–Hutter (GTH) pseudopo-

tentials [221, 222]. For the auxiliary plane-wave (PW) basis, used in the Gaussian-and-Plane-Waves method of CP2K, we have used 1000 Ry for the PW energy cutoff and 60 Ry for the reference grid cutoff. In FHI-aims, we have used both the Tier-1 and Tier-2 numerical orbital basis-sets. In VASP, we have used the default plane-wave basis with a 500 eV energy cutoff. It is important to point out that the aim of this study is not to judge the accuracy of the exchange–correlation functional; instead, we highlight the role of the basis set in predicting the catalytic dynamics (i.e., mechanistic outcomes) of a reaction. As such, we have only considered one of the most widely used exchange–correlation functionals, namely, PBE (although we anticipate that the dramatic basis set effects shown in this work will also apply to other functionals).

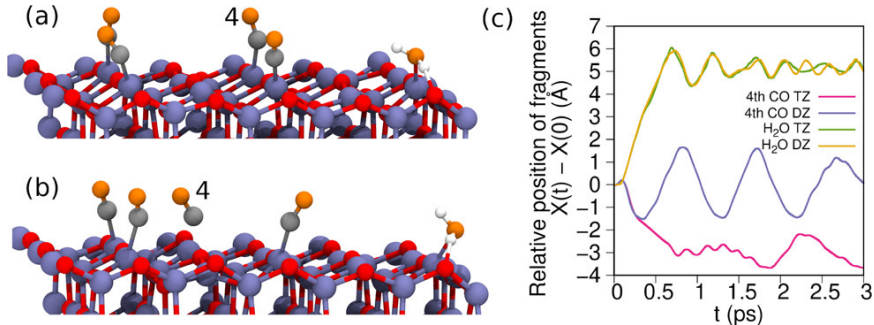


Figure 6.2: Panels (a) and (b) show how the configuration in Figure 6.1a has changed for a H_2O molecule with 0.6 eV of incident energy after a 3 ps NVE simulation performed with the DZVP and TZV2P basis sets, respectively. (c) Relative change in the positions of the 4th CO and H_2O molecules along the a-direction of the lattice.

For all of the electronic structure calculations, a 4x4 slab with a thickness of 4 layers was used. A vacuum of at least 15 Å was used to avoid any spurious interactions between the periodic images. Since the choice of a k-mesh did not affect our findings (see Appendix C), we only present calculations performed at the Γ -point of the Brillouin zone. Furthermore, since recent experiments have shown that CO molecules can easily hop along

the a-direction of the ZnO slab (due to the relatively smaller lattice spacing along that direction), ([223]) in all of our BOMD simulations, we have expanded our slab along the a-direction to capture this hopping mechanism. Thus, we performed all of our NVE and NVT calculations with a 4-layer-thick 6x3 slab (instead of a 4x4-slab). We have integrated the equations of motion with a 0.5 fs time step. The initial velocities and coordinates for the NVE runs were obtained by running an NVT simulation at 300 K. For the NVT simulations, we have used the Nosé–Hoover thermostat of the chain length three. In all the NVT and NVE runs, Grimme’s D3 dispersion correction was employed [224]. Various other computational details are given in Appendix C.

6.3 Results and Discussion

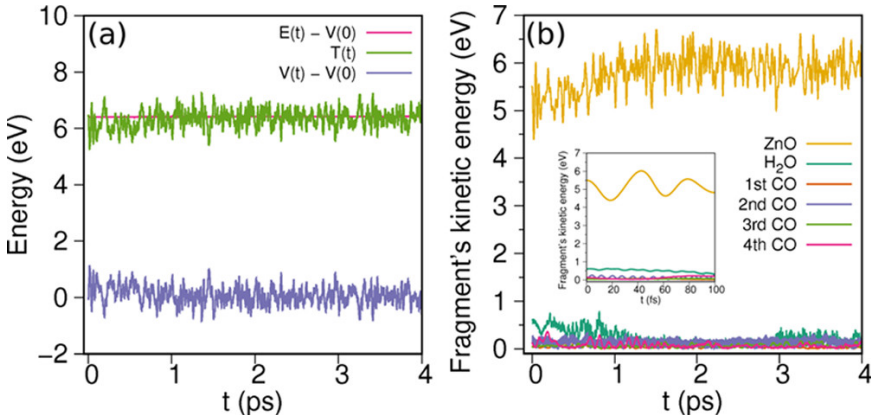


Figure 6.3: Panel (a) shows the changes in total (E), kinetic (T), and potential (V) energies of the entire system for the 0.6 eV collision simulated with the TZ basis. Panel (b) shows the changes in the kinetic energy of each of the fragments during the same NVE run. The inset in (b) shows the changes in kinetic energies of the fragments during the first 100 fs of the simulation.

Figure 6.1a shows the initial configuration used to study the interaction of an H₂O molecule colliding with a ZnO slab preadsorbed with CO molecules at 0.22 ML coverage.

For simplicity, we have only considered the normal incidence of H₂O (90° to the slab). We have studied the impact of H₂O with two incident energies, namely, 0.6 and 6 eV. Here, the lower incident energy of 0.6 eV was chosen since it is slightly above the binding energy of the CO molecule on a ZnO surface (0.26–0.53 eV) obtained at the PBE level of theory [223, 225]. In other words, it is the minimum amount of energy required for the CO molecule to hop or diffuse on the ZnO surface. To represent the high-energy collision, we used a 6 eV incidence, which is an order of magnitude larger than the binding energy of the adsorbate.

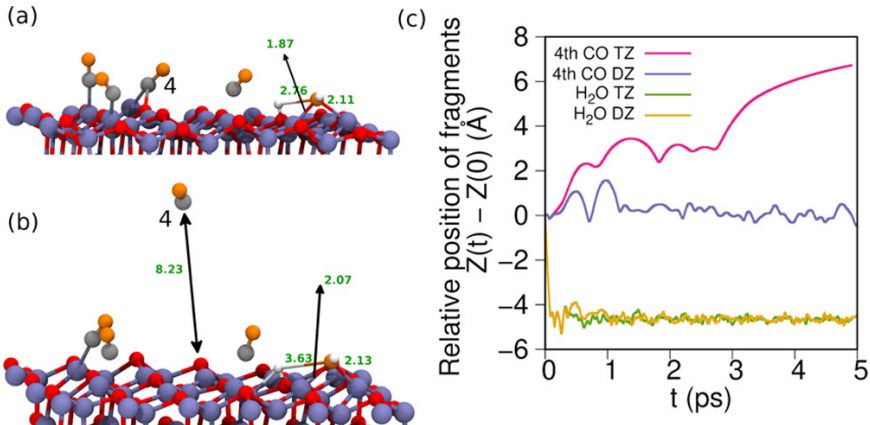


Figure 6.4: Panels (a) and (b) show how the configuration in Figure 6.1a has changed for the H₂O molecule with 6 eV of incident energy, after a 5 ps NVE simulation performed with the DZVP and TZV2P basis sets, respectively. (c) Relative change in the positions of the 4th CO and H₂O molecules along the *c*-direction of the lattice.

Before proceeding to the collision reaction results, we would first like to show how the adsorbed CO molecules evolve during an NVT simulation (performed at 300 K and with the TZV2P basis) in the absence of any collision. In Figure 6.1b, we have shown the positions of each of the CO molecules for every 10 fs during the entire trajectory. Clearly, throughout the 4.4 ps simulation time, all of the CO molecules remained in their binding sites and fluctuated around their mean positions. Apart from these fluctuations, we did not

find any significant changes (such as a hop, diffusion, or desorption of any CO) during the simulation. Similar results were also obtained with the DZVP basis (see Figure 6.8). Next, we present the effect of the basis set quality on the collision reactions.

6.3.1 Low-Energy (0.6 eV) Collision

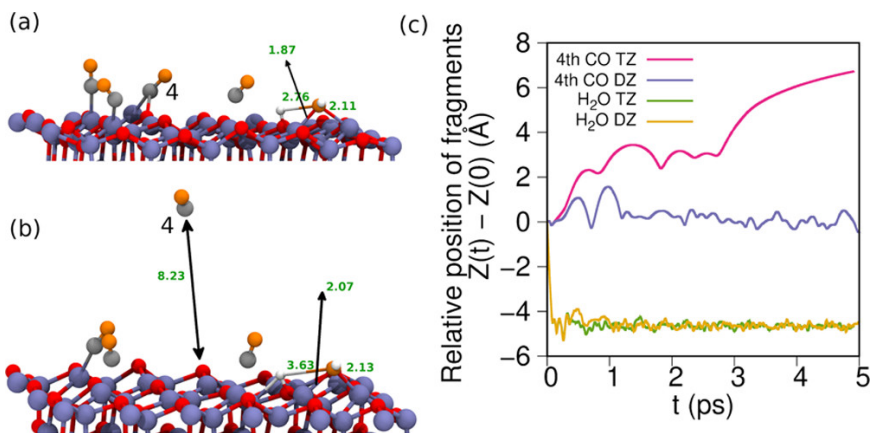


Figure 6.5: Panel (a) shows the changes in total (E), kinetic (T), and potential (V) energies of the entire system for the 6 eV collision simulated with the TZ basis. Panel (b) shows the changes in the kinetic energy of each of the fragments during the same NVE run. The inset in (b) shows the changes in kinetic energies of the fragments during the first 100 fs of the simulation.

We now present our NVE simulation results of a water molecule colliding with the ZnO surface with a 0.6 eV translational energy. In Figure 6.2 (a) and (b), we show the final configuration obtained after a 3 ps simulation performed with the double zeta (DZVP) and triple-zeta (TZV2P) basis, respectively. The relative change in the positions of the 4th CO and H₂O, along the a-direction of the lattice, for the entire trajectory is shown in Figure 6.2c. The changes depicted here are relative to the initial frame (shown in Figure 6.1a). Here, we would like to note that there is no head-on collision of the H₂O with the 4th CO. However, as it is the closest one to the impact area, it has the maximum effect.

Also, hereafter, we denote the TZV2P basis as TZ and the DZVP basis is denoted as DZ.

As shown in Figures 6.2a and 6.9a, with the DZ basis, none of the CO molecules has left its binding sites even after 3 ps of simulation time. Furthermore, during the entire simulation, their relative positions (only shown here for the 4th CO) were merely oscillating around zero (Figure 6.2c, blue curve). This oscillatory behavior around zero corresponds to the thermal fluctuations around their initial positions. Thus, the low-energy collision did not result in any significant changes in the behavior of the CO molecules on the ZnO surface. On the other hand, with the TZ basis, we observed a hop of the 4th CO from one Zn site to the other one (see Figures 6.2b and 6.9b as compared with Figure 6.1a). As shown in Figure 6.2c, this hop actually occurred within the first picosecond of the simulation time and is characterized by a decrement in the relative position of the 4th CO by ~ 3 units in the negative direction. Here, $\sim 3 \text{ \AA}$ is the distance between two Zn sites along the a -direction (3.28 \AA for the ground-state structure), and the negative change corresponds to the hop along the negative a -direction of the lattice. After this hop, the 4th CO continued to remain at the new Zn site while exhibiting thermal fluctuations (Figures 6.2c and 6.9b). Finally, the positive change in the position of the H_2O with both basis sets corresponds to hopping along the positive a -direction of the lattice. This change in the position of the H_2O and its adsorption onto the ZnO surface (after the collision) can be seen in both Figure 6.2a,b.

Next, we discuss the changes in the energetics of this 0.6 eV collision reaction. In Figure 6.3a, we have given the total (E), kinetic (T or KE), and potential (V or PE) energies of the entire system during the NVE run with the TZ basis. Clearly, the total energy of the entire system is conserved, demonstrating the NVE nature of this simulation.

Also, due to the low-impact energy, there are only minor changes to the total KE or PE of the system. As shown in Figure 6.3b, the incident H₂O molecule transferred some (~ 0.3 eV) of its KE toward the 4th CO molecule during the initial stages of the run (< 1 ps), resulting in the hopping of the 4th CO from one Zn site to the other (as discussed earlier). At a later time, the KE of both the H₂O and the 4th CO has been completely transferred to the ZnO slab. This energy transfer between fragments can be clearly observed in Figure 6.3b as a rise in the KE of the ZnO and the decay of the H₂O and 4th CO molecule's KE after ~ 2 ps simulation time. During the rest of the simulation time, all of the molecules (four COs and one H₂O) remained vibrating around their mean positions. Thus, for 0.6 eV of collision energy, during the entire run, we only observed hopping of the 4th CO and the adsorption of H₂O to the ZnO surface. The energetics of the simulation with the DZ basis is given in Figure 6.10.

6.3.2 High-Energy (6 eV) Collision

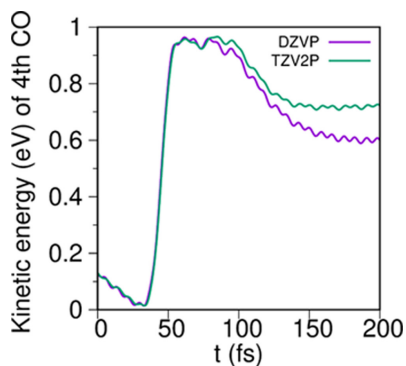


Figure 6.6: Similar results in the rise of the KE of the 4th CO using both the DZ and TZ basis during the first 100 fs of the simulation.

Next, we present the results of the high-energy collision (6 eV). In Figures 6.4 (a) and (b), we show the final configuration obtained after a 5 ps NVE simulation performed with the DZ and TZ basis, respectively. First, unlike the low-energy collision, we observed the dissociation of the colliding H₂O molecule into OH and H species with both basis sets. Also, we find that the dissociated OH and H species were within the bonding region of the Zn and O sites of the ZnO lattice, respectively. Unlike the case of H₂O, we obtained significantly different results for the 4th CO with the DZ and TZ basis. After a 5 ps simulation with the TZ basis, we find the complete desorption of the 4th CO (Figure 6.4b); however, with the DZ basis, we only find the diffusion of the 4th CO (Figure 6.4a).

In Figure 6.4c, we present the relative change in the positions of the 4th CO and H₂O along the *c*-direction of the lattice for the entire trajectory. Once again, the depicted changes are relative to the first frame (shown in Figure 6.1a). Here, for the H₂O, a relative position of zero Angstroms corresponds to its initial state of ~ 6.5 Å above the surface, and a relative position of -5 Å corresponds to its contact with the ZnO surface. Thus, as shown in Figure 6.4c, the H₂O molecule made contact with the ZnO surface within ~ 80 fs for both basis sets and remained on the surface until the end of the simulation. On the other hand, the 4th CO showed qualitatively different results with both basis sets. In this figure, for the 4th CO, a relative position of zero Angstroms corresponds to its initial adsorption configuration on the ZnO surface, and a relative position of greater than 6 Å corresponds to its complete desorption from the ZnO surface. As such, with the TZ basis, we observed an initial dissociation (>1 Å) of the 4th CO from its Zn binding site in less than a picosecond (Figure 6.4c). This is followed by its diffusion on the ZnO surface (~ 2 - 3 Å) until ~ 3 ps and

finally a complete desorption from the ZnO surface at ~ 4 ps of the run. On the other hand, with the DZ basis, the CO always remained below 2 \AA from the ZnO surface, suggesting a diffusion-like behavior. As noted earlier, for the TZ basis calculation, the CO molecule was also diffusing during the first 3 ps simulation time, albeit at a larger distance ($>2 \text{ \AA}$) from the ZnO surface compared to the CO when the DZ basis was used. A visual representation of these changes for the entire trajectory is given in Figure 6.11.

Table 6.1: Adsorption Energies of the Molecules Calculated Using CP2K with Different Basis Sets and Differences in the Adsorption Energies between Basis Sets.

System	Adsorption Energy with the DZVP basis, E_{DZVP} (eV)	Adsorption Energy with the TZV2P basis, E_{TZV2P} (eV)	Difference in Adsorption Energy, $E_{\text{TZV2P}} - E_{\text{DZVP}}$ (eV)
ZnO slab + CO	-0.884	-0.504	0.380
ZnO slab + CO ₂	-0.177	-0.175	0.002
ZnO slab + H ₂	-0.088	-0.083	0.005
ZnO slab + H ₂ O	-0.990	-0.961	0.029

Next, in Figures 6.5 (a) and (b), we present the energy profiles of the 6 eV collision with the TZ basis. Once again, the total energy of the entire system is conserved. However, unlike the low-energy collision, the changes in the KE or PE of the system are quite apparent for the 6 eV collision. Here, we find an ultrafast (<200 fs) decay (rise) in the total kinetic energy (potential energy) of the system, as shown in Figure 6.5a. To understand this ultrafast decay in the total KE, we have plotted the changes in the KE of the individual fragments in Figure 6.5b. Clearly, the ultrafast decay of the total KE can be directly correlated with the ultrafast decay in the H₂O molecule’s KE (inset of Figure 6.5b). However, although the H₂O molecule has lost most of its KE (~ 5 eV) in less than 100 fs, only a part of it (~ 3 eV) has been transferred to the other fragments (~ 1 eV to the

4th CO and ~ 2 eV to the ZnO lattice) in the form of KE (see Figure 6.5b). The rest is transformed into the PE of the system, leading to its increase. The reason for such a rise in the PE of the system is due to the formation and dissociation of various bonds during the first 200 fs of the run. At ~ 80 fs, the H₂O molecule made contact with the ZnO surface (see Figures 6.4c and 6.12), and, subsequently, the dissociation process of water was initiated. Following the water dissociation (into OH and H), we observed the formation of Zn–OH and O–H bonds with the Zn and O atoms of the slab, respectively (see Figure 6.12). Also, during this time, due to the increased KE of the 4th CO, the Zn–CO bond subsequently dissociated (see Figures 6.4c and 6.12). Thus, the ultrafast decay in the KE profile is due to a combination of many factors as explained above. We have given the energetics of the same 6 eV collision with the DZ basis in Figure 6.13.

It is important to emphasize here that the simulations with both the TZ and DZ basis behaved almost similarly during the 0–200 fs simulation time. This similarity is apparent both in their energetics and dynamics (compare Figures 6.5 and 6.13; also see Figure 6.12). The differences between the TZ and DZ basis start after ~ 200 fs, and, as noted earlier, the differences are quite apparent in the dynamics of the 4th CO. However, even for the 4th CO, the energetics and dynamics are quite similar to the DZ and TZ basis during the first few hundred femtoseconds, as shown in Figures 6.6 and 6.4c. Clearly, both the KE gain by the 4th CO (Figure 6.6) and its relative position (Figure 6.4c) are almost the same, until ~ 200 fs for both the DZ and TZ runs (also see Figure 6.12). Thus, the differences observed in the dynamics of the 4th CO are not due to any changes in its KE gain (from H₂O) but only due to the difference in the interactions among various fragments

of the system introduced by the employed basis set. In the next section, we study how the interaction between the adsorbate molecule and the ZnO surface vary with a change in the basis set.

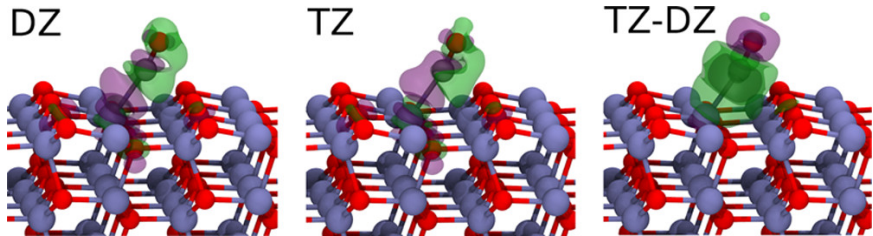


Figure 6.7: The first two columns show electron density differences obtained for a single CO molecule adsorbed on a ZnO slab with the DZVP (DZ) and TZV2P (TZ) basis sets, respectively. Here, each of these columns is obtained by subtracting the total electron density of the entire system from its individual components (i.e., $\rho[\text{diff}] = \rho[\text{ZnO} + \text{mol}] - \rho[\text{ZnO}] - \rho[\text{mol}]$). The last column shows the difference in the densities between the first two columns (i.e., $\rho^{\text{basis}}[\text{diff}] = \rho^{\text{TZ}}[\text{diff}] - \rho^{\text{DZ}}[\text{diff}]$). We have used an isovalue of $0.001 \text{ e}\text{\AA}^{-3}$ for the first two columns and $0.0003 \text{ e}\text{\AA}^{-3}$ for the last column.

Until now, we have found that irrespective of the incident energy of the H_2O molecule, we have always obtained different results using the DZ and TZ basis sets for the same initial configuration and for the same initial rise (loss) in the KE energy by the 4th CO (H_2O). Although the differences are just noticeable in the case of the 0.6 eV collision, they are drastically different for the case of the high-energy collision. For the 6 eV collision with the DZ basis, we have only observed the diffusion of the 4th CO on the ZnO surface, whereas we observed a qualitatively different mechanism for the TZ basis, namely a complete desorption (apart from the initial diffusion). Furthermore, for the 6 eV collision, the differences in the reaction outcomes between the DZ and TZ basis persisted with a change in the impact position (verified by impacting the H_2O molecule in the vicinity of the 2nd CO molecule instead of the 4th CO molecule) even when dispersion corrections were included/neglected. These additional results are presented in Appendix C. This significant

difference in the collision reaction outcomes obtained by merely changing the basis is one of the central findings of our work.

6.3.3 Basis Set Effect on the Adsorption Energies

Table 6.2: Adsorption Energies of the Molecules Calculated Using FHI-aims with Two Different Basis Sets and Differences in the Adsorption Energies between Basis Sets.

System	Adsorption Energy with the Tier-1 basis, E_{Tier1} (eV)	Adsorption Energy with the Tier-2 basis, E_{Tier2} (eV)	Difference in Adsorption Energy, $E_{\text{Tier2}} - E_{\text{Tier1}}$ (eV)
ZnO slab + CO	-0.489	-0.398	0.091
ZnO slab + CO ₂	-0.196	-0.172	0.023
ZnO slab + H ₂	-0.095	-0.096	-0.001
ZnO slab + H ₂ O	-1.055	-0.999	0.056

To further understand the reason for this difference, we have performed a few adsorption energy calculations with different adsorbates on the ZnO surface. In all of these calculations, we have calculated the adsorption energy (E_{ads}) of a molecule on a ZnO slab as

$$E_{\text{ads}} = E_{(\text{slab}+\text{mol})} - E_{\text{slab}} - E_{\text{mol}} \quad (6.1)$$

where $E_{(\text{slab}+\text{mol})}$ is the energy of the entire system (slab + adsorbate molecule), E_{slab} is the energy of the ZnO slab, and E_{mol} is the energy of the adsorbate molecule.

In Table 6.1, we have given the adsorption energies of all four molecules calculated using CP2K with both the DZ and TZ basis. We have also given the difference in the adsorption energy ($E_{\text{diff}}^{\text{TD}}$) between the TZ and DZ basis. Except for CO, the adsorption energy of all other molecules is not sensitive to basis set effects (differences are on the order of room temperature). However, surprisingly for CO, the change in the adsorption

energy with a change in the basis is more than an order of magnitude (~ 0.4 eV) larger than room temperature (~ 0.026 eV). This huge difference in the CO adsorption energies with a change in the basis is the primary reason for the completely different results observed in the collision calculations (as shown in the earlier sections) despite their similarities during the initial stages of the run. From the adsorption energies, it is clear that CO is weakly bound to the ZnO substrate (by ~ 0.4 eV) when using the TZ basis. We note that this weakness in the interaction between ZnO and CO with the TZ basis manifests itself in both the facile dissociation and hopping of the 4th CO for the 0.6 eV collision and the complete desorption of the 4th CO for the 6 eV collision. Neither the hop nor the desorption of the CO was observed while using the DZ basis because of the stronger interaction between CO and ZnO with this basis set. This difference in the interaction between CO and ZnO with a change in the basis can also be noticed in the shorter (longer) distance between CO and ZnO during the diffusion of CO on the surface with the DZ (TZ) basis.

6.3.4 Electron Density Difference Maps

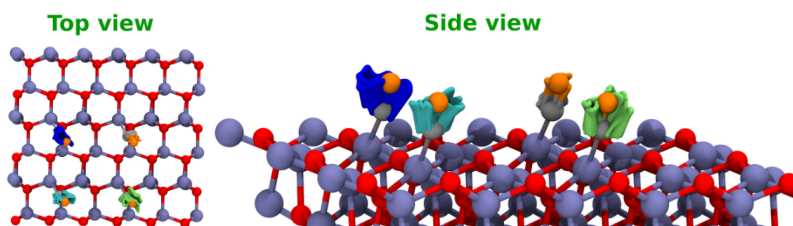


Figure 6.8: Both panels show how the CO molecules evolve during the 3.5 ps NVT simulation with DZVP basis sets at 300 K (in the absence of any collision). In both panels, the cyan, lime, and blue trajectories show the change in the positions of the 1st, 2nd, and 3rd CO molecules, respectively, for the entire trajectory (3.5 ps). The changes in the 4th CO position are shown with the stick model without changing its colors.

To gain further insight into the influence of the basis set on the interaction between the slab and the adsorbate molecule, we have plotted the electron density differences, as shown in Figure 6.7. First, for each basis set, we have computed the electron density difference ($\rho[\text{diff}]$) as

$$\rho[\text{diff}] = \rho[\text{ZnO} + \text{mol}] - \rho[\text{ZnO}] - \rho[\text{mol}] \quad (6.2)$$

where $\rho[\text{ZnO} + \text{mol}]$, $\rho[\text{ZnO}]$, and $\rho[\text{mol}]$ are the electron densities of the entire system (slab + adsorbate molecule), ZnO slab, and adsorbate molecule, respectively. Here, all of the $\rho[\text{diff}]$ plots are generated by fixing the fragment geometries to the optimized geometry of the entire system (optimized at the TZV2P/PBE level of theory). These electron density differences ($\rho[\text{diff}]$) using both the DZ and TZ basis are shown in the first two columns of Figure 6.7. Here, the regions colored in purple (green) correspond to regions that gained (lost) electron density in the composite system. Clearly, with both basis sets, the major changes in the electron density are near the adsorption site, with a clear gain in the electron density between the carbon atom of the CO and the Zn atom proximal to the CO. This gain in the electron density suggests the formation of a Zn–C bond in the combined system.

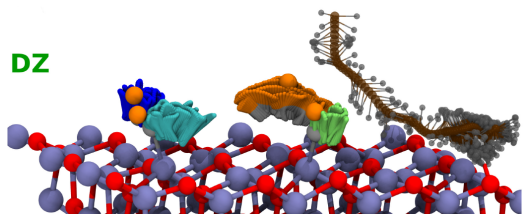


Figure 6.9(a) DZVP

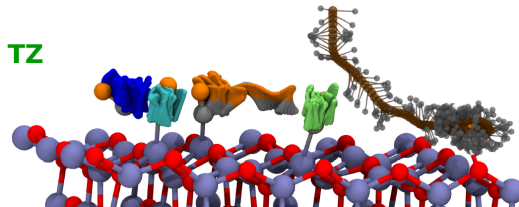


Figure 6.9(b) TZV2P

Figure 6.9: Panels (a) and (b) show how the CO molecules evolve during the 3 ps NVE simulation with the DZVP and TZV2P basis sets, respectively, for the **0.6 eV** collision. In both panels, the cyan, lime, and blue sticks show the change in the positions of 1st, 2nd, and 3rd CO molecules, respectively, for the entire trajectory (3 ps). The changes in the 4th CO position are shown with the stick model without changing its colors, and the changes in the water molecule are shown with the glass ball-and-stick model. Larger fluctuations in the position of 4th CO molecule can be seen in both cases. This large fluctuation is due to its closeness to the impact area. However, a hopping mechanism is only observed with the TZ basis.

In the last column of Figure 6.7, we have shown the difference in the interaction between the adsorbate and the slab with a change in the basis by subtracting the $\rho[\text{diff}]$ values calculated at each basis set (i.e., $\rho^{\text{basis}}[\text{diff}] = \rho^{\text{TZ}}[\text{diff}] - \rho^{\text{DZ}}[\text{diff}]$). Here, the regions colored in purple (green) correspond to regions that have larger electron density with the TZ basis (DZ basis). Clearly, for the DZ basis (colored in green), there is a larger delocalization of the electron density between Zn and C and can be related to the higher adsorption energy of CO in the DZ basis. On the other hand, for the TZ basis (colored in purple), there is a larger electron density localized at the C and O atoms (of CO), suggesting a relatively weaker Zn–C bond. Thus, the differences in the collision results with a change in the basis

can be closely related to the differences in the interaction between an adsorbate molecule and the slab with a change in the basis. We have also generated similar plots for the CO₂ molecule that are given in Figure 6.14. For the case of CO₂, $\rho^{basis}[\text{diff}]$ is negligible, suggesting that its interaction with the ZnO slab is less sensitive to changes in the basis set. This similarity in the interaction between CO₂ and ZnO, irrespective of the employed basis sets, can also be seen in the adsorption energies (Table 6.1).

6.3.5 Results with Other Software Packages

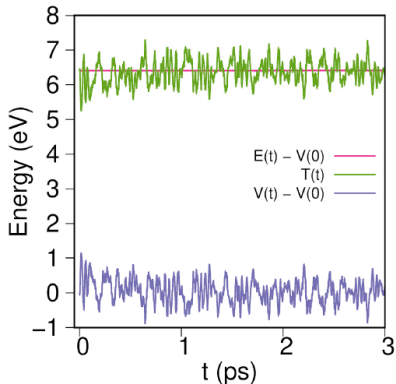


Figure 6.10(a) Energy Change

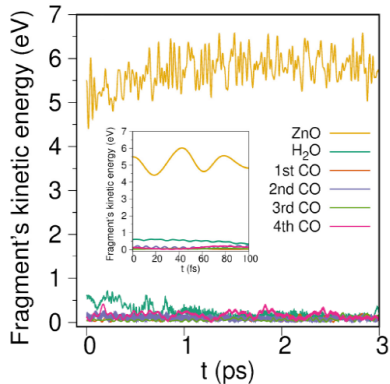


Figure 6.10(b) Kinetic Energy Change

Figure 6.10: Panel (a) shows the changes in total (E), kinetic (T), and potential (V) energies of a ZnO slab covered with four CO molecules impinged by an H₂O molecule having **0.6 eV** of kinetic energy during an NVE run (at PBE/**DZVP**). Panel (b) shows the changes in the kinetic energy of each of the fragments during the NVE run corresponding to the left panel. The inset in the right panel shows the changes in kinetic energies of the fragments during the first 100 fs of the simulation.

To verify the above findings, we have repeated the adsorption energy calculations with the FHI-aims software package with two different basis sets (tier-1 and tier-2, with tier-2 being the higher quality basis set), and the results are reported in Table 6.2. Once again, among all of the adsorbate molecules, CO is the outlier with a difference of ~ 0.1 eV

in the adsorption energy with a change in the basis, proving the robustness of our findings. We have also plotted (not shown here) $\rho^{basis}[\text{diff}]$ for CO and CO₂ using the FHI-aims package and found them to be similar to the ones in Figure 6.7. Thus, we confirmed that the basis set does have an impact on the interaction between the adsorbate and surface, and this impact is immense for the case of CO with ZnO. We have also verified these adsorption energy results with VASP (see Table C.12). Finally, we have also demonstrated that the above findings are robust against the employed exchange–correlation functional as well as the inclusion/omission of dispersion interactions (see Appendix C).

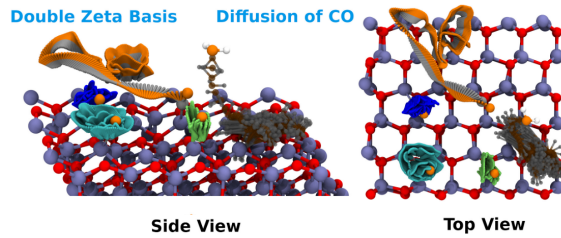


Figure 6.11(a) DZVP

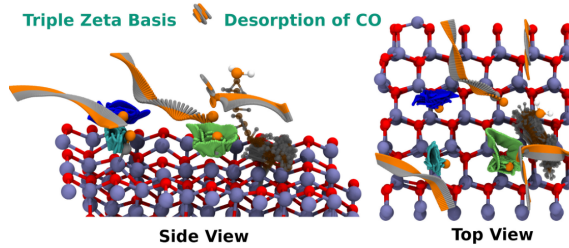


Figure 6.11(b) TZV2P

Figure 6.11: Both panels show how the CO molecules evolve during the 5 ps NVE simulation with the DZVP and TZV2P basis sets, respectively, for the 6 eV collision. In both panels, the cyan, lime, and blue trajectories show the change in the positions of 1st, 2nd, and 3rd CO molecules, respectively. The changes in the 4th CO position are shown without changing its color scheme, and the changes in the water molecule are shown with the glass ball-and-stick model. With the TZ basis, the 4th CO trajectory does not seem continuous because it crossed the periodic boundaries at the points of visual discontinuity.

6.4 Conclusion & Outlook

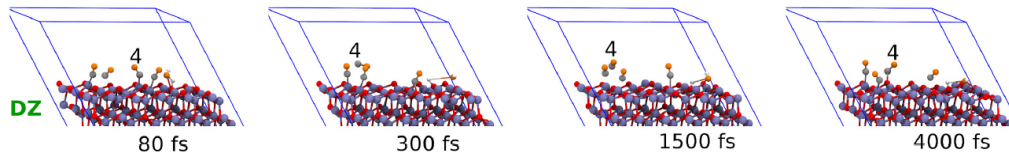


Figure 6.12(a) DZVP

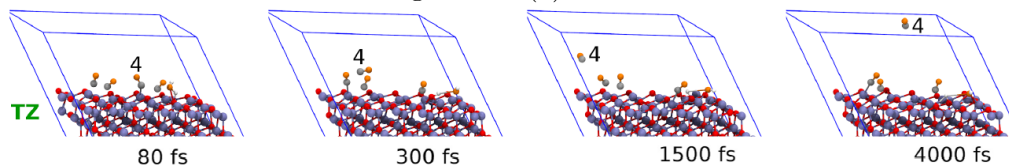


Figure 6.12(b) TZV2P

Figure 6.12: Snapshots of an NVE run of H_2O colliding with a ZnO slab at 6 eV of incident energy. (a) [(b)] panel shows the results obtained using the DZVP (TZV2P) basis set. The 4th CO is identified with the number 4 in all the frames.

Using ab initio molecular dynamics, we have examined a collision between the reactants of a water–gas shift reaction (i.e., CO and H_2O) with a ZnO surface and have shown how the resulting mechanism and dynamics are affected by the quality of the employed basis set. For both the low- and high-energy collisions, we have shown significant differences in the outcomes of a BOMD simulation by merely changing the basis-set quality. Although the differences are noticeable with the low-energy collision, they are quite substantial in the case of the high-energy collision. By analyzing the energetics of these collision reactions, we have shown that the KE gain by the adsorbate molecule from the colliding molecule is the same irrespective of the employed basis set. Using this result, we have determined the source of these differences to arise from the interaction between the surface and the adsorbate. Furthermore, through our binding energy calculations and electron density difference maps, we have clearly shown that the interaction between the adsorbate and the substrate

relies heavily on the quality of the employed basis set. These results have significant ramifications for understanding catalytic dynamics since merely changing the quality of the basis for CO translates to more than an order of magnitude (~ 0.4 eV) energy difference compared to room temperature (~ 0.026 eV). We have shown that these surprising results are robust by obtaining very similar trends with other software packages. Due to this sensitive dependence on basis sets, previous studies that have employed the commonly used double zeta basis sets may need to be re-examined (especially for CO-based reactions).

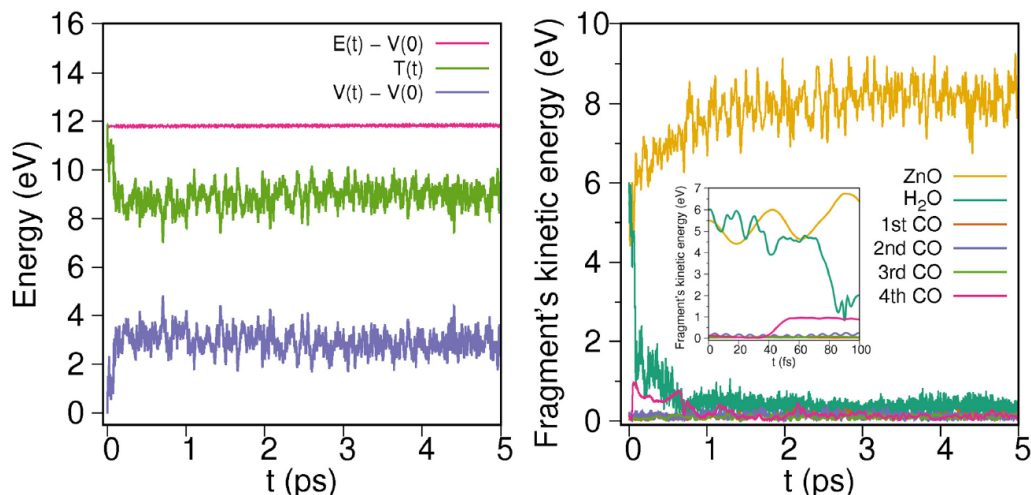


Figure 6.13: The **left** panel shows the changes in total (E), kinetic (T), and potential (V) energies of a ZnO slab covered with four CO molecules impinged by an H₂O molecule having **6 eV** of kinetic energy, during an NVE run (at the PBE/**DZVP** level of theory). The **right** panel shows the changes in the kinetic energy of each of the fragments during the NVE run corresponding to the left panel. The inset in the right panel shows the changes in kinetic energies of the fragments during the first 100 fs of the simulation.

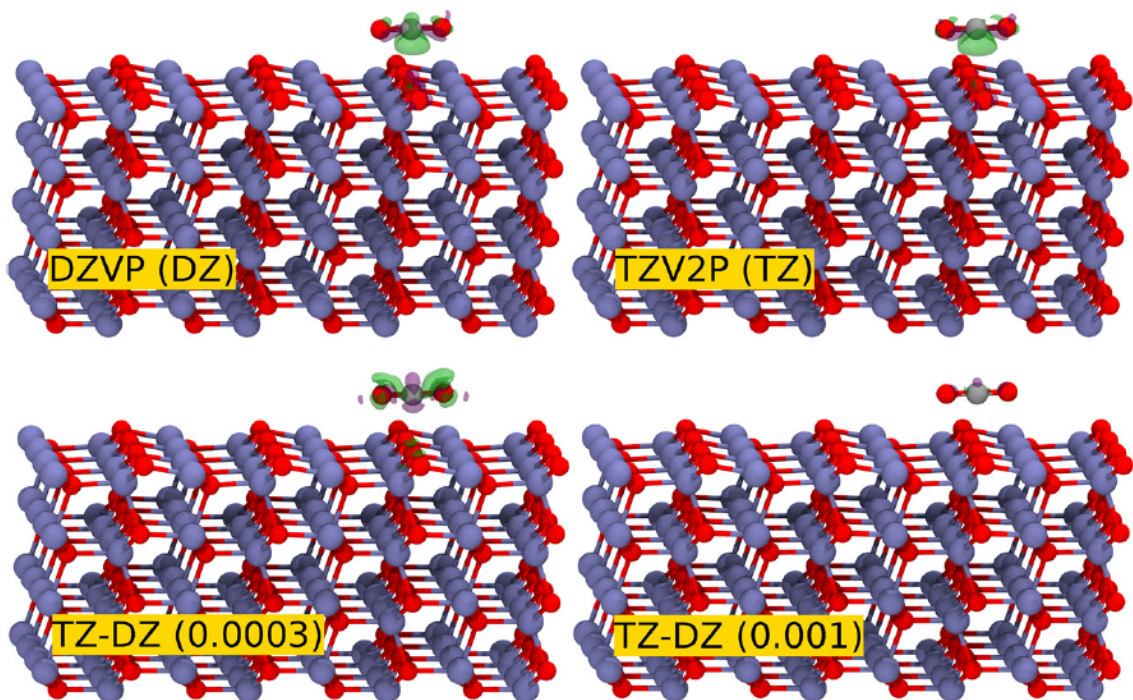


Figure 6.14: The top panel shows the electron density differences obtained for a single CO_2 molecule adsorbed on a ZnO slab obtained using the DZVP (DZ) and TZV2P (TZ) basis sets. Each column in the top panel is obtained by subtracting the total electron density of the entire system from its individual components (i.e. $\rho[\text{diff}] = \rho[\text{ZnO}+\text{mol}] - \rho[\text{ZnO}] - \rho[\text{mol}]$). The bottom panel shows the difference in the densities between the columns of the top panel (i.e., $\rho^{basis}[\text{diff}] = \rho^{TZ}[\text{diff}] - \rho^{DZ}[\text{diff}]$) plotted at two different iso values. For the top panel, we have used an iso-value of $0.001 \text{ e}\text{\AA}^{-3}$, and in the bottom panel, both 0.0003 and $0.001 \text{ e}\text{\AA}^{-3}$ iso-values are used. Clearly, even at a $0.0003 \text{ e}\text{\AA}^{-3}$ iso-value, the changes in the density with different basis sets are only observed in the CO_2 molecule but not in the interaction region between the CO_2 and ZnO. For the case of CO, there are huge changes in the interaction region, as shown in Figure 6.5.

Moving forward, we suggest that conducting a single-point binding energy calculation with different quality basis sets be first carried out as a diagnostic test to ensure that a suitable/reliable basis set is chosen prior to carrying out the computationally demanding BOMD collision reactions. Specifically, if the resulting binding energies obtained with two different quality basis sets are similar, then the resulting BOMD dynamics can still reliably (and efficiently) use the lower-quality basis set. However, if the computed single-point

binding energies from two different quality basis sets are dissimilar, extreme care should be taken since the resulting BOMD simulations will exhibit qualitatively different dynamics with these basis sets. Finally, although the conclusions drawn in this study pertain to basis set effects from BOMD simulations (where only the ground-state PES is involved), we believe that these effects are quite general and will apply even to nonadiabatic molecular dynamics simulations, which is the focus of a future study.

Chapter 7

Charge Density Wave Hampers

Exciton Condensation in 1*T*-TiSe₂

This chapter provides *ab initio* atomic-level framework for rationalizing recent experiments and further manipulating exciton condensates in TiSe₂. This chapter work resulted from a collaboration with Dr. Chao Lian. Dr. Lian implemented the linear-response Time-Dependent Density Functional Theory (lr-TDDFT) used in this chapter to help me understand the theory and expand my analysis techniques before moving toward more complicated concepts. The article was published in *the Journal of Physical Review B* [226], where the full analysis is presented. Dr. Chao Lian later extended the work presented in this chapter [227].

7.1 Introduction

Bose-Einstein condensates (BECs) exhibit exotic transport phenomena such as superfluidity in liquid Helium and superconductivity via Cooper pairs. Other bosonic quasiparticles, such as excitons [228–243], polaritons [244–253], and magnons [254–258] can also form a BEC, with exciton condensation particularly drawing immense recent attention [259]. The exciton condensate is predicted to form a superfluid current [260, 261], which is not only an exotic emergent phenomenon in fundamental quantum research but also vital for designing next-generation, scattering-free electronic devices.

From a Bardeen-Cooper-Schrieffer (BCS)-like Hamiltonian, Kohn, Jérôme, Rice, and Halperin proposed the phenomenon of exciton condensation in the 1960s [262–264]. For materials with an indirect gap, excitons spontaneously form if the exciton binding energy E_B is larger than the band gap, E_G . When $E_G < E_B$, the total energy decreases by creating excitons with an identical momentum $q = w$; i.e., the exciton condensate, where w is the reciprocal vector connecting the valence band maximum (VBM) and the conduction band minimum (CBM). The exciton condensate can be detected with momentum-resolved spectroscopy such as electron energy loss spectroscopy (EELS) [265] and resonant inelastic X-ray scattering (RIXS) [266]. The incident electron or X-ray beam can induce a non-zero momentum excitation in these spectroscopic techniques. As shown in Fig. 7.1(a), when $E_G < E_B$, the exciton condensate elastically scatters the incident beam: the beam exchanges a momentum w without energy loss [259, 262]. This creates a soft mode in the interband plasmon at the momentum $q \rightarrow w$ [Fig. 7.1(c)]. In comparison, when $E_G > E_B$, creating excitons consumes the energy of the incident beam [Fig. 7.1(b)]; thus, the energy

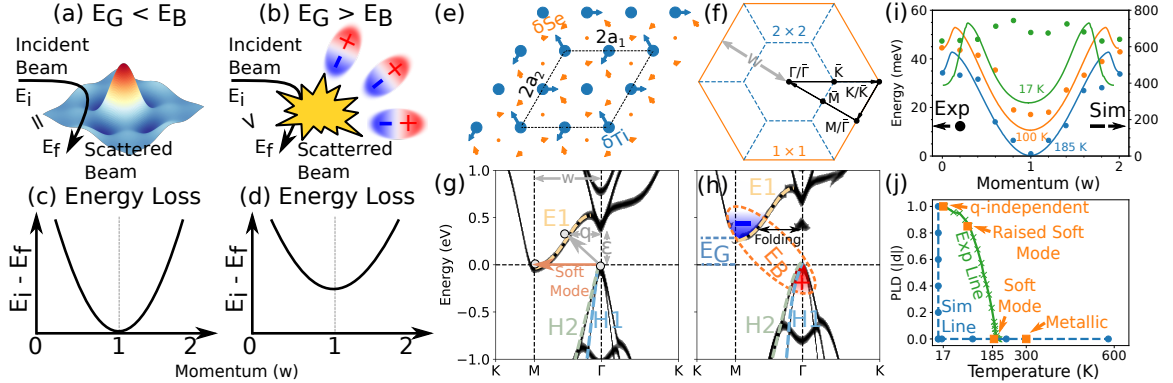


Figure 7.1: (a-d) Schematic diagram of the exciton condensation. E_G and E_B are the band gap and excitation binding energy, respectively, as shown in (h). E_i and E_f are the energies of the incident and scattered beam, respectively. w is the momentum between the CBM and the VBM as shown in (g). (e) Atomic structure of 1T-TiSe₂. The blue and orange circles denote the Ti atoms and the Se atoms, respectively. The arrows denote the periodic lattice distortion (PLD) displacements $\{\mathbf{d}_i\}$. (f) The Brillouin zone (BZ) of TiSe₂. The solid orange and dashed blue lines denote the BZ of the 1×1 and 2×2 TiSe₂ cell, respectively. $\{\Gamma, M, K\}$ and $\{\bar{\Gamma}, \bar{M}, \bar{K}\}$ denote the special k points in the 1×1 and 2×2 BZ, respectively. (g)(h) Band structures of 1T-TiSe₂ at (g) its normal state and (h) charge density wave state. (i) Experimental [259] (dots) and simulated (solid lines) plasmon dispersions at different temperatures. (j) Phase diagram of the plasmon in TiSe₂ as a function of PLD and temperature. The blue circles and orange squares denote the experimental measurements and simulation results, respectively. The experimental line describes the PLD as a function of temperature, reproduced from [267].

loss is always positive [Fig. 7.1(d)] [259, 262].

Among the materials that meet this band structure requirement, TiSe₂ [Fig. 7.1(e-h)] is an up-and-coming candidate for observing exciton condensation effects [268]. Above the transition temperature $T_c \sim 190$ K, TiSe₂ has a negative indirect gap with the VBM at the M point and the CBM at the Γ point, as shown in Fig. 7.1(g). Besides its semimetallic nature, the quasi-two-dimensional (2D) structure of TiSe₂ weakens the Coulomb screening and favors exciton binding. In a pioneering experimental study, Kogar et al. recently observed exciton condensation in TiSe₂ using EELS [259]. At T_c , the exciton condensation emerges as a soft plasmon mode near the edge of the Brillouin zone (BZ), which indicates

that zero energy consumption is required to excite the electrons. In contrast to conventional BEC effects that are always stable at a sufficiently low temperature, the soft mode is fragile to both an increase and decrease of the temperature [259]. As summarized in Fig. 7.1(i) and (j), instead of a sustainable soft plasmon mode, a momentum-independent plasmon and a regular metallic plasmon are observed at 17 and 300 K, respectively. The highly temperature-dependent exciton condensation reflects the rich physics in TiSe₂. Specifically, the well-known charge density wave (CDW) emerges below T_c , accompanied with a 2×2 periodic lattice distortion (PLD) $\{\mathbf{d}_i\}$ shown in Fig. 7.1(e). Although the CDW state has been extensively studied as a prototypical example of the CDW-superconductivity (SC) transition [269–293], the driving force of the CDW is still under debate regarding whether it is a pure electronic exciton-related mechanism [294–319] or an electron-phonon coupling (EPC) mechanism [320–335]. The complex interplay among the thermal field, CDW order, and exciton condensation are beyond existing conventional models based on the assumption of fixed ions and single-particle bands [262–264, 268]. As such, a first-principles-based *ab initio* framework is essential for understanding this entangled system and tuning the properties of the exciton condensation.

In this article, we accurately reproduce and characterize the experimental spectra using linear-response time-dependent density functional theory (lr-TDDFT). At the normal state, we observe a soft plasmon mode directly representing the exciton condensation, which can be attributed to interband electronic transitions. At the CDW state below T_c , the periodic lattice distortions introduce a CDW gap and hamper the spontaneous formation of the exciton. Above T_c , the higher electronic temperature prevents condensation by

introducing a dissipation channel. The combined effect of the CDW and finite electronic temperature explains the fragile temperature dependence of the exciton condensation. Our work provides the first *ab initio* atomic-level framework, beyond widely-used simplistic analytical models, for rationalizing recent experiments and further manipulating exciton condensates in TiSe₂.

7.2 Methodology

The GPAW [336–338] package was used to carry out the ground state and lr-TDDFT calculations. The projector augmented-waves (PAW) [339] method and Perdew-Burke-Ernzerhof (PBE) XC functional [340] were used. The plane-wave cutoff energy was set to 750 eV. The on-site Coulomb interaction of the 3d orbitals of Ti was 3.5 eV. The Brillouin zone was sampled using the Monkhorst-Pack (MP) scheme [341] with a $48 \times 48 \times 1$ k-point mesh. The computational cell examined in this work was a 2×2 supercell of TiSe₂, containing 4 Ti atoms and 12 Se atoms. In the lr-TDDFT calculations, the Bootstrap XC kernel [342] was utilized together with the random phase approximation (RPA) as a comparison.

We calculate the EELS, i.e. the frequency and wave-vector dependent density response functions, based on the lr-TDDFT formalism [343–345]. The non-interacting density response function in real space is written as

$$\chi^0(\mathbf{r}, \mathbf{r}', \omega) = \sum_{\mathbf{k}, \mathbf{q}} \sum_{n, n'}^{\text{BZ}} \frac{f_{n\mathbf{k}} - f_{n'\mathbf{k}+\mathbf{q}}}{\omega + \epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}+\mathbf{q}} + i\eta} \times \psi_{n\mathbf{k}}^*(\mathbf{r}) \psi_{n'\mathbf{k}+\mathbf{q}}(\mathbf{r}) \psi_{n\mathbf{k}}(\mathbf{r}') \psi_{n'\mathbf{k}+\mathbf{q}}^*(\mathbf{r}'), \quad (7.1)$$

where n is the band index, \mathbf{k} is the k index, \mathbf{q} stands for the Bloch vector of the incident

wave, $\eta \rightarrow 0$, and $\epsilon_{n\mathbf{k}}$ and $\psi_{n\mathbf{k}}(\mathbf{r})$ are the eigenvalues and eigenvectors of the ground state Hamiltonian, respectively. The full interacting density response function is obtained by solving Dyson's equation from its non-interacting counterpart χ^0 as

$$\begin{aligned} \chi(\mathbf{r}, \mathbf{r}', \omega) = & \chi_0(\mathbf{r}, \mathbf{r}', \omega) \\ & + \iint_{\Omega} d\mathbf{r}_1 d\mathbf{r}_2 \chi_0(\mathbf{r}, \mathbf{r}_1, \omega) K(\mathbf{r}_1, \mathbf{r}_2) \chi(\mathbf{r}_2, \mathbf{r}', \omega), \end{aligned} \quad (7.2)$$

where the kernel is the summation of the coulomb and exchange-correlation (XC) interaction

$$K(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} + f_{xc}. \quad (7.3)$$

Here, $f_{xc} = \partial V_{xc}[n]/\partial n$ is the XC kernel. The commonly-used XC kernels include the adiabatic local density approximation (ALDA) [340], Bootstrap approximation [342], etc.

One of the simplest cases is the random phase approximation (RPA), with $f_{xc} = 0$.

For a system possessing translational symmetry, it is more convenient to represent χ^0 in the reciprocal lattice space:

$$\chi^0(\mathbf{r}, \mathbf{r}', \omega) = \frac{1}{\Omega} \sum_{\mathbf{q}} \sum_{\mathbf{G}\mathbf{G}'} e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} \chi_{\mathbf{G}\mathbf{G}'}^0(\mathbf{q}, \omega) e^{-i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'}, \quad (7.4)$$

where Ω is the normalization volume and $\mathbf{G}(\mathbf{G}')$ are reciprocal lattice vectors. The Fourier coefficients $\chi_{\mathbf{G}\mathbf{G}'}^0(\mathbf{q}, \omega)$ are written as

$$\chi_{\mathbf{G}\mathbf{G}'}^0(\mathbf{q}, \omega) = \sum_{n,n'} \chi_{\mathbf{G}\mathbf{G}'n,n'}^0(\mathbf{q}, \omega) \quad (7.5)$$

where

$$\begin{aligned} \chi_{\mathbf{G}\mathbf{G}'n,n'}^0(\mathbf{q}, \omega) = & \frac{1}{\Omega} \sum_{\mathbf{k}} \frac{f_{n\mathbf{k}} - f_{n'\mathbf{k}+\mathbf{q}}}{\omega + \epsilon_{n\mathbf{k}} - \epsilon_{n'\mathbf{k}+\mathbf{q}} + i\eta} \\ & \times \langle \psi_{n\mathbf{k}} | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} | \psi_{n'\mathbf{k}+\mathbf{q}} \rangle_{\Omega_{\text{cell}}} \\ & \times \langle \psi_{n\mathbf{k}} | e^{i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'} | \psi_{n'\mathbf{k}+\mathbf{q}} \rangle_{\Omega_{\text{cell}}}, \end{aligned} \quad (7.6)$$

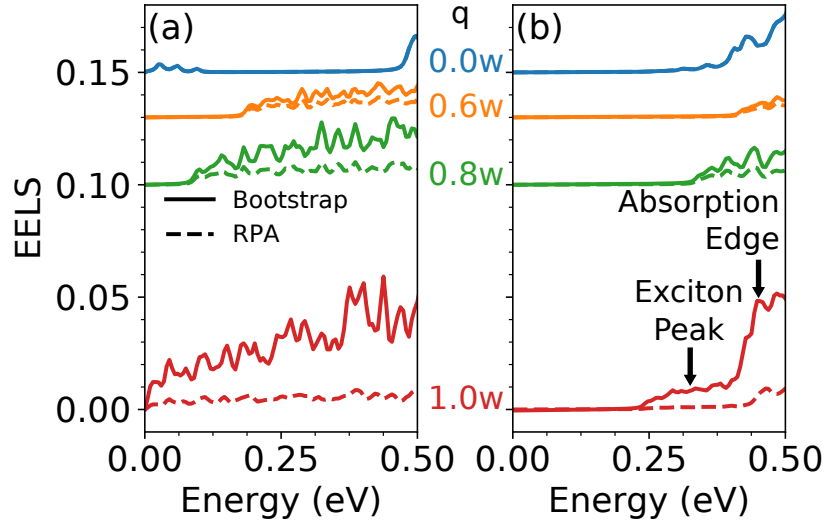


Figure 7.2: Comparison of the EELS with the random phase approximation (RPA) vs Bootstrap XC kernels for (a) the normal state (b) the CDW state.

Dyson's equation is expressed in the \mathbf{G} basis as

$$\begin{aligned} \chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega) &= \chi_{\mathbf{G}\mathbf{G}'}^0(\mathbf{q}, \omega) \\ &+ \sum_{\mathbf{G}_1 \mathbf{G}_2} \chi_{\mathbf{G}\mathbf{G}_1}^0(\mathbf{q}, \omega) K_{\mathbf{G}_1 \mathbf{G}_2}(\mathbf{q}) \chi_{\mathbf{G}_2 \mathbf{G}'}(\mathbf{q}, \omega). \end{aligned} \quad (7.7)$$

The dielectric function can be expressed with $\chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega)$ as

$$\epsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}, \omega) = \delta_{\mathbf{G}\mathbf{G}'} - \sum_{\mathbf{G}_1} K_{\mathbf{G}\mathbf{G}_1}(\mathbf{q}) \chi_{\mathbf{G}_1 \mathbf{G}'}(\mathbf{q}, \omega). \quad (7.8)$$

The macroscopic dielectric function is defined by

$$\epsilon_M(\mathbf{q}, \omega) = \frac{1}{\epsilon_{00}^{-1}(\mathbf{q}, \omega)}, \quad (7.9)$$

and the electron energy loss spectrum (EELS) is

$$\text{EELS} = -\text{Im} \frac{1}{\epsilon_M(\mathbf{q}, \omega)}. \quad (7.10)$$

We briefly discuss the accuracy of lr-TDDFT in describing the exciton in TiSe_2 .

It is known that semi-local XCs (e.g. PBE) poorly describe long-range Coulomb screen-

ing [342], and long-range corrections, such as in the Bethe–Salpeter equation (BSE) and long-range-corrected XC functionals [346–351], are required for improved accuracy. However, the exciton in TiSe_2 is formed by an attractive interaction $V(w)$ between the electron pocket at the M point and the hole pocket at the Γ point, as shown in Fig. 7.1(d). Here, $w = \pm 0.5|\mathbf{b}_1|$ is the reciprocal vector and \mathbf{b}_1 ($i = 1, 2$) is the reciprocal lattice vector along the i th direction, as shown in Fig. 7.1(c). Thus, $V(w)$ is a medium-ranged interaction with a characteristic distance of $1/w = a$, where $a = |\mathbf{a}_i|$ and \mathbf{a}_i is the lattice vector. $V(w)$ is distinctly different from typical long-range interactions that are characteristic of vertical excitonic excitations in momentum space.

We quantitatively characterize the exciton effect by comparing the EELS calculated with the RPA and the Bootstrap XC kernel, as shown in Fig. 7.2. The Bootstrap XC kernel is designed to correct the long-range error of the Coulomb interaction and generates accurate exciton peaks at a similar computational cost to ALDA [342]. The bootstrap XC kernel can yield a good description of exciton when screening is not high, which is suitable for many low-dimensional and semi-metallic systems. At $q \sim 0$, the EELS spectra with the RPA and Bootstrap are generally the same, indicating the absence of the exciton-type excitation. We note that with increasing q , the difference increases and reaches a maximum at $q = w$, corresponding to the emergence of exciton excitations. As shown in Fig. 7.2(b), a new exciton peak can be observed only in the Bootstrap calculation at $q \sim 0$ around 320 meV, which is 100 meV lower than the absorption edge. This indicates that the exciton binding energy E_B in the Bootstrap calculations is consistent with the results from BSE, $E_B = 75$ meV [316].

The CDW phase of TiSe₂ is a 2×2 cell of the normal phase. The energy bands are folded from the 1×1 BZ to the 2×2 BZ. In contrast, ARPES measurements still span over the 1×1 BZ. To bridge the gap between the DFT bands and the measured ARPES spectra, the band unfolding technique is used to calculate the effective band structure (EBS) [352, 353] of the supercell (SC). Expanding the adiabatic basis $|\phi_{i,\mathbf{k}}\rangle$ of the 2×2 SC in the wavefunction $|\Phi_{I,\mathbf{K}}\rangle$ of primitive 1×1 cell (PC), we get

$$|\phi_{i,\mathbf{k}}(\mathbf{G}, t)\rangle = \sum_{I,\mathbf{K}} a(I, \mathbf{K}; i, \mathbf{k}; t) |\Phi_{I,\mathbf{K}}(\mathbf{G}, t)\rangle, \quad (7.11)$$

where $\mathbf{K} = \mathbf{k} + \mathbf{B}$, \mathbf{B} is the reciprocal basis vector of SC, and $a(I, \mathbf{K}; i, \mathbf{k}; t)$ is the coefficient of $|\Phi_{I,\mathbf{K}}\rangle$ as the basis of $|\phi_{i,\mathbf{k}}\rangle$. The spectral function is the EBS along the \mathbf{K} path in PCBZ:

$$A(\mathbf{K}, E, t) = \sum_i P(\mathbf{K}; \mathbf{k}, i; t) \delta(E - \epsilon_{i,\mathbf{k}}), \quad (7.12)$$

where $\epsilon_{i,\mathbf{k}}$ is the eigenvalue of band i at the \mathbf{k} point, E is the energy, and

$$P(\mathbf{K}; \mathbf{k}, i) = \sum_{\mathbf{G}} |\phi_{i,\mathbf{k}}(\mathbf{G} + \mathbf{K} - \mathbf{k}, t)|^2, \quad (7.13)$$

We can introduce an extra weight function $w(i, \mathbf{k})$

$$A(\mathbf{K}, E, t) = \sum_i P(\mathbf{K}; \mathbf{k}, i; t) w_{i,\mathbf{k}} \delta(E - \epsilon_{i,\mathbf{k}}). \quad (7.14)$$

The choice of $w_{i,\mathbf{k}}$ is arbitrary [354]. Here, we use the Fermi-Dirac distribution

$$w_{i,\mathbf{k}} = \left[1 + \exp\left(\frac{\epsilon_{i,\mathbf{k}} - E_F}{T_e}\right) \right]^{-1}, \quad (7.15)$$

where E_F is the Fermi energy and T_e is the electronic temperature.

7.3 Results and Discussion

We calculate the structures of the normal 1×1 and 2×2 CDW phase and obtain the optimized PLD displacements $\{\mathbf{d}_i\}$ shown in Fig. 7.1(e), with $\delta_{Ti} = 0.083 \text{ \AA}$, $\delta_{Se} = 0.027 \text{ \AA}$, and $\delta_{Ti}/\delta_{Se} = 3.07 : 1$. These results accurately reproduce the experimental measurements of $\delta_{Ti} = 0.085 \pm 0.014 \text{ \AA}$ and $\delta_{Ti}/\delta_{Se} \sim 3 : 1$ [267]. To directly compare with the angle-resolved photoemission spectroscopy (ARPES) measurements, we unfold the energy bands from the 2×2 BZ to the 1×1 BZ to generate the effective band structures (EBS) along the K-M- Γ -K symmetry points, as shown in Fig. 7.1(f). In its normal state, TiSe₂ is semimetallic [Fig. 7.1(g)]: the conduction band E1 touches the Fermi energy at the M point; two valence bands, H1 and H2, touch the Fermi energy at the Γ point. At low temperatures, the CDW state of 1T-TiSe₂ is a semiconductor. The 2×2 CDW order backfolds the H1 band to the M point. Induced by the avoided crossing, a repulsive force raises the E1 band and opens a gap of $E_G \sim 0.27 \text{ eV}$ [Fig. 7.1(h)]. Via the same mechanism, the backfolded E1 band at the Γ point cuts the H1 and H2 bands into valence and conduction parts and upshifts the conduction parts of the H1 and H2 bands.

As previously mentioned, the soft mode is highly temperature-dependent and distinguishable from other BEC phenomena. Thus, the knowledge of thermal field effects is essential for understanding exciton condensation. We propose that the thermal field mainly influences the electronic structures in two aspects: **(I)** The decrease in temperature stabilizes the PLD and introduces a CDW gap. **(II)** The increase in temperature of the electron system, T_e , creates more thermal carriers. As shown in Fig. 7.1(j), both T_e and PLD cannot be dis-entangled in the experiment, and only the states on the experimental line can

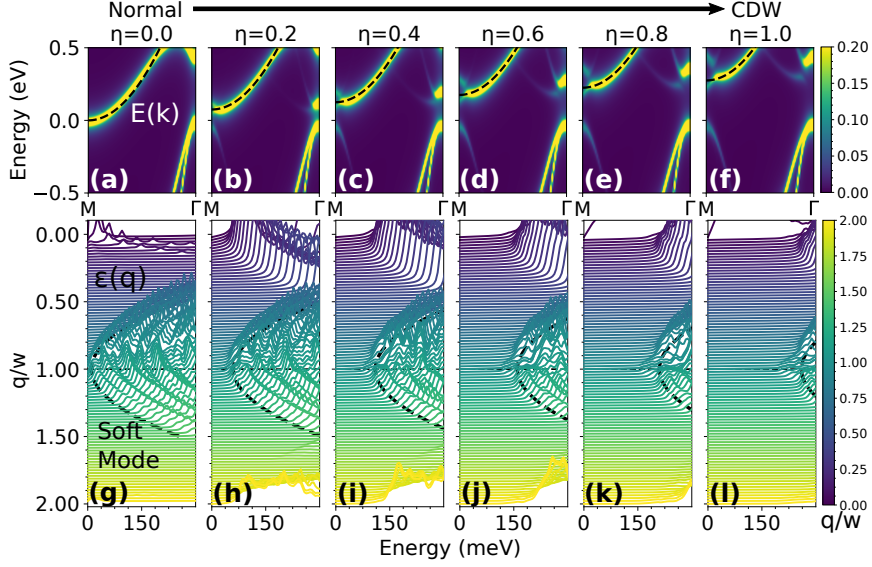


Figure 7.3: (a-f) Effective band structures (EBS) and (g-l) electron energy loss spectra (EELS) at different PLDs $\{\eta\mathbf{d}_i\}$. The colorbars for (a-f) and (g-l) denote the spectral weight [Eq. 7.14] of the EBS and the momentum q of the EELS [Eq. 7.10], respectively. The dashed lines in panels (a-f) and (g-l) denote the parabolic fittings of the band structures $E(k)$ [Eq. 7.16] near the M point and the plasmon dispersions $\epsilon(q)$ [Eq. 7.17] near w , respectively.

be observed. Nevertheless, our *ab initio* techniques are not limited to these experimentally-observed states and can further sample the phase diagram and characterize the effects of the PLD and T_e separately.

We first examine the effects of the PLD by calculating the EBS and EELS for 6 structures with different PLDs $\{\eta\mathbf{d}_i\}$ at $T_e = 0$ K, where $\eta = 0.0, 0.2, \dots, 1.0$, i.e., the states along the Y axis in Fig 7.1(j). $\eta = 0.0$ and $\eta = 1.0$ correspond to the normal and CDW state, respectively. The intermediate η values describe the structures at low but finite temperatures $0 < T < T_c$.

We observed the soft mode [Fig 7.1(g)] at the normal state with $\eta = 0$. As shown in Fig. 7.3(a) and (g), the interband plasmon energy $\epsilon(q)$ decreases with $q < w$, reaches the zero point $\epsilon(q) = 0$ eV at $q = w$ and then increases with $q > w$. This behavior is

consistent with the experimental observations [259], although the rate of change is higher in our simulation. The experimental energy range of the soft mode is 0–40 meV, whereas we obtain a range of 0–400 meV from our simulations due to the underestimation of the Coulomb screening in the semi-local exchange-correlation functionals.

The CDW gap increases with η , as shown in Fig. 7.3(a-f). To quantitatively describe the effect of the PLD, we construct an expression for this band-plasmon correspondence. As shown in Fig. 7.3(a-f), the low-energy-range band can be accurately described by the expression

$$E(k) = \max\{\alpha(k - k_M)^2 + E_G, 0\}, \quad (7.16)$$

where k_M is the k coordinates of the M point, $\alpha = 0.95 \text{ eV}/\text{\AA}^2$, and $E_G = 250\eta \text{ meV}$ are the CDW gaps. As shown in Fig. 7.3(g-l), we find that the interband plasmon dispersion near w can be accurately described by the expression

$$\epsilon(q) = \max\{\alpha(q - w)^2 + E_G, 0\}, \quad (7.17)$$

with the same values of $\alpha = 0.95 \text{ eV}/\text{\AA}^2$ and $E_G = 250\eta \text{ meV}$ in Eq. 7.16. This indicates that the incident beam pumps electrons from Γ to the E1 band, producing the EELS signal near w , as shown in Fig. 7.1(g). We observed a soft mode at the normal state with zero excitation energy at $q = w$. In addition, the band gap increases with η and raises the soft mode to higher energy. The excitation energy at $q = w$ concomitantly increases from zero to a finite value, preventing spontaneous exciton condensation.

Next, we discuss the influence of electronic temperature T_e on the exciton condensation. The finite electronic temperature T_e is approximated by broadening the Fermi-Dirac

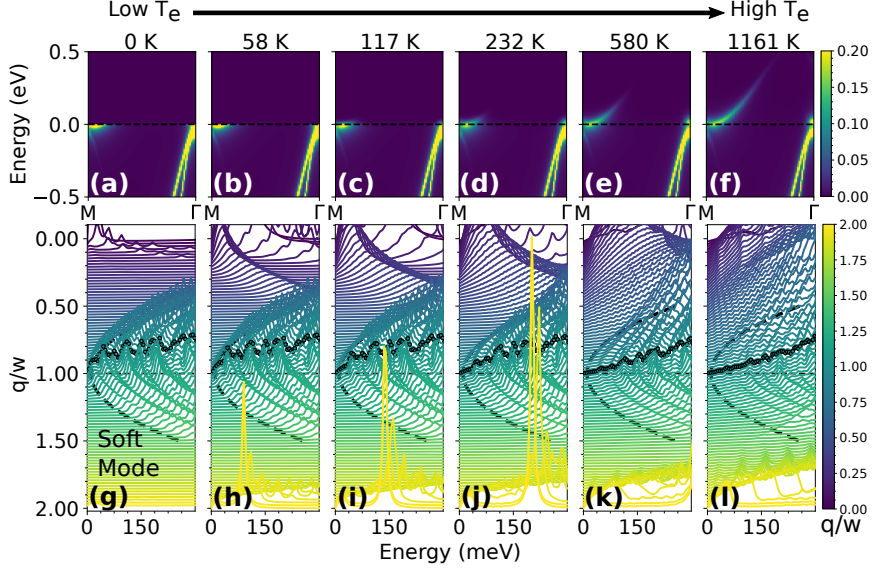


Figure 7.4: (a-f) Effective band structures (EBS) and (g-l) electron energy loss spectra (EELS) at different electronic temperatures T_e . The colorbars for (a-f) and (g-l) denote the electron population of EBS [Eq. 7.14 and 7.15] and the momentum q of the EELS [Eq. 7.10], respectively. The dashed lines in (g)-(l) denote the plasmon dispersions [Eq. 7.17] near w . The dark dots denote the EELS at $q = w$.

distribution as

$$f(E) = \frac{1}{1 + \exp[(E - E_F)/k_B T_e]}, \quad (7.18)$$

where E is the energy of the electron, E_F is the Fermi energy, and k_B is the Boltzmann constant. Due to the semi-metallic nature of TiSe_2 , the carrier density is fairly sensitive to the temperature. We simulated the EELS at a different electronic temperature T_e with $\eta = 0.0$, as shown in Fig. 7.4. The increase in T_e changes the electron distribution but barely alters the band structure [Fig. 7.4(a)-(f)]. Accordingly, the thermal carriers gradually fill the empty bands as T_e increases. Compared with the unified quantum state of the exciton condensate, the thermal carriers form excitons with a wide range of momenta. The thermally-excited excitons with random momenta collide with and accelerate the decoherence of the exciton condensate. As shown in Fig. 7.4(g)-(l), the intensity of the exciton

excitation decreases due to the partially-occupied initial and final state. This effect broadens the peak and decreases the peak height without changing the peak positions of the soft mode.

Besides the dissipation introduced by the thermal charge carriers, the increasing phonon density may also destroy the exciton condensation. Both linear-response time-dependent density functional theory (lr-TDDFT) [343–345] and density-functional perturbation theory (DFPT) [355, 356] are required to describe the phonon-induced linewidth broadening in the plasmon spectra. To the best of our knowledge, although temperature-dependent quasiparticle linewidths and optical spectra can be calculated with several DFT packages such as Abinit [357], Yambo [358], and EPW [359], \mathbf{k} -resolved algorithms to calculate temperature-dependent plasmon spectra are still under development. Thus, we only considered the dissipation induced by thermal charge carriers. A large density of acoustic phonons will accelerate the dissipation of exciton condensation at high temperatures, while its mechanism is still an important open question.

We note that frequency-independent xc kernels such as the adiabatic local density approximation and bootstrap [342] neglect disorder and phonon scattering. In addition, the damping due to electronic many-body effects is absent. This effect can be introduced in the non-adiabatic xc kernel based on the time-dependent current density functional formalism [360–365], which shows good quantitative agreement with experimental linewidth data [366–370]. Furthermore, increasing the temperature introduces extra damping from the enhanced phonon scattering, leading to a faster vanishing of the soft mode signals when the temperature increases above T_c . Please refer to Appendix D for a concise comparison

of the results and Ref. [226] for the complete analysis.

7.4 Conclusion

In summary, we have developed a new *ab initio* atomic-level framework of the exciton condensation in TiSe_2 . Our lr-TDDFT approach accurately reproduces the experimental spectra and explains the complex interplay among the thermal field and CDW order effects that were recently observed in exciton condensation experiments. The soft mode that characterizes the exciton condensation can be attributed to interband electronic transitions. Furthermore, the fragile temperature dependence of the exciton condensation is the combined effect of the CDW and thermal carriers: Below T_c , the periodic lattice distortions hamper the spontaneous formation of the exciton by introducing a CDW gap; above T_c , a higher electronic temperature produces sufficient thermal carriers and introduces a dissipation channel. This explains why the soft mode is only observed at T_c . Our *ab initio* framework provides critical mechanistic insight into recent exciton condensation experiments and presents additional avenues to experimentalists for further manipulating exciton condensates in TiSe_2 .

Chapter 8

Implementation of Real-Time TDDFT for Periodic Systems in the Open-Source PySCF Software Package

This chapter introduces rt-TDDFT, then proceeds to the complete formalism in a velocity gauge framework for isolated and periodic systems. The purpose of the study is to validate the rt-TDDFT implementation. The chapter concludes by presenting validation results. The majority of the work in this chapter resulted from a collaboration with Dr. Kota Hanasaki. Aside from calculations, he helped successfully develop and implement the theory. The manuscript is still being written at the time of writing this chapter [\[371\]](#).

8.1 Introduction

The broad range of field-induced real-time (RT) dynamics in condensed phases has many potential applications, including field-induced transitions [1, 2], surface reactions [3], photocatalysis [4, 5], and strong-field dynamics [6–8]. These applications can help pave the way for real-time dynamics to quantitatively investigate the dynamics of charge carriers. Theoretical reproduction and understanding of field-induced dynamics are also critical for the design and analysis of nanoscale devices [9]. Dynamical process analyses in solids require a size-scalable excited-state calculation tool. The time-dependent density functional theory (TDDFT) [372] is undoubtedly one of the most promising tools for real-time simulations. The dipole-gauge formulation is challenging to incorporate external field effects, commonly used in molecular calculations, in condensed phases since a strong field can break the Hamiltonian’s translational symmetry [373]. Although real-time dynamics in condensed phases has many potential applications, due to the challenges posed by field effects, implementations of TDDFT for periodic systems under the influence of external fields are scarce.

Several RT-TDDFT implementations, applicable to condensed-phase systems, have recently been developed, including plane-wave-based codes such as TDAP [374] and Q-box [375], a real-space grid-based code Salmon [376, 377], a full-potential linearized augmented-plane wave (FP-LAPW) code called Elk [378], and a numerical atomic orbital based code Siesta [379], to name a few. These programs are either in-house, not open-source, written in a low language, or some other limiting factor, making it difficult for developers or practitioners to extend or modify the implemented methods, making availability difficult. To fill this

void, we implement a Gaussian-based RT-TDDFT method in the velocity gauge formalism, enabling molecular scientists to, conveniently, extend their work to periodic systems.

Gaussian basis sets have many computational advantages, one being their compact size. They are tailored for each atom to reproduce the highly non-homogeneous densities in isolated systems. Gaussian basis sets are small and accurate compared to plane waves, which require a high cutoff and large supercell to reproduce the electronic densities near nuclei. They also allow the analytic evaluation of the kinetic energy and Coulombic potential matrices. The advantage is less prominent in condensed-phase calculations in the pseudopotential formulation since solids' extended (pseudo-) wave functions can be efficiently represented on a numerical grid or by plane-wave expansion. We, however, found that the Gaussian basis functions adapted for the pseudopotential DFT calculations [380] and several numerical techniques implemented in PySCF [381], such as the density-fitting [382–385], enable calculations of periodic systems in a reasonable time and cost. We also consider the linear combination of atomic orbital (LCAO) scheme to be a suitable choice for calculating surfaces and low-dimensional nanoscale devices, which are the main targets of our real-time simulations. Aside from having a large user base, we employ the velocity-gauge formulation, implementing our Gaussian-based RT-TDDFT code in the PySCF program suite [381].

The modular structure of the Python-based PySCF code makes modification and extension of the implemented methodologies easy, even for those unfamiliar with each implementation detail. Recently, there has been considerable growth in the use of the Python programming language within the scientific community. Python is regarded as possessing advantages in terms of code readability and re-usability [386]. Python supports rapid de-

velopment practices and easy program maintenance as a high-level language. Due to its favorable features, Qiming Sun and his group developed their open-source project, PySCF, in the Python language [381].

In this work, we aim to tackle the challenges mentioned above in Gaussian-based RT-TDDFT. We introduce a real-time ab initio approach based on Gaussian basis sets for simulating electron dynamics under laser excited conditions. We proceed with benchmark test to validate our implementation. Our results are complemented by a variety of analyses, including real-time electronic properties and time-dependent orbital occupations. Both the length and vector gauge of the electromagnetic field has been implemented. This approach enables RT-TDDFT calculations of solids and surfaces using relatively simple unit cells, reducing computational costs by several orders of magnitudes. Finally, we conclude with a discussion and summary of our results, with additional perspectives of future applications of RT-TDDFT that can have a broad impact in.

8.2 Methodology

8.2.1 Real-time propagation scheme

We first describe the real-time propagation scheme used in our custom PySCF implementation. We commence with the electronic ground state of a periodic system described by a given nuclear configuration with lattice vectors $\{\mathbf{a}_i\}$ and reciprocal vectors \mathbf{G} . Each Kohn-Sham (KS) orbital is characterized by a Bloch vector \mathbf{k} and is a solution of the

eigenvalue equation,

$$H_{\mathbf{k}}^{\text{KS}}\psi_{\mathbf{k}\alpha} = \varepsilon_{\mathbf{k}\alpha}\psi_{\mathbf{k}\alpha}, \quad (8.1)$$

where α is an orbital index and H^{KS} is the KS Hamiltonian given by

$$H^{\text{KS}} = \mathcal{T} + V_{\text{ext}} + V_{\text{H}} + V_{\text{xc}}, \quad (8.2)$$

where \mathcal{T} , V_{ext} , V_{H} and V_{xc} are the kinetic energy, external potential, Hartree potential, and exchange-correlation (xc) potential operators, respectively.

For a system is in its electronic ground state at time $t = 0$, its time evolution for future times, $t > 0$, is given by the time-dependent KS equation

$$i\hbar\frac{\partial}{\partial t}\psi_{\mathbf{k}\alpha}(\mathbf{r}, t) = H_{\mathbf{k}}^{\text{KS}}(t)\psi_{\mathbf{k}\alpha}(\mathbf{r}, t), \quad (8.3)$$

where $H_{\mathbf{k}}^{\text{KS}}(t)$ is the time-dependent KS Hamiltonian operator. The electron density of the system at time t is given by

$$\rho(\mathbf{r}, \mathbf{r}'; t) = \sum_{\mathbf{k}\alpha} \psi_{\mathbf{k}\alpha}(\mathbf{r}, t)w_{\mathbf{k}\alpha}\psi_{\mathbf{k}\alpha}^*(\mathbf{r}', t), \quad (8.4)$$

where $w_{\mathbf{k}\alpha}$ is a time-independent occupation number. In this paper, we use the adiabatic formulation [387] of TDDFT, which approximates the xc potential to be a functional of the density at time t . As representative examples of our implementation, we use local and semi-local xc functionals such as the local density approximation (LDA) [388] and the generalized gradient approximations (GGA) [389, 390]. We recognize that hybrid functionals give better results than LDA and GGA functionals, but the purpose of this study is to validate our implementation.

To numerically solve the time evolution in equation (8.3), we apply the Crank-Nicolson algorithm [391] and propagate the KS orbitals as

$$\psi_{\mathbf{k}\alpha}(t + \Delta t) \approx \prod_{j=0}^{N_s-1} \left[1 + \frac{idt}{2\hbar} H_{\mathbf{k}}^{\text{KS}}(t_{j+\frac{1}{2}}) \right]^{-1} \left[1 - \frac{idt}{2\hbar} H_{\mathbf{k}}^{\text{KS}}(t_{j+\frac{1}{2}}) \right] \psi_{\mathbf{k}\alpha}(t), \quad (8.5)$$

is the propagation time step, and $t_{j+\frac{1}{2}} \equiv t + (j + \frac{1}{2})dt$ are the time values. The Crank-Nicolson operator in Eq. (8.5) is a strictly unitary approximation of the exact time evolution operator $e^{-i\Delta t H^{\text{KS}}/\hbar}$ for a small step size Δt . The density matrix and Hamiltonian at time $t + \Delta t$ are solved self-consistently, whereas the Hamiltonians at the intermediate time points, $H_{\mathbf{k}}^{\text{KS}}(t_{j+\frac{1}{2}})$, are approximated by a linear interpolation of $H_{\mathbf{k}}^{\text{KS}}(t)$ and $H_{\mathbf{k}}^{\text{KS}}(t + \Delta t)$ as suggested in Ref. [374](b).

Within the PySCF [381] software package, the time-dependent KS orbitals are expanded in a set of Gaussian atomic orbitals (AOs) as

$$\psi_{\mathbf{k}\alpha}(\mathbf{r}, t) = \sum_{\mu} \chi_{\mathbf{k}\mu}(\mathbf{r}) C_{\mathbf{k}\alpha}^{\mu}(t), \quad (8.6)$$

where $C_{\mathbf{k}\alpha}^{\mu}(t)$ are the time-dependent coefficients, $\chi_{\mathbf{k}\mu}(\mathbf{r}) \equiv \sqrt{\frac{1}{\mathcal{N}}} \sum_{\mathbf{T}_n} \chi_{\mu}(\mathbf{r} - \mathbf{R}_{\mu} - \mathbf{T}_n) e^{i\mathbf{k} \cdot \mathbf{T}_n}$ are the Gaussian basis functions with \mathbf{R}_{μ} being the atomic center of orbital μ , \mathbf{T}_n is the lattice translation, and \mathcal{N} is a formal normalization factor. The electron density, Eq. (8.4), is expanded as

$$\rho(\mathbf{r}, \mathbf{r}'; t) = \sum_{\mathbf{k}} \sum_{\mu\nu} \chi_{\mathbf{k}\mu}(\mathbf{r}) D_{\mathbf{k}\mu\nu}(t) \chi_{\mathbf{k}\nu}^*(\mathbf{r}'), \quad (8.7)$$

where $D_{\mathbf{k}\mu\nu}(t)$ is the density matrix in the AO representation, given by

$$D_{\mathbf{k}\mu\nu}(t) \equiv \sum_{\alpha} C_{\mathbf{k}\alpha}^{\mu}(t) w_{\mathbf{k}\alpha} C_{\mathbf{k}\alpha}^{\nu*}(t). \quad (8.8)$$

With these definitions, the time-dependent KS equation in Eq. (8.3) can be written in matrix form as

$$i\hbar \sum_{\nu} S_{\mu\nu} \frac{\partial}{\partial t} C_{\mathbf{k}\alpha}^{\nu}(t) = \sum_{\nu} \left(H - i\hbar \dot{\mathbf{R}} \cdot \mathbf{X} \right)_{\mu\nu} C_{\mathbf{k}\alpha}^{\nu}(t), \quad (8.9)$$

where $X_{\mathbf{k}\mu\nu} \equiv \langle \chi_{\mathbf{k}\mu} | \frac{\partial}{\partial \mathbf{R}} | \chi_{\mathbf{k}\nu} \rangle$ is the derivative coupling, $\dot{\mathbf{R}}$ is the nuclear velocity, and $S_{\mu\nu}$ is the overlap matrix.

We can rewrite Eq. (8.9) as

$$i\hbar \sum_{\nu} \sqrt{S}_{\mu\nu} \dot{C}_{\mathbf{k}\alpha}^{\nu}(t) = \sum_{\kappa\nu} \left(\tilde{H} - i\hbar \dot{\mathbf{R}} \cdot \tilde{\mathbf{X}} \right)_{\mu\kappa} \sqrt{S}_{\kappa\nu} C_{\mathbf{k}\alpha}^{\nu}(t), \quad (8.10)$$

where $\tilde{H}_{\mathbf{k}\mu\nu} \equiv \sum_{\sigma\kappa} \left(1/\sqrt{S} \right)_{\mu\sigma} H_{\mathbf{k}\sigma\kappa} \left(1/\sqrt{S} \right)_{\kappa\nu}$ and $\tilde{X}_{\mathbf{k}\mu\nu} \equiv \sum_{\sigma\kappa} \left(1/\sqrt{S} \right)_{\mu\sigma} X_{\mathbf{k}\sigma\kappa} \left(1/\sqrt{S} \right)_{\kappa\nu}$.

The coefficients can be calculated with the following expression:

$$C_{\mathbf{k}\alpha}^{\nu}(t + \Delta t) \approx \sum_{\mu\kappa} \left(\frac{1}{\sqrt{S}} \prod_{j=0}^{N_s-1} \left[1 + \frac{idt}{2\hbar} \tilde{H}_{\mathbf{k}}(t_{j+\frac{1}{2}}) \right]^{-1} \left[1 - \frac{idt}{2\hbar} \tilde{H}_{\mathbf{k}}(t_{j+\frac{1}{2}}) \right] \right)_{\nu\mu} \left(\sqrt{S} \right)_{\mu\kappa} C_{\mathbf{k}\alpha}^{\kappa}(t). \quad (8.11)$$

The required self-consistency between the KS Hamiltonian $H^{\text{KS}}(t + \Delta t)$ and the electron density $\rho(\mathbf{r}, \mathbf{r}'; t + \Delta t)$ at time $t + \Delta t$ is achieved iteratively starting from an initial guess for $H^{\text{KS}}(t + \Delta t)$. We propagate the wavefunctions using Eq. (8.5), calculate the electron density at time $t + \Delta t$ using Eq. (8.7), and recalculate the corresponding Hamiltonian $H^{\text{KS}}(t + \Delta t)$. This iterative procedure is repeated until a self-consistent solution is reached. We require the norm of the electron density difference across successive iterations to be smaller than a predetermined ϵ_{tol} value, which we set to 5×10^{-7} . Convergence is accelerated by Pulay's direct inversion of iterative space (DIIS) algorithm [392] in which the trial density matrix at the p th iteration step $D_{\mathbf{k}\mu\nu}^{(p)}(t)$ is set as the input vector. For conciseness, we hereafter

focus on electron dynamics in the fixed-nucleus approximation and neglect the derivative coupling term in Eq. (8.9).

8.2.2 Velocity gauge formulation

The velocity gauge formulation is used to correctly incorporate the effects of external electric fields with periodic boundary conditions. In this formulation, the kinetic energy, \mathcal{T} , in Eq. (8.2) becomes

$$\mathcal{T} = \frac{1}{2m_e} \left(\frac{\hbar}{i} \nabla - \frac{q_e}{c} \mathbf{A}(t) \right)^2, \quad (8.12)$$

where m_e is the electron mass of an electron, q_e is its charge, c is the speed of light, and $\mathbf{A}(t)$ is the electromagnetic vector potential. Adopting the long-wavelength approximation and assuming the spatial uniformity of the electric field, $\mathbf{A}(t)$ is the integration of the electric field strength $\mathbf{E}(t)$: $\mathbf{A}(t) = -c \int^t dt' \mathbf{E}(t')$. The wavefunctions in the velocity gauge, ψ^{vg} , and the length gauge, ψ^{lg} , are related by a formal gauge transformation:

$$\psi^{\text{vg}}(\mathbf{r}, t) = e^{i \frac{q_e}{\hbar c} \mathbf{A}(t) \cdot \mathbf{r}} \psi^{\text{lg}}(\mathbf{r}, t). \quad (8.13)$$

Non-local potentials appearing in the pseudopotential formulation are therefore transformed according to the expression

$$V_{\text{NL}}^{\text{vg}}(\mathbf{r}, \mathbf{r}') = e^{i \frac{q_e}{\hbar c} \mathbf{A}(t) \cdot \mathbf{r}} V_{\text{NL}}^{\text{lg}}(\mathbf{r}, \mathbf{r}') e^{-i \frac{q_e}{\hbar c} \mathbf{A}(t) \cdot \mathbf{r}'}. \quad (8.14)$$

PySCF uses Goedecker Teter Hutter (GTH) separable dual-space pseudopotentials [393]. The matrix elements of the transformed GTH pseudopotential in a Gaussian basis set representation can be evaluated either analytically or numerically in reciprocal space. Observables containing the differential operator, such as the electronic velocity, are modified

in a covariant manner as

$$\boldsymbol{\xi}^{\text{vg}}(t) = \sum_{\mathbf{k}\alpha} w_{\mathbf{k}\alpha} \left\langle \psi_{\mathbf{k}\alpha}(t) \left| \frac{1}{m_e} \left(\frac{\hbar}{i} \nabla - \frac{q_e}{c} \mathbf{A}(t) \right) + \frac{1}{i\hbar} [\hat{\mathbf{r}}, \hat{V}_{\text{NL}}^{\text{vg}}] \right| \psi_{\mathbf{k}\alpha}(t) \right\rangle, \quad (8.15)$$

where the commutator term reflects the non-commutativity of the dipole and non-local pseudopotential operators [379].

8.2.3 Optical absorption spectrum

In the length gauge calculation of finite systems (i.e., molecules), the time-dependent dipole momenta is induced by an instantaneous electric field pulse of the form $\mathbf{E}(t) = F_0 \mathbf{n} \delta(t)$, where \mathbf{n} is the polarization direction, F_0 the electric field amplitude, and $\delta(t)$ is the Dirac delta function [394]. From linear response theory, the system's response to an infinitesimally small external field, $F_0 \mathbf{n}$, is given by

$$\langle \mu^k(\mathbf{r}, \omega) \rangle = \sum_l \int d^3 \mathbf{r}' \chi_0^{kl}(\mathbf{r}, \mathbf{r}', \omega) n^l F_0, \quad (8.16)$$

where $\mu^k(\mathbf{r}, \omega)$ is the Fourier transform of the dipole moment in the k th spatial direction at coordinate \mathbf{r} , where χ_0 is the response function given by

$$\chi_0^{kl}(\mathbf{r}, \mathbf{r}', \omega) = \sum_a \frac{\langle 0 | \hat{\mu}^k(\mathbf{r}) | a \rangle \langle a | \hat{\mu}^l(\mathbf{r}') | 0 \rangle}{\hbar\omega - E_{a0} + i\hbar\eta} - \sum_b \frac{\langle 0 | \hat{\mu}^l(\mathbf{r}') | b \rangle \langle b | \hat{\mu}^k(\mathbf{r}) | 0 \rangle}{\hbar\omega + E_{b0} + i\hbar\eta}, \quad (8.17)$$

where $|a\rangle$ ($|0\rangle$) denotes an excited (ground) energy eigenstate with energy E_a (E_0), η the broadening parameter, and the excitation energies are given by $E_{a0} \equiv E_a - E_0$.

In our RT-TDDFT formalism, we are only interested in the spatially uniform component, and the oscillator strength can be computed as $\sum_{kl} n^k \alpha^{kl}(\omega) n^l$ where $\alpha^{kl}(\omega)$ is calculated as

$$\alpha^{kl}(\omega) \equiv -\frac{2\hbar\omega}{\pi} \text{Im} \iint d^3 \mathbf{r} d^3 \mathbf{r}' \chi_0^{kl}(\mathbf{r}, \mathbf{r}'; \omega) = \hbar\omega \sum_{\alpha} \langle 0 | \hat{\mu}^k | \alpha \rangle \langle \alpha | \hat{\mu}^l | 0 \rangle \delta(\hbar\omega - E_{\alpha 0}). \quad (8.18)$$

The oscillator strength satisfies the frequency sum rule. In the velocity gauge formulation, the input field transforms to a step function vector field given by

$$\mathbf{A}(t) = -cF_0\mathbf{n}\theta(t), \quad (8.19)$$

where $\theta(t)$ is the Heaviside function.

In condensed-phase calculation, a convenient observable is, instead of the dipole, the electronic velocity, whose response is given by

$$\chi_{vd}^{kl}(\omega) = \sum_a \frac{\langle 0 | [\hat{\mu}^k, H^{KS}] / i\hbar | a \rangle \langle a | \hat{\mu}^l | 0 \rangle}{\hbar\omega - E_{a0} + i\hbar\eta} - \sum_b \frac{\langle 0 | \hat{\mu}^l | b \rangle \langle b | [\hat{\mu}^k, H^{KS}] / i\hbar | 0 \rangle}{\hbar\omega + E_{b0} + i\hbar\eta}, \quad (8.20)$$

where the subscript vd in the left hand side indicates the dipole to velocity response. As such, Eq. (8.18) can be succinctly written as

$$\alpha^{kl}(\omega) = -\frac{2}{\pi} \text{Re} \chi_{vd}^{kl}(\omega), \quad (8.21)$$

In real-time simulation, we calculate the electronic velocity [the dipole (if the target system is an isolated molecule)] as a time-dependent observable, which is then Fourier-transformed to numerically calculate the response function $\chi_{vd}^{kl}(\omega)$ [$\chi^{kl}(\omega)$]. The formal expression Eq. (8.21) or (8.18) is then used to calculate the oscillator strength from the response function.

The dielectric function can be obtained from the expression [395]

$$\varepsilon(\omega) = 1 + \frac{4\pi i \sigma(\omega)}{\omega}, \quad (8.22)$$

where the dynamical conductivity $\sigma(\omega)$ is given by

$$\sigma(\omega) = \frac{qe}{\Omega_c} \int_{\Omega_c} d^3\mathbf{r} \int dt e^{i\omega t} \langle \Psi_t | \hat{n} \cdot \hat{\mathbf{j}}(\mathbf{r}) | \Psi_t \rangle / F_0, \quad (8.23)$$

where Ω_c is the cell volume, and $\hat{\mathbf{j}}(\mathbf{r})$ is the electronic current operator. The electronic current integrated over the unit cell is equivalent to the electronic velocity expectation value given in Eq. 8.15. In our numerical implementation, the time propagation is truncated to a finite range from $0 \leq t \leq T$, where T is set to 24 fs. The time-series of dipoles/electronic velocities are multiplied by a Gaussian function of the form $e^{-w^2 t^2/2}$ ($w = 0.1$ eV is used in this work) before a Fourier transformation is taken, which introduces a finite spectral width. The nuclear dynamics formalism for our theory is available in Appendix E.

8.3 Results

This section provides numerical results of our velocity-gauge RT-TDDFT implementation in the PySCF software package. Our calculations were carried out on the comet and expanse cluster in the San Diego Supercomputer Center. We first discuss optical absorption spectra calculations and examine both non-periodic molecules (Section 8.3.1) and periodic solids (Section 8.3.1) to validate and verify our implementation. We then discuss real-time electron dynamics between a $(\text{ZnO})_4$ molecular cluster and a periodic graphene sheet in Section 8.3.2 as an example of real-time charge transfer in strong external fields.

8.3.1 Optical Absorption Spectra

To benchmark and validate our velocity-gauge RT-TDDFT implementation, we calculated the optical absorption spectra of both molecular and condensed-phase systems. The real-time response to a vector field step-function (Eq. 8.19) was calculated with F_0 set to 0.001 in the atomic unit (a.u.). The time step, Δt , was set to 8.0 as, and the wavefunction

was propagated for 3,000 steps.

Optical absorption spectra of molecules

As a first test of our approach we compare our results for real-time electron dynamics of an isolated CH_4 molecule with an all-electron basis set. Figure 8.1 shows the absorption spectrum of CH_4 obtained with the Vosko-Wilk-Nusair (VWN) local density approximation [396] with the aug-cc-pvtz [397] basis set (LDA/aug-cc-pvtz), where the continuous spectrum were calculated using Eq. 8.21. The molecular geometry was optimized using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) variant of the Davidon-Fletcher-Powell minimization algorithm [398].

Figure 8.1(a) compares our velocity-gauge RT-TDDFT optical absorption results against a standard linear-response (LR) TDDFT calculation. As expected, the low-energy optical spectrum obtained using RT-TDDFT is essentially equivalent to that obtained with LR-TDDFT, which validates our velocity-gauge implementation.

Our next validation test is shown in Fig. 8.1(b), which compares the oscillator strength results obtained from our length-gauge and velocity-gauge RT-TDDFT calculations, using Eqs. 8.18 and 8.21, respectively. We obtained the length-gauge results from the system's response to a delta-function electric field strength of $F_0 = 0.001$ a.u., whereas the velocity gauge results were obtained using the Heaviside function vector field Eq. (8.19) with the same field strength, $F_0 = 0.001$ a.u. The two results show nearly perfect agreement, indicating equivalence of the two gauge choices in the linear response regime.

Finally, Fig. 8.1(c) compares the same oscillator strength distribution calculated from the dipoles and electronic velocities (see Eqs. (8.18) and (8.21), respectively). Except

for the fictitious peak around $\omega \approx 0$ appearing in the velocity spectrum, these two results show good agreement. We attribute the appearance of the fictitious peak to the nature of the velocity spectrum in Eq. (8.21) [399]. The consistency between the results obtained from the length and velocity gauges validate our RT-TDDFT implementation.

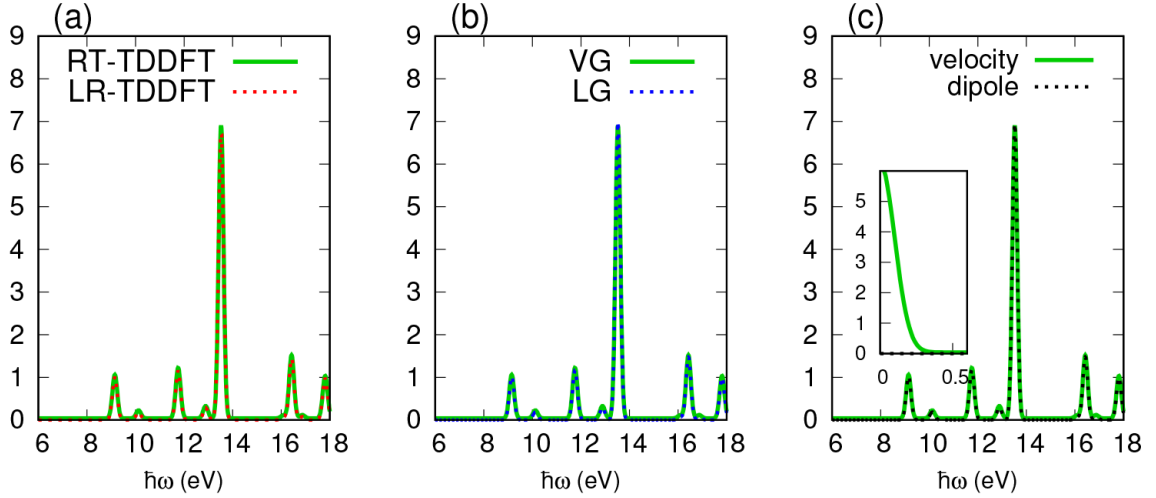


Figure 8.1: Optical absorption spectrum of the CH₄ molecule. Panel (a) shows the oscillator strength distribution of the CH₄ molecule obtained with the LDA xc potential and aug-cc-pVTZ [397] basis set. The solid green line shows the RT-TDDFT result, whereas the dotted red line shows the LR-TDDFT result. Panel (b) shows the RT-TDDFT spectra calculated using the velocity gauge (VG, solid green line) and the length gauge (LG, dotted blue line) formulations. Panel (c) compares the velocity (solid green line) and dipole (dotted black line) spectra. The inset shows the behavior of the spectra around $\omega = 0$.

Optical absorption spectra of solids

We next investigate the optical response of fully extended periodic systems. As a prototypical example, we examine a two-dimensional hexagonal boron nitride (*h*-BN) layer with a lattice constant of $a = 2.504 \text{ \AA}$ [400] and a vacuum layer of 20 \AA . We used a $15 \times 15 \times 1$ Monkhorst-Pack [401] mesh for the Brillouin zone sampling, and the VWN [396] formulation of the LDA [388] xc functional. The GTH separable dual-space pseudopotential

tial [393] and the double-zeta polarized basis set adapted for DFT calculations with the GTH pseudopotential (gth-dzvp) [380] was used. To enhance the computational efficiency of our calculations, the Gaussian and Plan-Wave density-fitting scheme [385], termed the fast Fourier Transformation (FFT) density-fitting in PySCF program, was applied. In calculating the dynamical conductivity $\sigma(\omega)$, we followed the procedure in Ref. [17] and introduced a damping factor of the form $e^{-\gamma t}$ with $\gamma = 0.5 \text{ eV}/\hbar$, to reduce numerical error due to the finite time propagation.

To further validate our implementation, we compute the dynamical conductivity of h -BN and compare it with Ref. [17]. Figure 8.2 shows the frequency-dependent conductivity, $\sigma(\omega)$, calculated from the real-time response of the system to a weak step-function vector field Eq. (8.19) with $F_0 = 0.001 \text{ a.u.}$ and \mathbf{n} in the x direction. Our calculation shows that our computed spectrum agrees relatively well with Ref. [17], where we attribute the minor differences to the basis set and \mathbf{k} -point sampling used in our calculations (Ref. [17] used a real-space grid and sampled over 32×32 \mathbf{k} -points). As such, the agreement between our RT-TDDFT calculations with the LDA result of Ref. [17] validates our RT-TDDFT PySCF implementation for periodic systems.

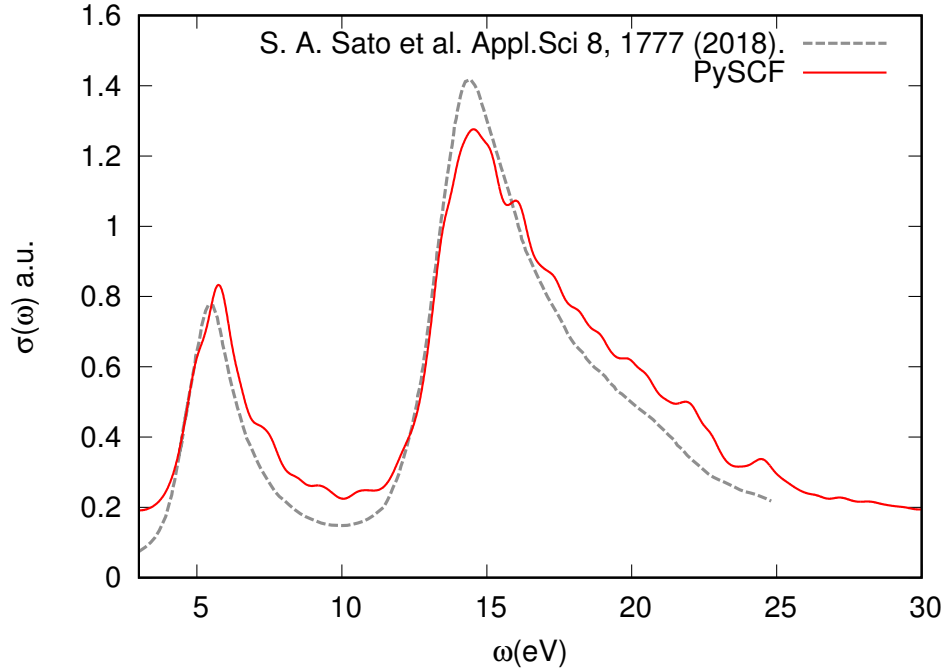


Figure 8.2: Dynamical conductivity $\sigma(\omega)$ of monolayer *h*-BN. The solid red line was obtained from our RT-TDDFT PySCF implementation, and the dashed gray line was obtained from Ref. [17].

8.3.2 Real-time dynamics

With our PySCF RT-TDDFT implementation validated, we next discuss real-time dynamics in strong external fields. This section examines laser-induced charge transfer dynamics for a two-dimensional periodic system whose unit cell consists of a $(\text{ZnO})_4$ molecular cluster and a 4×4 monolayer graphene sheet, which is a prototypical example of excited-state dynamics between a (non-periodic) molecule and an extended periodic system [9].

Figure 8.3 shows our simulation cell, which consists of a 4×4 supercell of graphene and a $(\text{ZnO})_4$ cluster. The figure was created using the Avogadro software [402]. The graphene sheet is held fixed with an inter-carbon distance of $r_{\text{CC}} = 1.42 \text{ \AA}$, whereas the nu-

clear configuration of $(\text{ZnO})_4$ cluster was first optimized by density functional tight binding (DFTB) calculations using the dftb+ toolkit [403] with the 3ob parameter set [404, 405], then shifted towards the graphene sheet by 1 Å. In this geometry, the Zn atom closest to the graphene sheet was 1.71 Å above the sheet. Our calculations utilized a single \mathbf{k} -point. To reproduce the gapless excitation spectrum of graphene, we shifted the \mathbf{k} -point origin by $\mathbf{K}_0 \equiv (\frac{4\pi}{3a}, 0, 0)$, where a ($= 2.4596 \text{ Å} = 4.648 \text{ a.u.}$) is the graphene sheets lattice constant. We applied a Fermi-Dirac distribution function corresponding to $T = 1000 \text{ K}$ for smearing the metallic Fermi surface. We used the Perdew, Burke, and Ernzerhof [390] GGA xc functional and the single-zeta molecularly optimized (molo)pt basis set (gth-szv-molo)pt-sr) [380] for our RT-TDDFT calculation. We applied an ultrashort trapezoidal pulse of width $\tau_p = 20 \text{ fs}$, with an ascending/descending time of 5 fs, shown in Fig. 8.4(a)'s inset. The wavelength and peak intensity of the pulse was set to $\lambda = 360 \text{ nm}$ and $I = 4.0 \times 10^{13} \text{ W/cm}^2$, respectively. We chose $\lambda = 360 \text{ nm}$ as our excitation energy since the local density of states of this hybrid system indicates that a UV photon with this wavelength would cause an excitation from the valence to the conduction band.

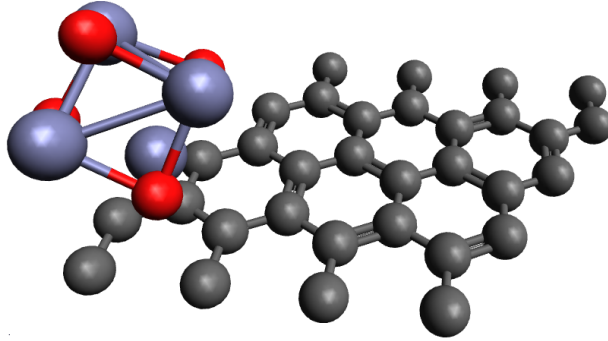


Figure 8.3: Geometry of our simulation cell consisting of a $(\text{ZnO})_4$ molecular cluster and a 4×4 periodic graphene sheet. The purple, red, and dark gray spheres represent zinc, oxygen, and carbon atoms, respectively.

Figure 8.4 shows the electronic charges time evolution on the $(\text{ZnO})_4$ cluster. We considered a dividing plane 1.0 \AA above the graphene sheet, $z = z_0$ with $z_0 = 1.0 \text{ \AA}$, and defined the total electric charge Q above the dividing plane as $Q(t) \equiv \int_{z \geq z_0} dz \iint dx dy \rho(\mathbf{r}, t)$, which represents the total number of electrons on the $(\text{ZnO})_4$ cluster. Panel (a) plots the time derivative dQ/dt as a function of time. The inset in panel (a) plots the electric field of the laser pulse, which confirms that the oscillation of the electric charge on the $(\text{ZnO})_4$ cluster is directly correlated with the oscillations of the laser field. To confirm these effects, Fig. 8.4(b) plots the summation of the Mulliken charges of Zn, C, and O represented by the blue, green, and red lines, respectively. To show the spatial distribution of the electronic density at an intermediate time $t = 11.22 \text{ fs}$, Fig. 8.5 plots the charge-density difference, $\Delta\rho(\mathbf{r}, t) \equiv \rho(\mathbf{r}, t) - \rho(\mathbf{r}, 0)$, which shows dynamic charge transfer between the $(\text{ZnO})_4$ cluster and periodic graphene sheet.

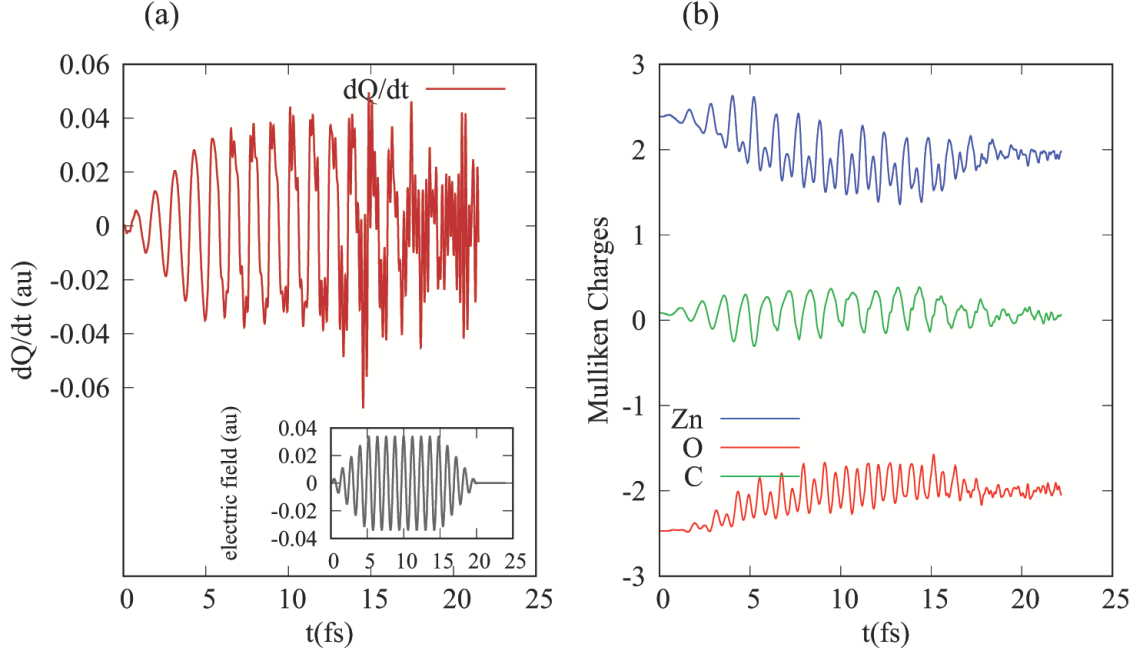


Figure 8.4: (a) Time evolution of the time-derivative of the total electronic charge Q above the $z = 1.0 \text{ \AA}$ dividing plane. The inset shows the electric field of the applied laser pulse. (b) Mulliken charges as a function of time. The blue, green, and red lines show the total summation of the Mulliken charges on the zinc, carbon, and oxygen atoms, respectively.

8.4 Discussion and Conclusion

In this work, we have implemented a new, Gaussian-based, velocity gauge RT-TDDFT formalism in the open-source PySCF software package. Our calculations in Section 8.3.1 show a close agreement between LR- and RT-TDDFT, different gauge choices, and oscillator strength distributions obtained from our time-dependent dipoles and electronic velocities. In addition, we also showed that our RT-TDDFT calculation of the dynamical conductivity, $\sigma(\omega)$, of a periodic h -BN system agrees with previously published benchmarks. As such, these detailed consistency tests validate our new RT-TDDFT implementation for solids in the open-source PySCF software package.

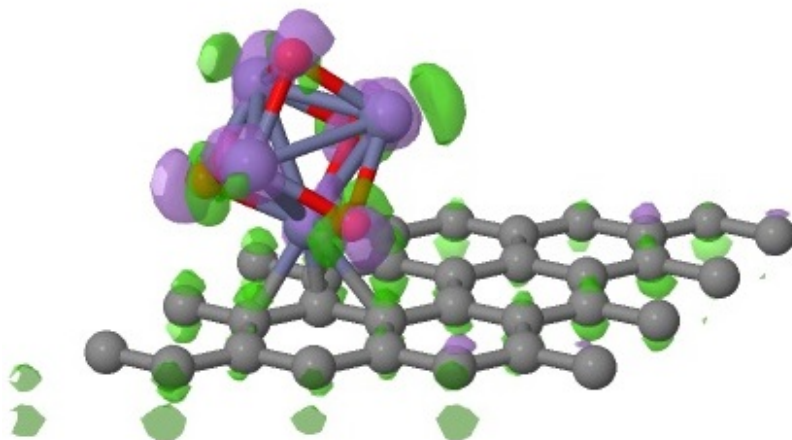


Figure 8.5: Dynamic charge transfer in a system made up of a $(\text{ZnO})_4$ molecular cluster and a periodic graphene sheet. The charge density difference was calculated between $t = 0$ and 11.22 fs, and the green and purple isosurfaces are plotted for $\Delta\rho = \pm 0.04$ a.u.

With these validation tests completed, we then examined real-time, laser-induced charge-transfer dynamics for the $(\text{ZnO})_4$ molecular cluster and graphene system. Our femtosecond laser-induced dynamics calculations in this system demonstrates that our implementation can be used to probe real-time electron dynamics in periodic systems, which is a new capability not previously available in the PySCF software package.

In this work, we have implemented a Gaussian-based velocity gauge RT-TDDFT on the PySCF program suite, in order to provide a platform to where future real-time formalisms can be easily extended. As an example, we validate our implementation and investigate ultrafast modification of optical properties in both, molecular and extended systems. Our first test, on an isolated CH_4 molecule, shows our velocity-gauge oscillator strength results agreeing well with the LR-TDDFT, and length-gauge TDDFT results. We proceeded to check the dynamical conductivity results on the 2D extended system h -BN, where our results closely matched with the results in Ref. [17]. The plots discrepancies were

due to the differences in basis set and k-point sampling used in our calculations. For our final check, we applied our program to field-induced dynamics on a graphene surface, which reproduces field-induced charge transfer between a metal oxide and a graphene sheet. Our simulation successfully showed charge transfer to the graphene surface. We expect future RT-TDDFT implementations to be extended to broader fields and become part of the standard tools for physicists, chemists, and materials scientists interested in laser-matter interactions.

Chapter 9

Indirect but Efficient: Laser-Excited Electrons can drive Ultrafast Polarization Switching in Ferroelectric Materials

This chapter describes a mechanism that enables the reversible switching of the ferroelectric polarization with optical pulses. This chapter resulted from a collaboration with Dr. Chao Lian, who developed and implemented the real-time time-dependent density functional theory (RT-TDDFT). The work presented was to introduce and teach me to analyze RT-TDDFT results. The article was published in *the Journal of Physical Chemistry Letters* [406], where the charge dynamics is also included in the complete analysis.

9.1 Introduction

Ferroelectric (FE) materials are characterized by an intrinsic spontaneous electric polarization that can be further harnessed for next-generation electronic and energy-harvesting materials. For example, tuning the FE polarization can vary the tunneling resistance over several orders of magnitude [407], enabling technological advancements such as non-volatile memory in digital electronic devices, [408] memristors, [409] and integrated neuromorphic networks [410, 411]. In addition, by enabling a steady-state photocurrent, FE polarization can substantially increase light-harvesting efficiency, particularly in hybrid organic-inorganic halide perovskite solar cells [412–423]. Finally, the variation in FE polarization on surfaces can also dramatically change the adsorption energetics in catalytic systems and could be further harnessed to enable other polarization-dependent surface mechanisms [424, 425].

These applications are intrinsically associated with FE polarization and can be further enhanced by tuning and controlling this intrinsic material property. Polarization switching is typically accomplished through a static electric field; however, the switching time is relatively sluggish (on the order of nanoseconds) due to the slow re-crystallization time of most materials (typically hundreds of picoseconds) [426, 427]. To accelerate this switching process, significant research has focused on enabling ultrafast polarization switching via optical processes [428, 429]. Terahertz (THz) sources provide ultrafast pulsed electric fields expected to reverse the polarization via the same mechanism as a static electric field. However, even with state-of-the-art THz sources, a single THz pulse does not possess enough strength to switch this polarization [430–436]. To circumvent THz sources, Subedi

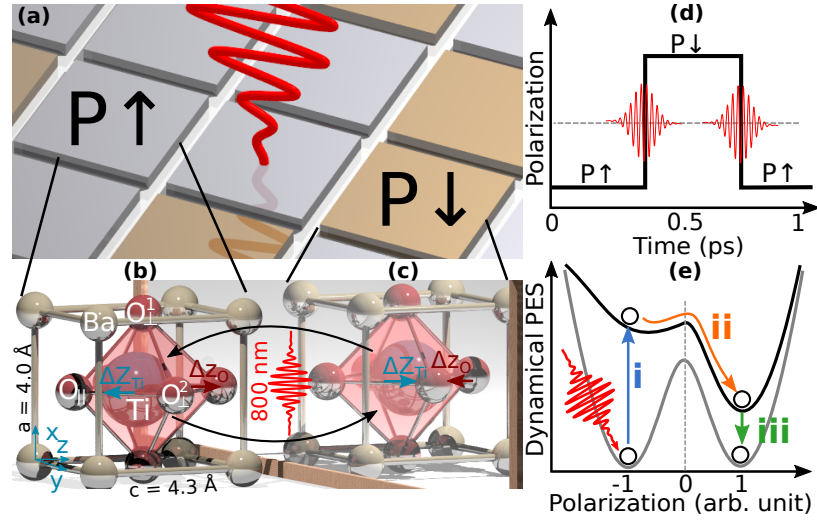


Figure 9.1: (a) Schematic diagram of a FE array. The bright and dark blocks denote the (b) up-polarized (c) and down-polarized structures of BaTiO₃, respectively. (d) Diagram of ultrafast optical polarization switching as a function of time. The red lines denote two sequential identical laser pulses. (e) Diagram of laser-induced modification of the dynamical potential energy surface (PES). The gray (black) line represents the ground- and excited-state PES, respectively.

recently proposed that activating an infrared-active phonon mode can also induce polarization switching effects [437]. Later, Mankowsky et al. [438] conducted experiments on LiNbO₃ but only observed a temporary and partial polarization switching. While other researchers have reported related phenomena such as ultrafast domain wall movements [439–443], photoinduced depolarization [444–447], and coherent ionic movement [448–450] ultrafast polarization switching in these FE materials has not been entirely successful. Ideally, one would like a more efficient mechanism, such as that shown in Fig. 9.1(a) and (d), in which a readily-available laser source could be utilized to switch the polarization with one pulse and reversibly switched back with a second identical pulse.

In this Letter, we propose a new mechanism that enables this efficient switching of the polarization in FE materials. By accounting for the detailed interactions between

the electrons, nuclei, and electromagnetic field, our real-time time-dependent density functional theory (RT-TDDFT) simulations show the ultrafast FE polarization switching in BaTiO₃ occurs within 200 femtoseconds (fs) after the material is excited by an 800-nm laser pulse. This polarization switching commences when the laser pulse pumps electrons from 2p orbitals in oxygen to the 3d orbitals in titanium. This dynamic excitation process yields desirable and highly-directional forces on the Ti and O atoms that are (1) always opposite to the original FE displacements and (2) can be harnessed to switch the material’s polarization consistently. By laser-tuning the dynamical potential energy surface, we show that this new mechanism can switch the FE polarization in both directions using identical pulses from experimentally-available laser sources.

9.2 Computational Details

We use the time-dependent *ab initio* package, (TDAP), [451–453] for our RT-TDDFT calculations [454–456], where the wavefunctions and charge densities are obtained from the Quantum Espresso [457, 458] software package. We used the projector augmented-waves method (PAW) [339] and the Perdew-Burke-Ernzerhof (PBE) exchange-correlation (XC) functional [340] in both our DFT and RT-TDDFT calculations. Pseudopotentials were generated using the *pslibrary* [459] software package. The plane-wave energy cutoff was set to 55 Ry, and the Brillouin zone was sampled using a Monkhorst-Pack scheme with an $8 \times 8 \times 8$ k -point mesh for the unit cell and a $2 \times 2 \times 2$ mesh for a $3 \times 3 \times 3$ supercell. To reproduce the experimental band gap, a scissor correction of 1.65 eV was added to both the ground state and time-dependent calculations. The electronic timestep, δt , was

set to 1.94×10^{-4} fs, and the ionic timestep, Δt , was 0.194 fs. The Gaussian-type laser pulse utilized in our study is given by $\mathbf{E}(t) = \mathbf{E}_0 \cos(\omega t) \exp\left[-\frac{(t-t_0)^2}{2\sigma^2}\right]$, where $|\mathbf{E}_0|$ is the electric field amplitude, $\omega = 1.55$ eV is the laser frequency, and $t_0 = 50$ fs is the temporal location of the electric field peak. The laser pulse is linearly polarized along the x direction, perpendicular to the ferroelectric polarization. The crystal orbital Hamilton population (COHP) analysis was calculated with the Lobster [460–462] software package.

9.3 Results and Discussion

Table 9.1: Symmetric positions α_I^{sym} ($\alpha = x, y, z$), FE displacements Δz_I and Born effective charge of the I -th atom. We categorize these three O atoms into two types: one O_{\parallel} atom that is parallel to the polarization and two O_{\perp} atoms that are perpendicular to the polarization.

Atom I	x_I^{sym}	y_I^{sym}	z_I^{sym}	Δz (\AA)	μ^* ($ e $)
Ba	0.0	0.0	0.0	0.00	2.73
Ti	0.5a	0.5a	0.5	0.08	7.17
O_{\perp}^1	0.0	0.5a	0.5c	-0.12	-2.02
O_{\perp}^2	0.5a	0.0	0.5c	-0.12	-2.02
O_{\parallel}	0.5a	0.5a	0.0	0.21	-5.74

We first optimized the structure of BaTiO_3 , which is shown in Fig. 9.1(b) and Table 9.1. The lattice is tetragonal with $a = b = 4.00$ \AA and $c/a = 1.075$. The ferroelectric properties of BaTiO_3 originate from the slight distortion of the Ti-O octahedron: the Ti atom deviates from the body-centered position along the z direction and the O atoms deviate from the face-centered positions along the $-z$ direction. These FE displacements

are characterized as the displacements from the symmetric positions $\Delta\alpha_I = \alpha_I^{\text{FE}} - \alpha_I^{\text{sym}}$, where α_I^{FE} and α_I^{sym} are the positions of atom I ($I = \text{Ti}, \text{Ba}, \text{O}_\perp$, and O_\parallel) along the $\alpha = x, y$, and z direction in the FE and symmetric phase, respectively. We calculate the FE polarization accordingly as $P_0 = \frac{1}{V} \sum_I \mu_I^* \Delta z_I = 2.01 \text{ e}/\text{\AA}^2 = 0.50 \text{ C/m}^2$, where $V = a \times a \times c = 68.8 \text{ \AA}^3$ is the volume, Δz_I is the ionic FE displacement, and μ_I^* is the Born effective charge, as shown in Fig. 9.1(b) and Table 9.1. Consistent with previous studies, the semilocal exchange-correlation functionals slightly overestimate the c/a ratio and the static polarization, as discussed in [463, 464].

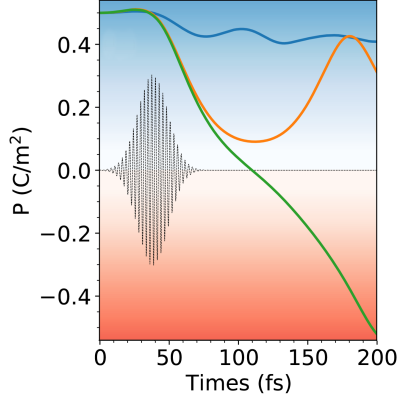


Figure 9.2(a) FE polarization P

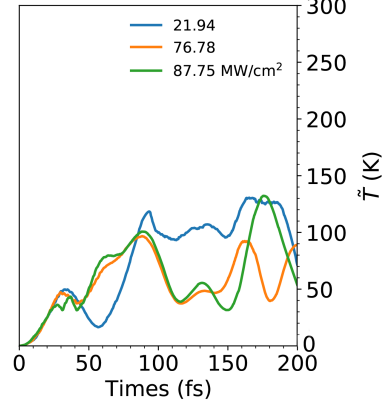


Figure 9.2(b) Effective temperature \tilde{T}

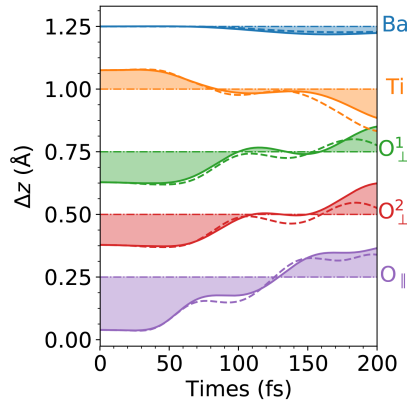


Figure 9.2(c) FE Displacement $\alpha = z$

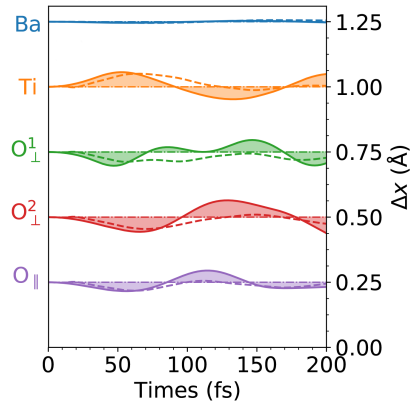


Figure 9.2(d) FE Displacement $\alpha = x$

Figure 9.2: (a) FE polarization P and (b) Effective temperature \tilde{T} as a function of time. FE displacements as a function of time for (c) $\alpha = z$ (d) $\alpha = x$. Solid lines denote the FE displacements calculated in the $1 \times 1 \times 1$ cell, and dashed lines denote the average FE displacements in the $3 \times 3 \times 3$ cell.

We apply laser pulses having a wavelength of $\lambda = 800$ nm, a polarization perpendicular to the FE polarization direction, and a duration lasting from $t_0 = 0$ fs to $t_f = 100$ fs. We characterize and monitor the laser-induced dynamics by calculating the dynamical polarization

$$P(t) = \frac{1}{V} \sum_I \mu_I^* \Delta z_I(t). \quad (9.1)$$

Figure 9.2(a) shows various laser fluence results, $F = 21.94, 76.78,$ and 87.75 MW/cm². We

identify three types of lattice dynamics: **(1)** at the lowest fluence of $F = 21.94 \text{ MW/cm}^2$, $P(t)$ slightly decreases and oscillates around P_0 . **(2)** when F increases to a critical fluence of 76.78 MW/cm^2 , $P(t)$ substantially decreases to around 0 within 100 fs, while $P(t)$ recovers to P_0 in the following 100 fs. **(3)** When F further increases to 87.75 MW/cm^2 , we observe a switching of the polarization. $P(t)$ continuously decreases past 0 until it reaches its maximum polarization in the opposite direction.

In contrast to other experimental strategies discussed previously, this laser-induced polarization change is a non-thermal process. To clearly demonstrate this, we plot the effective temperature $\tilde{T}(t) = E_{\text{kin}}(t)/k_B$ as a function of time, where E_{kin} is the kinetic energy of the ions and k_B is the Boltzmann constant. As shown in Fig. 9.2(b), $\tilde{T}(t)$ is much lower than room temperature throughout our entire dynamics simulations regardless of the laser fluence. In addition, $\tilde{T}(t)$ is different from the thermodynamic equilibrium temperature produced by random ionic movements. Instead, ionic movements are highly directional in the laser-induced polarization switching process. As shown in Fig. 9.2(c), the laser radiation triggers a set of FE displacements, $\Delta z_I(t)$, that are nearly anti-parallel with the original values, $\Delta z_I(t_0)$. $\Delta z_I(t)$ of the oxygen atoms O_{\perp} and O_{\parallel} change sign from negative to positive, while the titanium atom moves in the opposite direction from positive to negative. Considering that the oxygen and titanium atoms have effective Born charges with opposite signs (Table 9.1), the movements of all the atoms contribute consistently to the polarization switching. Moreover, along the directions perpendicular to the polarization, the displacement, $\Delta x_I(t)$, slightly oscillates around the equilibrium position, as shown in Fig. 9.2(d). This leads to an abnormal fluence-dependence of the effective temperature

$\tilde{T}(t)$, which slightly decreases as the fluence increases. Although the laser pulse induces larger ionic forces at higher fluence, the system needs to overcome the small energy barrier to reverse the polarization, as shown in Fig. 9.1(e), which consumes the kinetic energy and decreases the effective temperature. In comparison, with a lower influence, the ions can not overcome the energy barrier, resulting in a higher effective temperature for the oscillation. Thus, the highly directional FE movements predominantly contribute to the relatively low effective temperature $\tilde{T}(t)$, indicating a unique laser-induced non-thermal mechanism.

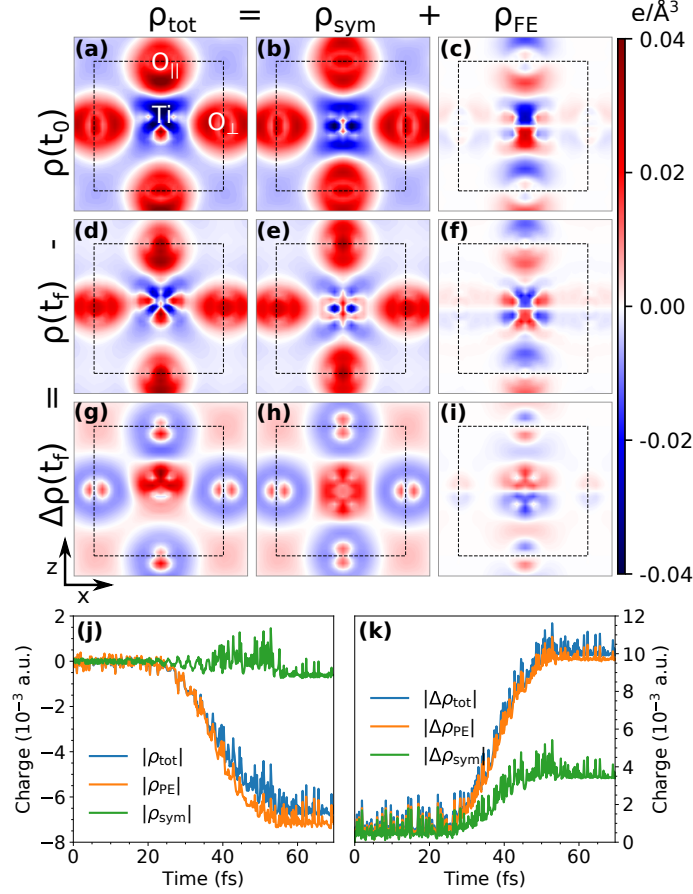


Figure 9.3: Two-dimensional contour plots of the charge densities at the plane $y = a/2$ for (a) $\rho_{\text{tot}}(x, z, t_0)$, (b) $\rho_{\text{sym}}(x, z, t_0)$, (c) $\rho_{\text{FE}}(x, z, t_0)$, (d) $\rho_{\text{tot}}(x, z, t_f)$, (e) $\rho_{\text{sym}}(x, z, t_f)$, (f) $\rho_{\text{FE}}(x, z, t_f)$, (g) $\Delta\rho_{\text{tot}}(x, z, t_f)$, (h) $\Delta\rho_{\text{sym}}(x, z, t_f)$, and (i) $\Delta\rho_{\text{FE}}(x, z, t_f)$, where t_0 and t_f are the start and end times of the laser pulse, respectively. Panels (j) and (k) depict the bonding and anti-bonding charge, respectively, as a function of time. For ease of comparison, the initial values of ρ_{tot} , ρ_{sym} , and ρ_{FE} in (j) are downshifted by 20.7×10^{-3} , 20.0×10^{-3} , and 5.70×10^{-3} a.u., respectively.

Coherence is another distinguishable feature of our polarization-switching mechanism in which the laser pulse drives the movement of the ions in a synchronized fashion. To verify that this coherence is not an artifact of size effects due to using a single unit cell, we performed the TDDFT calculations using a $3 \times 3 \times 3$ supercell for comparison. We analyzed the average laser-induced displacements as a function of time, $\langle \Delta z_I(t) \rangle = \sum_{s=1}^{27} \Delta z_I^s(t)$,

where s denotes the index of the unit cell. As shown in Fig. 9.2(c) and (d), $\langle \Delta z_I(t) \rangle$ in the $3 \times 3 \times 3$ supercell case is almost identical to $\Delta z_I(t)$ in the $1 \times 1 \times 1$ unit cell case, which indicates that the unified switching of the FE displacements occurs over all the laser-irradiated area. Consequently, no nucleation and growth of oppositely polarized domains are needed in this polarization switching mechanism, accelerating this switching process to finish within 200 fs.

We illustrate the atomistic mechanisms of the polarization switching by analyzing the carrier dynamics: i.e., we monitor the differential charge density $\rho_{\text{tot}}(x, y, z, t) = \rho_{\text{chg}}(x, y, z, t) - \rho_{\text{atom}}(x, y, z, t)$ as a function of time, where ρ_{chg} is the charge density, and ρ_{atom} is the superposition of the atomic charge densities. Thus, $\rho_{\text{tot}}(x, y, z, t)$ captures the spatial distribution of the bonding (+) and antibonding (-) densities. The FE polarization can be treated as the asymmetric disturbance of the symmetric phase, which is characterized by the centrosymmetric charge order. Thus, we divide the total charge $\rho_{\text{tot}}(x, y, z, t)$ into two parts, $\rho_{\text{sym}}(x, y, z)$ and $\rho_{\text{FE}}(x, y, z)$, where $\rho_{\text{sym}}(x, y, z) = [\rho_{\text{tot}}(x, y, c - z) + \rho_{\text{tot}}(x, y, z)]/2$, and $\rho_{\text{FE}}(x, y, z) = \rho_{\text{tot}}(x, y, z) - \rho_{\text{sym}}(x, y, z)$. Obviously, $\rho_{\text{sym}}(x, y, \frac{1}{2}c - z) = \rho_{\text{sym}}(x, y, \frac{1}{2}c + z)$ is the centrosymmetric part of the total charge that characterizes the original symmetric order, whereas $\rho_{\text{FE}}(x, y, z)$ is the asymmetric part that characterizes the FE order.

Figure 9.3(a)-(i) depicts a two-dimensional contour of the charge density, $\rho_i(x, z, t)$ ($i = \text{tot}, \text{sym}, \text{and FE}$), at the Ti-O surface $y = a/2$. Comparing Fig. 9.3(a)-(c), we find that $\rho_{\text{sym}}(x, z, t_0)$ is the dominant component of $\rho_{\text{tot}}(x, z, t_0)$ even when FE effects are included, whereas $\rho_{\text{FE}}(x, z, t_0)$ is localized around the Ti positions. As such, these figures show that an

electron transfers from a Ti 3d orbital to an O 2p orbital, which triggers a pseudo-Jahn-Teller effect (PJTE) [465] and stabilizes the FE phase. By comparing the laser-induced charge density difference, $\Delta\rho_i(x, z, t) = \rho_i(x, z, t) - \rho_i(x, z, t_0)$ ($i = \text{tot, sym, FE}$), with the ground state charge density $\rho_i(x, z, t_0)$, we find that: (1) the laser induces an electron transfer from a bonding to an anti-bonding area, where $\Delta\rho_{\text{tot}}(x, z, t_f)$ [Fig. 9.3(g)] is opposite to that of the ground-state bonding charge $\rho_{\text{tot}}(x, z, t_0)$ [Fig. 9.3(a)]; (2) the majority of the induced charge has a centrosymmetric periodicity, i.e. $\Delta\rho_{\text{sym}}(x, z, t_f)$ [Fig. 9.3(h)] dominates the induced total charge $\Delta\rho_{\text{tot}}(x, z, t_f)$ [Fig. 9.3(g)]; (3) The most important feature is that the induced FE charge density $\Delta\rho_{\text{FE}}(x, z, t_f)$ [Fig. 9.3(i)] is opposite to that of the original $\rho_{\text{FE}}(x, z, t_0)$ [Fig. 9.3(c)], indicating a decrease in the FE order. We quantitatively evaluate the laser-induced change in the charge density by analyzing the integrated charges $Q_i(t) = \int |\rho_i(\mathbf{r}, t)| d\mathbf{r}$ and $C_i(t) = \int |\rho_i(\mathbf{r}, t) - \rho_i(\mathbf{r}, t_0)| d\mathbf{r} = \int |\Delta\rho_i(\mathbf{r}, t)| d\mathbf{r}$, where $i = \text{tot, sym, and FE}$. The former characterizes the change in the bonding strength and the latter denotes the weakening of bonds, respectively. As shown in Fig. 9.3(j), the percentage decrease in the bond strength, $[Q_i(t_f) - Q_i(t_0)]/Q_i(t_0)$, is 32.45%, 35.92%, and 10.77% for $i = \text{tot, sym, and FE}$, respectively. Accordingly, $C_i(t)$ increases, with a ratio of $C_{\text{tot}}(t_f) : C_{\text{sym}}(t_f) : C_{\text{FE}}(t_f) = 1 : 0.97 : 0.27$, which is nearly the same as the ratio of the initial bonding charges $Q_{\text{tot}}(t_0) : Q_{\text{sym}}(t_0) : Q_{\text{FE}}(t_0) = 1 : 0.97 : 0.34$. Thus, the laser-induced bonding-antibonding transfer is nearly homogeneous, lowering both the centrosymmetric and FE order proportionally. Since the decrease in Q_{sym} affects all chemical bonds homogeneously, the overall effect of the decrease in Q_{tot} is to lower the stability of the FE polarization. Thus, the laser-induced charge-density change, $\Delta\rho_{\text{tot}}(x, z, t)$ is always opposite to the original FE

charge order and, therefore, weakens the original ionic bonding and PJTE, creating the unsymmetrical dynamical PES shown in Fig. 9.1(e). For the electronic structure dynamics, please refer to the complete publication [406].

9.4 Conclusions

Based on our photocarrier dynamic analysis of BaTiO₃, we have uncovered an electronically-driven polarization switching mechanism different from the conventional ionic-driven mechanism used in existing THz experiments. As shown in Fig. 9.1(e), our mechanism is based on a laser-tuned dynamical potential energy surface (PES), which can be generated with the time-dependent energy $E(t)$ and the time-dependent polarization $P(t)$ obtained in the TDDFT calculation as $E[P(t)]$. Since the dynamical PES is time-dependent, we only show the diagram to simplify the illustration. The laser radiation transforms the ground state PES into a dynamical PES by first exciting the material and automatically driving the ions to the opposite polarization. This PES-based mechanism encompasses three steps: **(1)** the laser pulse raises the PES of BaTiO₃ by pumping electrons from the 2p orbitals of oxygen to the 3d orbitals of titanium, inducing an anti-FE charge order [Fig. 9.3(i)]; **(2)** this dynamic excitation process yields desirable and highly-directional forces on the Ti and O atoms that are always opposite to the original FE displacements, transforming the lattice into a structure with an opposite FE polarization along the non-equilibrium TD-PES within 200 fs [Fig. 9.2(a)]; **(3)** the photo-excited system relaxes to the ground state in conjunction with the recombination of the photo-carriers (which occurs beyond our simulation time). We speculate that this dynamical-PES-based polarization switching is ubiquitous in

all the FE materials; therefore, the exact mechanism can be used to manipulate the polarization in other FE perovskite materials such as PbTiO_3 and LiNbO_3 , multiferroic materials, and even recently-discovered two-dimensional SnTe-based FE materials [466]. Most importantly, this new mechanism enables the switching of the FE polarization in both directions, with identical pulses produced from experimentally-available 800-nm laser sources. These findings provide a new mechanistic understanding of electronically-driven FE dynamics via laser-tuning the dynamical potential energy surface, which can accelerate the design of more efficient, ultrafast FE devices.

Chapter 10

Photo-induced Degradation of PFASs: Excited-State Mechanism from RT-TDDFT

This chapter explores the study of photo-induced degradation mechanisms of perfluorooctanoic acid (PFOA) contaminants by applying the real-time time-dependent density functional theory (RT-TDDFT). The majority of the work in this chapter resulted from a collaboration with Dr. Sharma S.R.K.C. Yamijala, Dr. Ravindra Shinde, and Dr. Kota Hanasaki. The work presented below extends Dr. Yamijala's previous work [467]. The results shown were published as an article in *the Journal of Hazardous Materials* [468], where perfluorooctanesulfonic acid (PFOS) is also analyzed. The molecular dynamics presented were only possible thanks to Dr. Yamijala and Dr. Hanasaki.

10.1 Introduction

Per- and polyfluoroalkyl substances (PFASs) are artificially-made compounds with strong carbon-fluorine (C–F) bonds that endow them with exceptional chemical stability. Because of their intrinsic longevity, these compounds have been used in various technologies, including non-stick coatings, water-resistant membranes, and fire-resistant foams [469–471]. However, the characteristics that enable this stability also pose severe health hazards since their presence in drinking-water sources is toxic and carcinogenic to humans [469, 472]. In particular, the intrinsic strength of the C–F bond in PFASs prevents most organisms from decomposing these persistent contaminants [473], which further exacerbates their bioaccumulation and toxicity. Moreover, due to their strong chemical stability, conventional treatments (such as chemical oxidation methods) are less effective in degrading PFASs, making them extremely difficult to remove [474].

Because of their resistance to conventional chemical treatments, immense interest has recently focused on photo-induced processes for directly decomposing PFAS contaminants [475–479]. In contrast to conventional filtration techniques that merely remove PFASs (which still require subsequent treatment after filtration), very recent studies have suggested that PFAS degradation can be accelerated with electromagnetic/optical fields, such as those used in **photocatalysis** or commercially available laser sources [475–480]. While these recent findings hold immense promise for directly treating PFASs, the exact mechanisms in these degradation processes remain unknown [476], and a guided path for rationally identifying photoactive materials and experimental conditions remains elusive.

To shed crucial mechanistic insight into these new degradation processes, we

present the first application of real-time time-dependent **density functional theory** (RT-TDDFT) [481, 482] for understanding photo-induced degradation mechanisms in PFASs. While DFT calculations have become more common in environmental research, prior DFT studies on PFASs have only focused on reactions on the electronic ground-state potential energy surface [467, 483]. In contrast, the RT-TDDFT formalism used in this work can describe electronic excited-state dynamics beyond conventional ground-state DFT to explore photo-induced mechanisms and bring a fundamental understanding of these degradation processes. In other words, the photo-induced degradation mechanisms in PFASs inherently occur on electronic-excited-state potential surfaces that cannot be treated with ground-state DFT, and excited-state RT-TDDFT is required for capturing the resulting **photochemical reaction** dynamics.

To this end, the present work constitutes the first quantum dynamical study of photo-induced degradation mechanisms of perfluorooctanoic acid (PFOA) contaminants. We first give a brief description of the RT-TDDFT formalism and our computational methods, which include our custom implementation of optical **electromagnetic fields** in the GPAW software package [484]. We then present a series of RT-TDDFT calculations and electronic diagnostics showing that this photo-induced process selectively degrades PFOA while keeping the water molecules in the surrounding environment intact. Our results are complemented by various analyses, including real-time electronic properties and time-dependent orbital occupations that explain the underlying excited-state mechanisms of the degradation process. Finally, we conclude with a discussion and summary of our results, with additional perspectives on future applications of RT-TDDFT that can have a broad

impact in probing photo-induced interactions of environmental contaminants.

10.2 Computational Details

We used RT-TDDFT with Ehrenfest-based molecular dynamics [485] to probe the degradation of PFOA solvated with 43 explicit water molecules under the influence of electromagnetic radiation. Previous work by us on PFOA has shown that 43 water molecules were sufficient to capture explicit solvent effects [467], which is necessary to mimic the natural solvation environment required for exploring the photo-induced dynamics of PFOA. The use of explicit solvent in this work provides a more accurate simulation of fluorine dissociation dynamics (and the possible formation of photo-induced reaction products) since hydrogen-bonding effects are more accurately captured by explicit solvent compared to coarse-grained polarizable continuum model approaches. The RT-TDDFT formalism with Ehrenfest-based molecular dynamics allows us to accurately capture non-adiabatic processes, such as the optically-induced, excited-state degradation dynamics examined in this work. The time-dependent orbitals were obtained by solving the self-consistent time-dependent Kohn-Sham equations:

$$i\hbar \frac{\partial \psi_n^{KS}(\mathbf{r}, t)}{\partial t} = [H^{KS}(\mathbf{r}, t) + V_{ext}(\mathbf{r}, t)] \psi_n^{KS}(\mathbf{r}, t), \quad (10.1)$$

where $\psi_n^{KS}(\mathbf{r}, t)$ is the n th time-dependent Kohn-Sham orbital, $H^{KS}(\mathbf{r}, t)$ is the Kohn-Sham Hamiltonian, and $V_{ext}(\mathbf{r}, t)$ includes the electric field of the applied optical pulse, which generates the subsequent electronic and atomic dynamics. In this study, we implemented custom optical pulses in the GPAW software package [484] having the form:

$$E(t) = E_0 e^{\frac{(t-t_0)^2}{2\tau_0^2}} \sin(\omega(t-t_0)), \quad (10.2)$$

where E_0 , t_0 , τ_0 , and ω are the electromagnetic pulse's amplitude, center, width, and frequency, respectively. We used the PBE exchange-correlation functional [219] within the projector augmented wave formalism [484] and a 0.3-Å real-space grid spacing for both the geometry optimizations and RT-TDDFT calculations. Using a 0.3-Å real-space grid in our simulations is on par with the settings recommended for RT-TDDFT simulations in GPAW [486]. To further assess the accuracy of our simulation settings, we also carried out two separate benchmark DFT calculations of a single PFOA molecule with a 0.3 and 0.1 Å real-space grid. Our benchmark calculations show that the density of states obtained with a 0.3-Å real-space grid is very similar to those obtained with 0.1 Å. Since the orbital energies near the Fermi level (which are responsible for most of the excitations) are very similar for both grid spacings, the use of a 0.3-Å real-space grid supports the robustness of our results.

The geometry of PFOA + 43H₂O molecules in a 29x24x22 Å³ simulation box was first optimized such that the residual forces were less than 0.01 eV/Å (a large box size was chosen to prevent periodic images of the system from interacting with each other). This optimized geometry was then used as an initial condition for the excited-state, photo-induced Ehrenfest dynamics, and the resulting atomic positions were monitored at each time step. We used an optical pulse with ω set to either 6 or 7.5 eV with a $\sim 10^{15}$ W/cm² intensity, which is commensurate with field strengths used in photocatalysis experiments [487, 488]. Our simulations were carried out for a total of 36 fs with a timestep of 10 attoseconds (the electromagnetic optical pulse (c.f. Fig. 10.1b) acts on the system for the first 20 fs). Finally, it is worth noting that these excited-state RT-TDDFT calculations are significantly more

expensive than conventional ground-state DFT calculations, and this work utilized a total of over 570,000 CPU hours on the XSEDE Comet computing cluster.

10.3 Results and Discussion

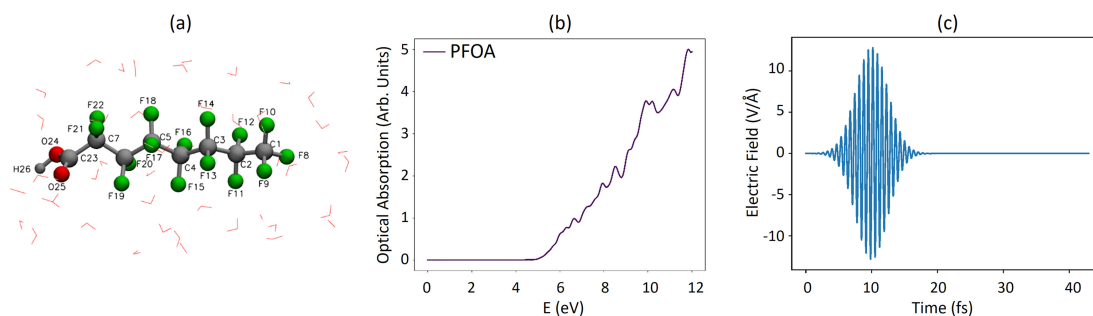


Figure 10.1: (a) PFOA molecule surrounded with 43 explicit water molecules (depicted as thin sticks for visual clarity). (b) Optical absorption spectra of PFOA + 43H₂O molecules. (c) Temporal dependence of the electric field of the applied optical pulse used in our calculations.

The RT-TDDFT methodology described previously was used to probe the dissociation of PFOA with a variety of optical fields. For clarity, we only discuss the dissociation dynamics of PFOA. To choose a relevant excitation energy (ω), we first computed the absorption spectrum of PFOA + 43H₂O molecules (depicted in Fig. 10.1(a)) using a standard RT-TDDFT procedure described in our previous work [72, 489–491]. Specifically, the Kohn-Sham orbitals for these optical absorption calculations were propagated for 25 fs with a time step of 8 as, and a 0.23 eV full-width-at-half-maximum Gaussian broadening was used to plot the spectrum. The optical absorption spectrum in Fig. 10.1(b) shows a near continuum of excitations beyond the optical gap of ~ 5 eV, which indicates that the system does not strongly absorb below 5 eV. As such, we set the optical pulse’s excitation frequency (ω) to 6 eV for all our photo-induced studies. Fig. 10.1(c) depicts the temporal dependence of the

electric field of the optical pulse used in our calculations. The polarization of the electric field is oriented perpendicular to the axis of the PFOA molecule.

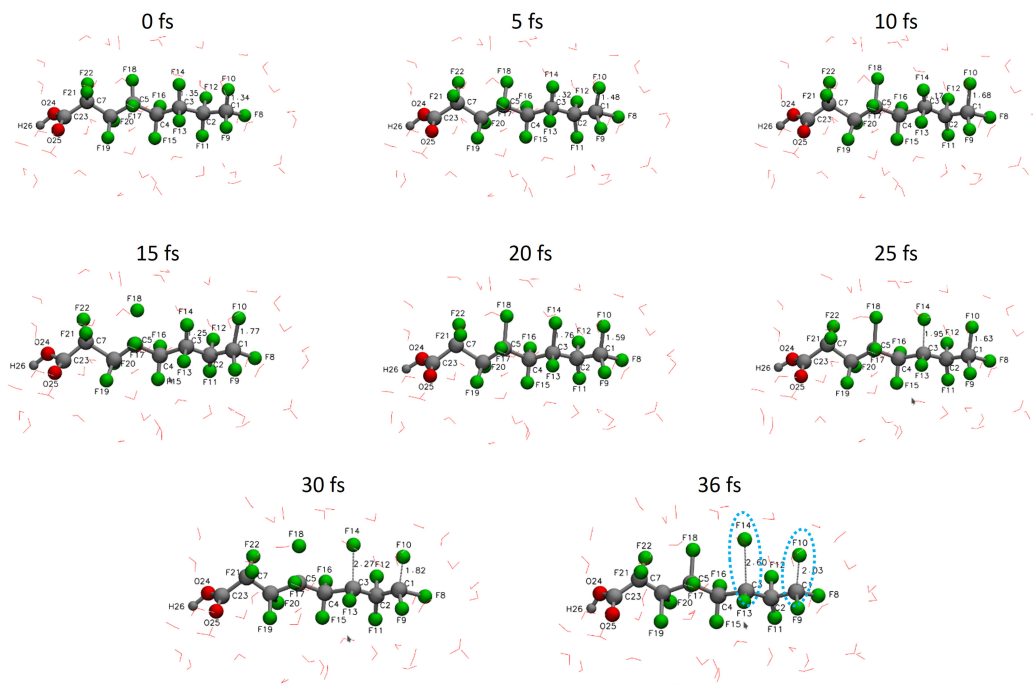


Figure 10.2: Variations in the dissociation dynamics of PFOA + 43H₂O molecules as a function of time. The encircled C3-F14 and C1-F10 bonds in PFOA dissociate at the end of 36 fs.

The dissociation dynamics of various bonds as a function of time are shown in Fig. 10.2. All the C–F bonds have an initial length of ~ 1.35 Å, and upon irradiation, most PFOA bonds begin to oscillate. Selected C–F bond lengths are plotted in greater detail in Fig. 10.3(a). Most notably, the C3–F14 bond initially contracts (with a minimum value at ~ 12 fs) and rises monotonically after that. After ~ 20 fs, all the C–F bonds oscillate until the C3–F14 and C1–F10 bonds finally dissociate. At the end of the simulation, the fluorine atom dissociates to 2.6 Å from the backbone carbon atom. When a higher-energy 7.5 eV optical pulse is used, the degradation of PFOA occurs on a similar timescale. It is interesting to

note that, apart from minor elongations of the O–H bond, the water molecules surrounding PFOA remain undissociated throughout the simulation. As such, our simulations indicate that these photo-induced processes effectively treat these systems without unintended side reactions (i.e., the surrounding water molecules remain intact since none of the O–H bonds dissociate).

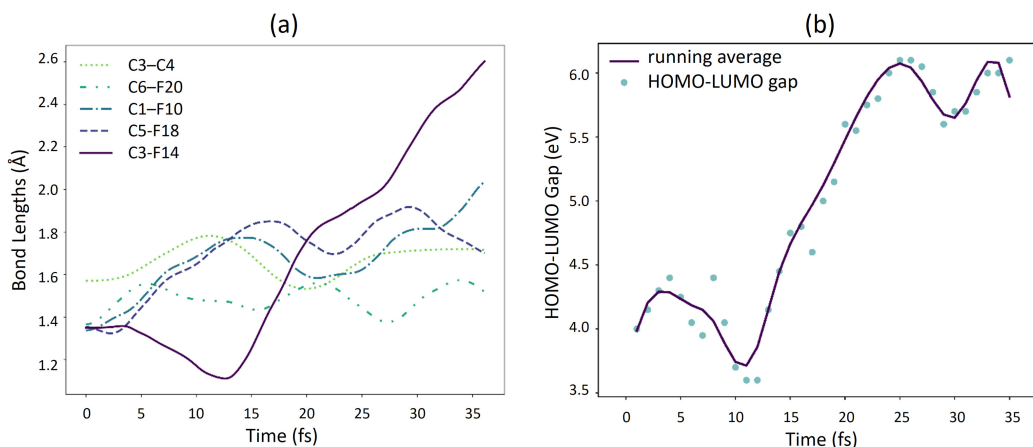


Figure 10.3: (a) Variation of selected C–F and C–C bond lengths in PFOA as a function of time (see Fig. 10.1(a) for atom numbering). After 12 fs, an irreversible elongation of the C3-F14 and C1-F10 bond in PFOA occurs. (b) Variation of the HOMO-LUMO gap of the system as a function of time. At 12 fs, the HOMO-LUMO gap reaches a minimum at which a photo-excited electron transfer occurs, and the C3–F14 bond starts to irreversibly break.

To further examine the induced mechanisms giving rise to this degradation process, Fig. 10.3(b) plots the HOMO-LUMO quasiparticle gap of the system as a function of time. The quasiparticle gap decreases to a critical value near 12 fs. At this point, the electromagnetic field polarizes the C3–F14 bond and induces a charge-transfer excitation that initiates a bond dissociation (and a widening of the quasiparticle gap, which is discussed further below). This excitation process also manifests itself in the C3–F14 bond length, which shows a similar variation in Fig. 10.3(a). This dynamic charge-transfer process is depicted in greater detail in Fig. 10.4(a), which plots the Bader charges for atoms C3

and F14 as a function of time. As the C3–F14 bond becomes polarized, the Bader charge of the C3 atom gradually increases within the 12–20 fs range, whereas the F14 atom shows a slight decrease in charge. Fig. 10.4 shows a visualization of the C3–F14 bond dissociation, which plots the total RT-TDDFT charge density as a function of time. For clarity, the charge density of the surrounding water molecules is not displayed, and the fluorine atom and its charge density dissociates at 20 ps.

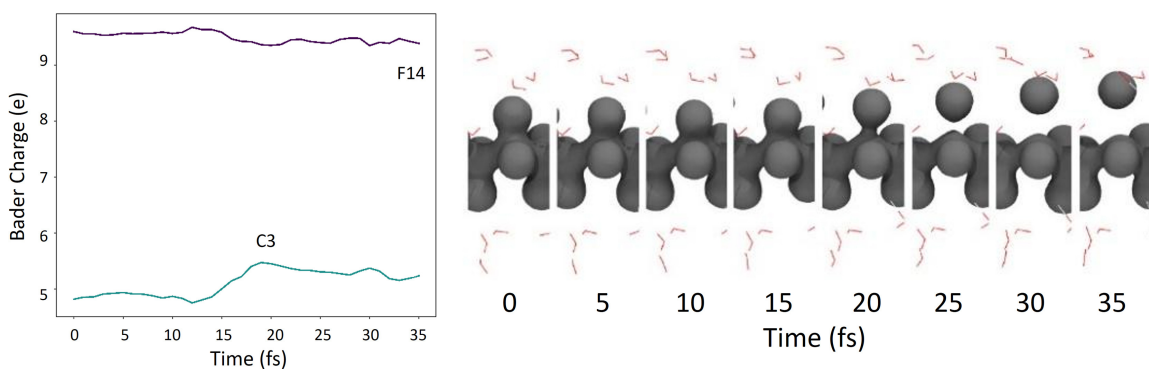


Figure 10.4: (left panel) Time-dependent Bader charges for the dissociated C3-F14 bond on PFOA; (right panel) variation of the total RT-TDDFT charge density of PFOA for the C3-F14 bond as a function of time.

Finally, to further examine the polarization and subsequent dissociation dynamics of the C–F bonds in PFOA, we computed the Kohn-Sham orbitals as a function of time in Fig. 10.5. Since these orbitals are complex-valued (i.e., Eq. 10.2 is a complex-valued integro-differential equation), we used the following visualization scheme in Fig. 10.5: the opacity denotes the magnitude of the complex-valued orbital, and its color (ranging from red to blue) represents its phase. At $t = 0$, the highest occupied molecular orbital (HOMO) is real-valued and almost entirely localized on nearby water molecules. As time progresses, the HOMO starts to delocalize onto the PFOA molecule during 8–16 fs, and some electronic charge transfers from the water to PFOA. This excess charge occupies the previously empty

$3s_F$ and $2p_C$ atomic orbitals on PFOA, which possesses a strong σ^* anti-bonding character. Once these empty orbitals become occupied, a dissociative **electron attachment** (DEA) process occurs: the subsequent electron-nuclei motion causes the C–F bond to elongate and destabilize (since the HOMO now has an anti-bonding character) until it irreversibly dissociates. This dynamic mechanism is reminiscent of the pseudo-Jahn-Teller effect that enables a charge density redistribution and produces highly directional ionic forces that destabilize specific bonds in chemical/material systems [492]. Finally, the surrounding water molecules stabilize the positively charged hole remaining in the solvent (which arises from a charge-transfer excitation). This general dissociation mechanism is also similar to that observed in high-resolution gas-phase electron-collision experiments in which electron attachment to a **CF**₄ molecule also leads to C–F bond dissociation [493]. However, the significant difference is that electrons in these prior experiments were artificially produced with an electron beam/gun. In contrast, the source of electrons in our solvated PFAS system naturally arises from nearby water molecules that are dynamically excited with an optical pulse. Finally, it is worth mentioning that our approach highlights the importance of utilizing these RT-TDDFT techniques for probing excited-state, photo-induced degradation of PFAS contaminants, which have started to garner immense attention in the scientific community [475, 478]. Most importantly, these excited-states photo-induced mechanisms *cannot be gleaned from conventional ground-state DFT calculations*. The RT-TDDFT techniques enable a new understanding of the excited-state dynamical mechanisms in these critical degradation processes.

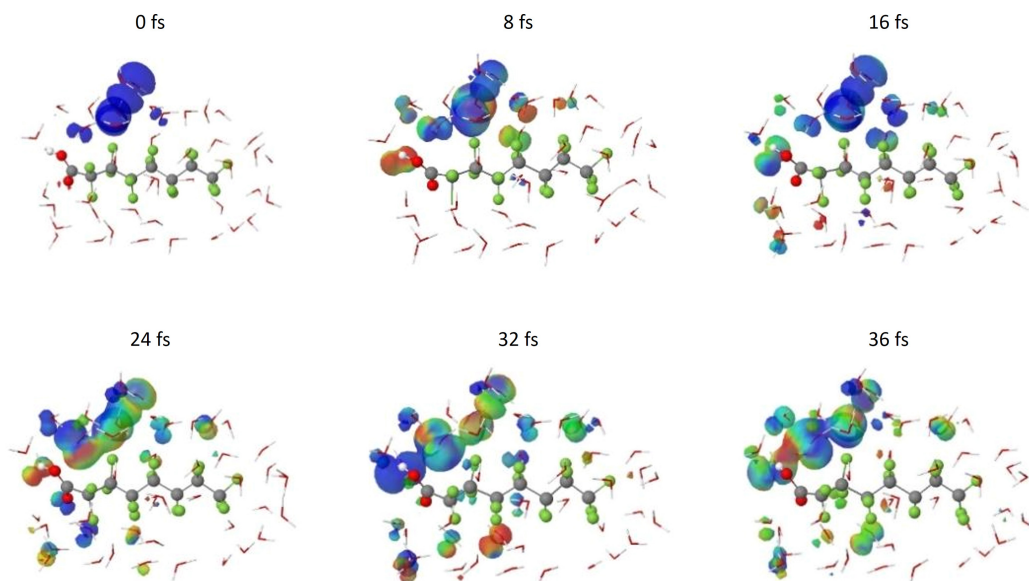


Figure 10.5: Variation of the HOMO for PFOA + 43H₂O molecules as a function of time. As time progresses, charge dynamically transfers from the water to an anti-bonding orbital on PFOA. The C-F bond begins to subsequently destabilize and dissociate as it continues to vibrate dynamically. The positively charged hole remaining in the solvent arising from the afore-mentioned charge-transfer excitation becomes stabilized by the surrounding water molecules.

10.4 Conclusion

In conclusion, we have conducted the first RT-TDDFT study of excited-state dynamics in PFAS pollutants to understand their photo-induced degradation mechanisms. By explicitly accounting for non-adiabatic excited-state interactions in solvated PFOA, we show that these photo-induced excitations enable a charge-transfer process that polarizes the C-F bond, resulting in a dynamic dissociation on a femtosecond time scale. Moreover, we show that this photo-induced process is highly selective and only affects the PFOA molecules while keeping the surrounding water molecules intact. The parameters used in these calculations are commensurate with commercially available monochromatic light/laser

sources and shed crucial mechanistic insight into their excited-state degradation dynamics.

Looking forward, we anticipate that these state-of-the-art RT-TDDFT approaches could accelerate PFAS remediation efforts in two quick ways. First, our predictive calculations demonstrate that electromagnetic/optical fields are a viable and *direct* approach to degrade PFAS pollutants as opposed to conventional filtration techniques that merely remove PFASs (which still require treatment after filtration). Second, these RT-TDDFT approaches can guide ongoing experimental efforts by rationalizing or high-throughput screening new photocatalytic materials/surfaces, which have started to garner immense attention for enhanced PFAS degradation. As such, the RT-TDDFT techniques used in this work provide fresh opportunities for actively exploring excited-state, photo-induced degradation dynamics in the remediation of PFASs and other environmental contaminants.

Chapter 11

Conclusions

This dissertation has presented advances in developing and implementing numerical methods for real-time time-dependent density functional theory (rt-TDDFT) and studying complex systems' excited states. Each chapter is summarized below. Please refer to the published version of each chapter for the complete analysis.

In Chapter 2, we performed computational work using quantum mechanical (QM) calculations to investigate Diels-Alder reactions with cyclodextrins as catalysts. We studied the reaction of 9-anthracene-methanol with N-cyclohexylmaleimide catalyzed by β -cyclodextrins. Our QM calculations showed that cyclodextrins reduced the activation entropy. The formation of the complex is the key to catalysis because it increases the rate that reactant molecules would meet each other and lowers the entropy and thus the activation free energy of this reaction.

In Chapter 3, we carried out high-level CCSD(T) analyses on the DMP⁺ cation to show that DFT performs quite well for this system, which is in contrast to their conclusion

that all DFT functionals fail. Whereas the PZ-SIC approach used by Weber et al. [57] is the outlier that is inconsistent with the high-level CCSD(T) (coupled-cluster with single and double excitations and perturbative triples) calculations.

In Chapter 4, we present a new QSAR model using semi-supervised metric learning techniques to assess which chemical functional groups affect bioactivities toward specific biological targets. Our approach is integrated with molecular docking calculations to predict possible bioactivities of PFAS molecules based on their chemical functional groups and specific biological targets. We combine dimension reduction and clustering methods to classify PFASs based on their molecular structures. We then apply a semi-supervised metric learning method to improve classification accuracy. Finally, we use a molecular docking approach to illuminate the physicochemical reasons for their bioactivity. Our study provides the first unsupervised/semi-supervised learning approach for screening potentially bioactive PFAS molecules beyond conventional supervised learning or QSAR approaches.

In Chapter 5, we calculate the formation energies of various silicon carbide (SiC) and zinc oxide (ZnO) configurations to highlight the efficiencies of our DFTB+CASM implementation. Our extensive calculations show that researchers can use DFTB as an efficient screening tool to compute complex materials' numerous formation energies (and convex hull if it exists). To highlight the efficiency and accuracy of our approach, we calculated and compared the formation energies of SiC and ZnO with both DFT and DFTB. We find that the DFTB approach enables highly efficient calculations of formation energies in a completely unbiased manner to predict low-lying metastable phases over the entire composition space.

In Chapter 6, we have studied the collision of a water (H_2O) molecule with a ZnO surface precovered with carbon monoxide (CO) with both double- and triple-zeta quality basis sets using BOMD simulations. To gain further insight, we have also studied the interaction of CO, CO_2 , H_2 , and H_2O with the ZnO surface with both of these basis sets. Based on the results from both BOMD and DFT, we propose that single-point binding energy calculations can be used as a simple diagnostic tool to estimate the impact of basis-set quality before carrying out a computationally expensive BOMD simulation. Finally, we note that our simulations are relevant to the industrially relevant water–gas shift reaction, where ZnO is one of the co-catalysts, and all of the molecules mentioned above are either products or reactants of this reaction.

In Chapter 7, we accurately reproduce and characterize the experimental spectra using linear-response time-dependent density functional theory (lr-TDDFT). At the normal state, we observe a soft plasmon mode directly representing the exciton condensation, which we attributed to interband transitions. At the CDW state below T_c , the periodic lattice distortions introduce a CDW gap and hamper the spontaneous formation of the exciton. Above T_c , the higher electronic temperature prevents condensation by introducing a dissipation channel. The combined effect of the CDW and finite electronic temperature explains the fragile temperature dependence of the exciton condensation.

In Chapter 8, we present the real-time TDDFT formalism. We demonstrated the implementations of the RT-TDDFT, and the advantages of using RT-TDDFT over LR-TDDFT for particular cases are discussed and illustrated. We validated our program by optical absorption calculations, which show our implementation’s validity and indicate

our code’s limited accuracy. We then applied our program to field-induced dynamics on a graphene surface, reproducing field-induced charge transfer between a metal oxide and a graphene sheet. These detailed consistency tests validate our new RT-TDDFT implementation for solids.

In Chapter 9, we presented a new mechanism that enables this efficient switching of the polarization in FE materials. By accounting for the detailed interactions between the electrons, nuclei, and electromagnetic field, our rt-TDDFT simulations show that the ultrafast FE polarization switching in BaTiO₃ occurs within 200 fs after the material is excited by an 800 nm laser pulse. This polarization switching commences when the laser pulse pumps electrons from 2p orbitals in oxygen to the 3d orbitals in titanium. This dynamic excitation process yields desirable and highly directional forces on the Ti and O atoms that are (1) always opposite to the original FE displacements and (2) can be harnessed to switch the material’s polarization consistently. By laser-tuning the dynamical potential energy surface, we show that this new mechanism can switch the FE polarization in both directions using identical pulses from experimentally available laser sources.

In Chapter 10, we carried out an rt-TDDFT study of excited-state dynamics in PFAS pollutants to understand their photo-induced degradation mechanisms. By explicitly accounting for non-adiabatic excited-state interactions in solvated PFOA, we show that these photo-induced excitations enable a charge-transfer process that polarizes the C–F bond, resulting in a dynamic dissociation on a femtosecond time scale. Moreover, we show that this photo-induced process is highly selective and only affects the PFOA molecules while keeping the surrounding water molecules intact.

In conclusion, this dissertation demonstrates the importance of quantum simulation using the real-time time-dependent density functional theory (rt-TDDFT). It shows how rt-TDDFT can be used to help experimental groups with analysis by giving further insight into experimentally challenging mechanisms.

Bibliography

- [1] K. Krieger et al. “Laser-Induced Demagnetization at Ultrashort Time Scales: Predictions of TDDFT”. In: *Journal of Chemical Theory and Computation* 11.10 (2015), pp. 4870–4874. DOI: [10.1021/acs.jctc.5b00621](https://doi.org/10.1021/acs.jctc.5b00621).
- [2] C. Lian et al. “Ultrafast charge ordering by self-amplified exciton–phonon dynamics in TiSe₂”. In: *Nature Communications* 11.1 (Jan. 2020). DOI: [10.1038/s41467-019-13672-7](https://doi.org/10.1038/s41467-019-13672-7).
- [3] Y. Zhang et al. “Surface-Plasmon-Driven Hot Electron Photochemistry”. In: *Chemical Reviews* 118.6 (Nov. 2017), pp. 2927–2954. DOI: [10.1021/acs.chemrev.7b00430](https://doi.org/10.1021/acs.chemrev.7b00430).
- [4] N. Kolobov et al. “UV-LED photocatalytic oxidation of carbon monoxide over TiO₂ supported with noble metal nanoparticles”. In: *Chemical Engineering Journal* 314 (Apr. 2017), pp. 600–611. DOI: [10.1016/j.cej.2016.12.018](https://doi.org/10.1016/j.cej.2016.12.018).
- [5] Q. Guo et al. “Fundamentals of TiO₂ Photocatalysis: Concepts, Mechanisms, and Challenges”. In: *Advanced Materials* 31.50 (Aug. 2019), p. 1901997. DOI: [10.1002/adma.201901997](https://doi.org/10.1002/adma.201901997).
- [6] G. Vampa et al. “Semiclassical analysis of high harmonic generation in bulk crystals”. In: *Physical Review B* 91.6 (Feb. 2015). DOI: [10.1103/physrevb.91.064302](https://doi.org/10.1103/physrevb.91.064302).
- [7] I. Floss et al. “*Ab initio* multiscale simulation of high-order harmonic generation in solids”. In: *Physical Review A* 97.1 (Jan. 2018). DOI: [10.1103/physreva.97.011401](https://doi.org/10.1103/physreva.97.011401).
- [8] N. Tancogne-Dejean et al. “Impact of the Electronic Band Structure in High-Harmonic Generation Spectra of Solids”. In: *Physical Review Letters* 118.8 (Feb. 2017). DOI: [10.1103/physrevlett.118.087403](https://doi.org/10.1103/physrevlett.118.087403).
- [9] V. Q. Dang et al. “Ultrahigh Responsivity in Graphene-ZnO Nanorod Hybrid UV Photodetector”. In: *Small* 11.25 (Feb. 2015), pp. 3054–3065. DOI: [10.1002/smll.201403625](https://doi.org/10.1002/smll.201403625).

- [10] M. Sentef et al. “Theory of Floquet band formation and local pseudospin textures in pump-probe photoemission of graphene”. In: *Nature Communications* 6.1 (May 2015). DOI: [10.1038/ncomms8047](https://doi.org/10.1038/ncomms8047).
- [11] T. Oka and H. Aoki. “Photovoltaic Hall effect in graphene”. In: *Physical Review B* 79.8 (Feb. 2009). DOI: [10.1103/physrevb.79.081406](https://doi.org/10.1103/physrevb.79.081406).
- [12] U. D. Giovannini, H. Hübener, and A. Rubio. “A First-Principles Time-Dependent Density Functional Theory Framework for Spin and Time-Resolved Angular-Resolved Photoelectron Spectroscopy in Periodic Systems”. In: *Journal of Chemical Theory and Computation* 13.1 (Dec. 2016), pp. 265–273. DOI: [10.1021/acs.jctc.6b00897](https://doi.org/10.1021/acs.jctc.6b00897).
- [13] M. Holler et al. “Attosecond Electron Wave-Packet Interference Observed by Transient Absorption”. In: *Physical Review Letters* 106.12 (Mar. 2011). DOI: [10.1103/physrevlett.106.123601](https://doi.org/10.1103/physrevlett.106.123601).
- [14] A. R. Beck et al. “Attosecond transient absorption probing of electronic superpositions of bound states in neon: detection of quantum beats”. In: *New Journal of Physics* 16.11 (Nov. 2014), p. 113016. DOI: [10.1088/1367-2630/16/11/113016](https://doi.org/10.1088/1367-2630/16/11/113016).
- [15] E. Runge and E. K. U. Gross. “Density-Functional Theory for Time-Dependent Systems”. In: *Physical Review Letters* 52.12 (Mar. 1984), pp. 997–1000. DOI: [10.1103/physrevlett.52.997](https://doi.org/10.1103/physrevlett.52.997).
- [16] C. A. Ullrich. *Time-Dependent Density-Functional Theory*. Oxford University Press, Dec. 2011. DOI: [10.1093/acprof:oso/9780199563029.001.0001](https://doi.org/10.1093/acprof:oso/9780199563029.001.0001).
- [17] S. Sato et al. “Ab Initio Simulation of Attosecond Transient Absorption Spectroscopy in Two-Dimensional Materials”. In: *Applied Sciences* 8.10 (Sept. 2018), p. 1777. DOI: [10.3390/app8101777](https://doi.org/10.3390/app8101777).
- [18] F. Schlaepfer et al. “Attosecond optical-field-enhanced carrier injection into the GaAs conduction band”. In: *Nature Physics* 14.6 (Mar. 2018), pp. 560–564. DOI: [10.1038/s41567-018-0069-0](https://doi.org/10.1038/s41567-018-0069-0).
- [19] M. Zürich et al. “Direct and simultaneous observation of ultrafast electron and hole dynamics in germanium”. In: *Nature Communications* 8.1 (June 2017). DOI: [10.1038/ncomms15734](https://doi.org/10.1038/ncomms15734).
- [20] M. Schultze et al. “Attosecond band-gap dynamics in silicon”. In: *Science* 346.6215 (Dec. 2014), pp. 1348–1352. DOI: [10.1126/science.1260311](https://doi.org/10.1126/science.1260311).

- [21] M. Lucchini et al. “Attosecond dynamical Franz-Keldysh effect in polycrystalline diamond”. In: *Science* 353.6302 (Aug. 2016), pp. 916–919. DOI: [10.1126/science.aag1268](https://doi.org/10.1126/science.aag1268).
- [22] W. Kohn and L. J. Sham. “Self-Consistent Equations Including Exchange and Correlation Effects”. In: *Phys. Rev.* 140 (4A Nov. 1965), A1133–A1138. DOI: [10.1103/PhysRev.140.A1133](https://doi.org/10.1103/PhysRev.140.A1133).
- [23] W. Chen et al. “An MM and QM Study of Biomimetic Catalysis of Diels-Alder Reactions Using Cyclodextrins”. In: *Catalysts* 8.2 (Jan. 2018), p. 51. DOI: [10.3390/catal8020051](https://doi.org/10.3390/catal8020051).
- [24] R. Challa et al. “Cyclodextrins in drug delivery: An updated review”. In: *AAPS PharmSciTech* 6.2 (June 2005), E329–E357. DOI: [10.1208/pt060243](https://doi.org/10.1208/pt060243).
- [25] E. D. Valle. “Cyclodextrins and their uses: a review”. In: *Process Biochemistry* 39.9 (May 2004), pp. 1033–1046. DOI: [10.1016/s0032-9592\(03\)00258-9](https://doi.org/10.1016/s0032-9592(03)00258-9).
- [26] M. Singh, R. Sharma, and U. Banerjee. “Biotechnological applications of cyclodextrins”. In: *Biotechnology Advances* 20.5-6 (Dec. 2002), pp. 341–359. DOI: [10.1016/s0734-9750\(02\)00020-4](https://doi.org/10.1016/s0734-9750(02)00020-4).
- [27] M. E. Davis and M. E. Brewster. “Cyclodextrin-based pharmaceuticals: past, present and future”. In: *Nature Reviews Drug Discovery* 3.12 (Dec. 2004), pp. 1023–1035. DOI: [10.1038/nrd1576](https://doi.org/10.1038/nrd1576).
- [28] W.-F. Lai. “Cyclodextrins in non-viral gene delivery”. In: *Biomaterials* 35.1 (Jan. 2014), pp. 401–411. DOI: [10.1016/j.biomaterials.2013.09.061](https://doi.org/10.1016/j.biomaterials.2013.09.061).
- [29] L. Marchetti and M. Levine. “Biomimetic Catalysis”. In: *ACS Catalysis* 1.9 (Aug. 2011), pp. 1090–1118. DOI: [10.1021/cs200171u](https://doi.org/10.1021/cs200171u).
- [30] Y. A. Tayade et al. “ β -Cyclodextrin as a supramolecular catalyst for the synthesis of dihydropyrano[2,3-c]pyrazole and spiro[indoline-3,4'-pyrano[2,3-c]pyrazole] in aqueous medium”. In: *Tetrahedron Letters* 56.19 (May 2015), pp. 2441–2447. DOI: [10.1016/j.tetlet.2015.03.084](https://doi.org/10.1016/j.tetlet.2015.03.084).
- [31] I. Tabushi. “Cyclodextrin catalysis as a model for enzyme action”. In: *Accounts of Chemical Research* 15.3 (Mar. 1982), pp. 66–72. DOI: [10.1021/ar00075a001](https://doi.org/10.1021/ar00075a001).
- [32] F. Hapiot et al. “Recent breakthroughs in aqueous cyclodextrin-assisted supramolecular catalysis”. In: *Catalysis Science & Technology* 4.7 (2014), p. 1899. DOI: [10.1039/c4cy00005f](https://doi.org/10.1039/c4cy00005f).

- [33] R. Breslow, H. Kohn, and B. Siegel. "Methylated cyclodextrin and a cyclodextrin polymer as catalysts in selective anisole chlorination". In: *Tetrahedron Letters* 17.20 (May 1976), pp. 1645–1646. DOI: [10.1016/s0040-4039\(00\)92912-5](https://doi.org/10.1016/s0040-4039(00)92912-5).
- [34] D. J. DiScenza et al. "Cyclodextrin-Promoted Detection of Aromatic Toxicants and Toxicant Metabolites in Urine". In: *Analytical Chemistry Letters* 6.4 (June 2016), pp. 345–353. DOI: [10.1080/22297928.2016.1210021](https://doi.org/10.1080/22297928.2016.1210021).
- [35] R. Breslow and S. D. Dong. "Biomimetic Reactions Catalyzed by Cyclodextrins and Their Derivatives". In: *Chemical Reviews* 98.5 (May 1998), pp. 1997–2012. DOI: [10.1021/cr970011j](https://doi.org/10.1021/cr970011j).
- [36] D. C. Rideout and R. Breslow. "Hydrophobic acceleration of Diels-Alder reactions". In: *Journal of the American Chemical Society* 102.26 (Dec. 1980), pp. 7816–7817. DOI: [10.1021/ja00546a048](https://doi.org/10.1021/ja00546a048).
- [37] H.-J. Schneider and N. K. Sangwan. "Diels–Alder reactions in hydrophobic cavities: a quantitative correlation with solvophobicity and rate enhancements by macrocycles". In: *J. Chem. Soc., Chem. Commun.* 24 (1986), pp. 1787–1789. DOI: [10.1039/c39860001787](https://doi.org/10.1039/c39860001787).
- [38] S. Chaudhuri, T. Phelan, and M. Levine. "Cyclodextrin-promoted Diels Alder reactions of a polycyclic aromatic hydrocarbon under mild reaction conditions". In: *Tetrahedron Letters* 56.13 (Mar. 2015), pp. 1619–1623. DOI: [10.1016/j.tetlet.2015.01.185](https://doi.org/10.1016/j.tetlet.2015.01.185).
- [39] S. P. Kim, A. G. Leach, and K. N. Houk. "The Origins of Noncovalent Catalysis of Intermolecular Diels-Alder Reactions by Cyclodextrins, Self-Assembling Capsules, Antibodies, and RNAses". In: *The Journal of Organic Chemistry* 67.12 (May 2002), pp. 4250–4260. DOI: [10.1021/jo011180d](https://doi.org/10.1021/jo011180d).
- [40] D. Boger. *Modern Organic Synthesis*. San Diego, CA, USA, 1999.
- [41] J. B. Siegel et al. "Computational Design of an Enzyme Catalyst for a Stereoselective Bimolecular Diels-Alder Reaction". In: *Science* 329.5989 (June 2010), pp. 309–313. DOI: [10.1126/science.1190239](https://doi.org/10.1126/science.1190239).
- [42] V. E. Gouverneur et al. "Control of the exo and endo Pathways of the Diels-Alder Reaction by Antibody Catalysis". In: *Science* 262.5131 (Oct. 1993), pp. 204–208. DOI: [10.1126/science.8211138](https://doi.org/10.1126/science.8211138).
- [43] V. B. Luzhkov and C. A. Venanzi. "Computer Modeling of Phenyl Acetate Hydrolysis in Water and in Reaction with β -Cyclodextrin: Molecular Orbital Calculations with the Semiempirical AM1 Method and the Langevin Dipole Solvent Model". In: *The*

- Journal of Physical Chemistry* 99.8 (Feb. 1995), pp. 2312–2323. DOI: [10.1021/j100008a012](https://doi.org/10.1021/j100008a012).
- [44] M. Eto et al. “Molecular Orbital Studies on Pericyclic Reactions of Cinnamyl Xanthates in β -Cyclodextrin Cavities”. In: *Chemical and Pharmaceutical Bulletin* 48.11 (2000), pp. 1652–1659. DOI: [10.1248/cpb.48.1652](https://doi.org/10.1248/cpb.48.1652).
- [45] A. M. Granados et al. “Effect of β -cyclodextrin on the hydrolysis of N-phenylphtalamide and N-adamantylphtalamide—a two-sided semiempirical approach”. In: *Journal of Molecular Structure: THEOCHEM* 619.1-3 (Dec. 2002), pp. 91–100. DOI: [10.1016/s0166-1280\(02\)00549-3](https://doi.org/10.1016/s0166-1280(02)00549-3).
- [46] Y. Wan, X. Wang, and N. Liu. “The reactivity of phenancyl bromide under β -cyclodextrin as supramolecular catalyst: a computational survey”. In: *Journal of Molecular Modeling* 21.5 (May 2015). DOI: [10.1007/s00894-015-2680-7](https://doi.org/10.1007/s00894-015-2680-7).
- [47] P. Li et al. “Investigating the Reactivity of 1,4-Anthracene-Incorporated Cycloparaphenylene”. In: *Organic Letters* 18.7 (Mar. 2016), pp. 1574–1577. DOI: [10.1021/acs.orglett.6b00430](https://doi.org/10.1021/acs.orglett.6b00430).
- [48] H. Daver et al. “Quantum Chemical Modeling of Cycloaddition Reaction in a Self-Assembled Capsule”. In: *Journal of the American Chemical Society* 139.43 (Oct. 2017), pp. 15494–15503. DOI: [10.1021/jacs.7b09102](https://doi.org/10.1021/jacs.7b09102).
- [49] W. Humphrey, A. Dalke, and K. Schulten. “VMD: Visual molecular dynamics”. In: *Journal of Molecular Graphics* 14.1 (Feb. 1996), pp. 33–38. DOI: [10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
- [50] R. D. Dennington, T. A. Keith, and J. Millam. *GaussView Version 5*. Semichem Inc. Shawnee Mission KS. 2009.
- [51] M. J. Frisch et al. *Gaussian 09 Revision E.01*. Wallingford, CT, USA, 2009.
- [52] D. McQuarrie and J. Simon. *Molecular Thermodynamics*. Sausalito, CA, USA: University Science Books, 1999.
- [53] X. Li and M. J. Frisch. “Energy-Represented Direct Inversion in the Iterative Subspace within a Hybrid Geometry Optimization Method”. In: *Journal of Chemical Theory and Computation* 2.3 (Apr. 2006), pp. 835–839. DOI: [10.1021/ct050275a](https://doi.org/10.1021/ct050275a).
- [54] C. Peng and H. B. Schlegel. “Combining Synchronous Transit and Quasi-Newton Methods to Find Transition States”. In: *Israel Journal of Chemistry* 33.4 (1993), pp. 449–454. DOI: [10.1002/ijch.199300051](https://doi.org/10.1002/ijch.199300051).

- [55] C. Peng et al. “Using redundant internal coordinates to optimize equilibrium geometries and transition states”. In: *Journal of Computational Chemistry* 17.1 (Jan. 1996), pp. 49–56. DOI: [10.1002/\(sici\)1096-987x\(19960115\)17:1<49::aid-jcc5>3.0.co;2-0](https://doi.org/10.1002/(sici)1096-987x(19960115)17:1<49::aid-jcc5>3.0.co;2-0).
- [56] Z. A. Ali, F. W. Aquino, and B. M. Wong. “The diamine cation is not a chemical example where density functional theory fails”. In: *Nature Communications* 9.1 (Nov. 2018), p. 4733. DOI: [10.1038/s41467-018-07266-y](https://doi.org/10.1038/s41467-018-07266-y).
- [57] X. Cheng et al. “Charge localization in a diamine cation provides a test of energy functionals and self-interaction correction”. In: *Nature Communications* 7.1 (Mar. 2016). DOI: [10.1038/ncomms11013](https://doi.org/10.1038/ncomms11013).
- [58] J. P. Perdew and A. Zunger. “Self-interaction correction to density-functional approximations for many-electron systems”. In: *Physical Review B* 23.10 (May 1981), pp. 5048–5079. DOI: [10.1103/physrevb.23.5048](https://doi.org/10.1103/physrevb.23.5048).
- [59] H. Gudmundsdóttir et al. “Self-interaction corrected density functional calculations of molecular Rydberg states”. In: *The Journal of Chemical Physics* 139.19 (Nov. 2013), p. 194102. DOI: [10.1063/1.4829539](https://doi.org/10.1063/1.4829539).
- [60] Y. Zhang, P. M. Weber, and H. Jónsson. “Self-Interaction Corrected Functional Calculations of a Dipole-Bound Molecular Anion”. In: *The Journal of Physical Chemistry Letters* 7.11 (May 2016), pp. 2068–2073. DOI: [10.1021/acs.jpcclett.6b00742](https://doi.org/10.1021/acs.jpcclett.6b00742).
- [61] S. Lehtola, M. Head-Gordon, and H. Jónsson. “Complex Orbitals, Multiple Local Minima, and Symmetry Breaking in Perdew–Zunger Self-Interaction Corrected Density Functional Theory Calculations”. In: *Journal of Chemical Theory and Computation* 12.7 (June 2016), pp. 3195–3207. DOI: [10.1021/acs.jctc.6b00347](https://doi.org/10.1021/acs.jctc.6b00347).
- [62] S. Lehtola, E. Ó. Jónsson, and H. Jónsson. “Effect of Complex-Valued Optimal Orbitals on Atomization Energies with the Perdew–Zunger Self-Interaction Correction to Density Functional Theory”. In: *Journal of Chemical Theory and Computation* 12.9 (Aug. 2016), pp. 4296–4302. DOI: [10.1021/acs.jctc.6b00622](https://doi.org/10.1021/acs.jctc.6b00622).
- [63] J. Seo et al. “The impact of environment and resonance effects on the site of protonation of aminobenzoic acid derivatives”. In: *Physical Chemistry Chemical Physics* 18.36 (2016), pp. 25474–25482. DOI: [10.1039/c6cp04941a](https://doi.org/10.1039/c6cp04941a).
- [64] M. Melander et al. “Implementation of Constrained DFT for Computing Charge Transfer Rates within the Projector Augmented Wave Method”. In: *Journal of Chemical Theory and Computation* 12.11 (Oct. 2016), pp. 5367–5378. DOI: [10.1021/acs.jctc.6b00815](https://doi.org/10.1021/acs.jctc.6b00815).

- [65] E. Skúlason and H. Jónsson. “Atomic scale simulations of heterogeneous electrocatalysis: recent advances”. In: *Advances in Physics: X* 2.3 (May 2017), pp. 481–495. DOI: [10.1080/23746149.2017.1308230](https://doi.org/10.1080/23746149.2017.1308230).
- [66] Y. Zhang et al. “Observation of Structural Wavepacket Motion: The Umbrella Mode in Rydberg-Excited *N*-Methyl Morpholine”. In: *The Journal of Physical Chemistry Letters* 8.16 (July 2017), pp. 3740–3744. DOI: [10.1021/acs.jpcllett.7b01274](https://doi.org/10.1021/acs.jpcllett.7b01274).
- [67] A. A. Ischenko, P. M. Weber, and R. J. D. Miller. “Capturing Chemistry in Action with Electrons: Realization of Atomically Resolved Reaction Dynamics”. In: *Chemical Reviews* 117.16 (July 2017), pp. 11066–11124. DOI: [10.1021/acs.chemrev.6b00770](https://doi.org/10.1021/acs.chemrev.6b00770).
- [68] Y. Zhang, H. Jónsson, and P. M. Weber. “Coherence in nonradiative transitions: internal conversion in Rydberg-excited *N*-methyl and *N*-ethyl morpholine”. In: *Phys. Chem. Chem. Phys.* 19.38 (2017), pp. 26403–26411. DOI: [10.1039/c7cp05244h](https://doi.org/10.1039/c7cp05244h).
- [69] D. you Kao et al. “Self-consistent self-interaction corrected density functional theory calculations for atoms using Fermi-Löwdin orbitals: Optimized Fermi-orbital descriptors for Li–Kr”. In: *The Journal of Chemical Physics* 147.16 (Oct. 2017), p. 164107. DOI: [10.1063/1.4996498](https://doi.org/10.1063/1.4996498).
- [70] J. C. Rienstra-Kiracofe et al. “Atomic and Molecular Electron Affinities: Photoelectron Experiments and Theoretical Computations”. In: *Chemical Reviews* 102.1 (Jan. 2002), pp. 231–282. DOI: [10.1021/cr990044u](https://doi.org/10.1021/cr990044u).
- [71] T. B. Adler, G. Knizia, and H.-J. Werner. “A simple and efficient CCSD(T)-F12 approximation”. In: *The Journal of Chemical Physics* 127.22 (Dec. 2007), p. 221106. DOI: [10.1063/1.2817618](https://doi.org/10.1063/1.2817618).
- [72] M. B. Oviedo, N. V. Ilawe, and B. M. Wong. “Polarizabilities of π -Conjugated Chains Revisited: Improved Results from Broken-Symmetry Range-Separated DFT and New CCSD(T) Benchmarks”. In: *Journal of Chemical Theory and Computation* 12.8 (July 2016), pp. 3593–3602. DOI: [10.1021/acs.jctc.6b00360](https://doi.org/10.1021/acs.jctc.6b00360).
- [73] H. Kwon, Z. A. Ali, and B. M. Wong. “Harnessing Semi-Supervised Machine Learning to Automatically Predict Bioactivities of Per- and Polyfluoroalkyl Substances (PFASs)”. In: *Environmental Science & Technology Letters* (Aug. 2022). DOI: [10.1021/acs.estlett.2c00530](https://doi.org/10.1021/acs.estlett.2c00530).
- [74] E. Hepburn et al. “Contamination of groundwater with per- and polyfluoroalkyl substances (PFAS) from legacy landfills in an urban re-development precinct”. In: *Environmental Pollution* 248 (May 2019), pp. 101–113. DOI: [10.1016/j.envpol.2019.02.018](https://doi.org/10.1016/j.envpol.2019.02.018).

- [75] B. E. Blake et al. “Associations between longitudinal serum perfluoroalkyl substance (PFAS) levels and measures of thyroid hormone, kidney function, and body mass index in the Fernald Community Cohort”. In: *Environmental Pollution* 242 (Nov. 2018), pp. 894–904. DOI: [10.1016/j.envpol.2018.07.042](https://doi.org/10.1016/j.envpol.2018.07.042).
- [76] T. Guillette et al. “Elevated levels of per- and polyfluoroalkyl substances in Cape Fear River Striped Bass (*Morone saxatilis*) are associated with biomarkers of altered immune and liver function”. In: *Environment International* 136 (May 2020), p. 105358. DOI: [10.1016/j.envint.2019.105358](https://doi.org/10.1016/j.envint.2019.105358).
- [77] OECD. *Series on Risk Management*. May 2018 [Online].
- [78] I. T. Cousins et al. “Strategies for grouping per- and polyfluoroalkyl substances (PFAS) to protect human and environmental health”. In: *Environmental Science: Processes Impacts* 22.7 (2020), pp. 1444–1460. DOI: [10.1039/d0em00147c](https://doi.org/10.1039/d0em00147c).
- [79] C. Hansch and T. Fujita. “ p - σ - π Analysis. A Method for the Correlation of Biological Activity and Chemical Structure”. In: *Journal of the American Chemical Society* 86.8 (Apr. 1964), pp. 1616–1626. DOI: [10.1021/ja01062a035](https://doi.org/10.1021/ja01062a035).
- [80] A. Cherkasov et al. “QSAR Modeling: Where Have You Been? Where Are You Going To?” In: *Journal of Medicinal Chemistry* 57.12 (Jan. 2014), pp. 4977–5010. DOI: [10.1021/jm4004285](https://doi.org/10.1021/jm4004285).
- [81] B. J. Neves et al. “QSAR-Based Virtual Screening: Advances and Applications in Drug Discovery”. In: *Frontiers in Pharmacology* 9 (Nov. 2018). DOI: [10.3389/fphar.2018.01275](https://doi.org/10.3389/fphar.2018.01275).
- [82] A. Raza et al. “A Machine Learning Approach for Predicting Defluorination of Per- and Polyfluoroalkyl Substances (PFAS) for Their Efficient Treatment and Removal”. In: *Environmental Science & Technology Letters* 6.10 (Sept. 2019), pp. 624–629. DOI: [10.1021/acs.estlett.9b00476](https://doi.org/10.1021/acs.estlett.9b00476).
- [83] W. Cheng and C. A. Ng. “Using Machine Learning to Classify Bioactivity for 3486 Per- and Polyfluoroalkyl Substances (PFASs) from the OECD List”. In: *Environmental Science & Technology* 53.23 (Oct. 2019), pp. 13970–13980. DOI: [10.1021/acs.est.9b04833](https://doi.org/10.1021/acs.est.9b04833).
- [84] Y. Wang et al. “PubChem BioAssay: 2014 update”. In: *Nucleic Acids Research* 42.D1 (Nov. 2013), pp. D1075–D1082. DOI: [10.1093/nar/gkt978](https://doi.org/10.1093/nar/gkt978).
- [85] S. G. Rohrer and K. Baumann. “Maximum Unbiased Validation (MUV) Data Sets for Virtual Screening Based on PubChem Bioactivity Data”. In: *Journal of Chemical Information and Modeling* 49.2 (Jan. 2009), pp. 169–184. DOI: [10.1021/ci8002649](https://doi.org/10.1021/ci8002649).

- [86] D. Krewski et al. “Toxicity Testing in the 21st Century: A Vision and a Strategy”. In: *Journal of Toxicology and Environmental Health, Part B* 13.2-4 (June 2010), pp. 51–138. DOI: [10.1080/10937404.2010.483176](https://doi.org/10.1080/10937404.2010.483176).
- [87] G. Subramanian et al. “Computational Modeling of β -Secretase 1 (BACE-1) Inhibitors Using Ligand Based Approaches”. In: *Journal of Chemical Information and Modeling* 56.10 (Oct. 2016), pp. 1936–1949. DOI: [10.1021/acs.jcim.6b00290](https://doi.org/10.1021/acs.jcim.6b00290).
- [88] I. F. Martins et al. “A Bayesian Approach to *in Silico* Blood-Brain Barrier Penetration Modeling”. In: *Journal of Chemical Information and Modeling* 52.6 (June 2012), pp. 1686–1697. DOI: [10.1021/ci300124c](https://doi.org/10.1021/ci300124c).
- [89] D. Rogers and M. Hahn. “Extended-Connectivity Fingerprints”. In: *Journal of Chemical Information and Modeling* 50.5 (Apr. 2010), pp. 742–754. DOI: [10.1021/ci100050t](https://doi.org/10.1021/ci100050t).
- [90] G. Landrum et al. *rdkit/rdkit: 2020_03_1 (Q1 2020) Release*. Accessed 2021-06-29. 2020. DOI: [10.5281/ZENODO.3732262](https://doi.org/10.5281/ZENODO.3732262).
- [91] P. J. Rousseeuw. “Silhouettes: A graphical aid to the interpretation and validation of cluster analysis”. In: *Journal of Computational and Applied Mathematics* 20 (Nov. 1987), pp. 53–65. DOI: [10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).
- [92] G. M. Morris et al. “AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility”. In: *Journal of Computational Chemistry* 30.16 (Dec. 2009), pp. 2785–2791. DOI: [10.1002/jcc.21256](https://doi.org/10.1002/jcc.21256).
- [93] A. B. Raies and V. B. Bajic. “*In silico* toxicology: computational methods for the prediction of chemical toxicity”. In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 6.2 (Jan. 2016), pp. 147–172. DOI: [10.1002/wcms.1240](https://doi.org/10.1002/wcms.1240).
- [94] H. Yang et al. “Computational Approaches to Identify Structural Alerts and Their Applications in Environmental Toxicology and Drug Discovery”. In: *Chemical Research in Toxicology* 33.6 (Feb. 2020), pp. 1312–1322. DOI: [10.1021/acs.chemrestox.0c00006](https://doi.org/10.1021/acs.chemrestox.0c00006).
- [95] M. Davies et al. “ChEMBL web services: streamlining access to drug discovery data and utilities”. In: *Nucleic Acids Research* 43.W1 (Apr. 2015), W612–W620. DOI: [10.1093/nar/gkv352](https://doi.org/10.1093/nar/gkv352).
- [96] X. Cheng and C. D. Klaassen. “Perfluorocarboxylic Acids Induce Cytochrome P450 Enzymes in Mouse Liver through Activation of PPAR- α and CAR Transcription Factors”. In: *Toxicological Sciences* 106.1 (July 2008), pp. 29–36. DOI: [10.1093/toxsci/kfn147](https://doi.org/10.1093/toxsci/kfn147).

- [97] J. O. Miners and D. J. Birkett. “Cytochrome P450C9: an enzyme of major importance in human drug metabolism”. In: *British Journal of Clinical Pharmacology* 45.6 (June 1998), pp. 525–538. DOI: [10.1046/j.1365-2125.1998.00721.x](https://doi.org/10.1046/j.1365-2125.1998.00721.x).
- [98] J. Ashburner and S. Klöppel. “Multivariate models of inter-subject anatomical variability”. In: *NeuroImage* 56.2 (May 2011), pp. 422–439. DOI: [10.1016/j.neuroimage.2010.03.059](https://doi.org/10.1016/j.neuroimage.2010.03.059).
- [99] C. Chu et al. “Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images”. In: *NeuroImage* 60.1 (Mar. 2012), pp. 59–70. DOI: [10.1016/j.neuroimage.2011.11.066](https://doi.org/10.1016/j.neuroimage.2011.11.066).
- [100] *OECD portal on per and poly fluorinated chemicals - OECD portal on per and poly fluorinated chemicals*. Accessed 2021-07-01.
- [101] A. Kumar, Z. A. Ali, and B. M. Wong. “Efficient Predictions of Formation Energies and Convex Hulls from Density Functional Tight Binding Calculations”. In: (). Submitted (2022).
- [102] C. Freysoldt et al. “First-principles calculations for point defects in solids”. In: *Reviews of modern physics* 86.1 (2014), p. 253.
- [103] M. Ai et al. “Role of vacancies in photocatalysis: a review of recent progress”. In: *Chemistry—An Asian Journal* 15.22 (2020), pp. 3599–3619.
- [104] C. G. Van de Walle and A. Janotti. “Advances in electronic structure methods for defects and impurities in solids”. In: *physica status solidi (b)* 248.1 (2011), pp. 19–27.
- [105] R. M. Nieminen. “Issues in first-principles calculations for defects in semiconductors and oxides”. In: *Modelling and Simulation in Materials Science and Engineering* 17.8 (Nov. 2009), p. 084001.
- [106] P. Hohenberg and W. Kohn. “Inhomogeneous Electron Gas”. In: *Phys. Rev.* 136 (3B Nov. 1964), B864–B871.
- [107] W. Kohn and L. J. Sham. “Self-Consistent Equations Including Exchange and Correlation Effects”. In: *Phys. Rev.* 140 (4A Nov. 1965), A1133–A1138.
- [108] T. Frauenheim et al. “Density-Functional-Based Construction of Transferable Nonorthogonal Tight-Binding Potentials for Si and SiH”. In: *Phys. Rev. B* 52 (15 Oct. 1995), pp. 11492–11501.

- [109] D. Porezag et al. “Construction of Tight-Binding-Like Potentials on the Basis of Density-Functional Theory: Application to Carbon”. In: *Phys. Rev. B* 51 (19 May 1995), pp. 12947–12957.
- [110] J. Widany et al. “Density-Functional-Based Construction of Transferable Nonorthogonal Tight-Binding Potentials for B, N, BN, BH, and NH”. In: *Phys. Rev. B* 53 (8 Feb. 1996), pp. 4443–4452.
- [111] S. S. Yamijala, M. B. Oviedo, and B. M. Wong. “Density Functional Tight Binding Calculations for Probing Electronic-Excited States of Large Systems”. In: *Reviews in Computational Chemistry, Volume 32*. John Wiley & Sons, Ltd, 2022. Chap. 2, pp. 45–79.
- [112] N. V. Ilawe, M. B. Oviedo, and B. M. Wong. “Real-Time Quantum Dynamics of Long-Range Electronic Excitation Transfer in Plasmonic Nanoantennas”. In: *J. Chem. Theory Comput.* 13.8 (2017), pp. 3442–3454.
- [113] N. V. Ilawe, M. B. Oviedo, and B. M. Wong. “Effect of Quantum Tunneling on the Efficiency of Excitation Energy Transfer in Plasmonic Nanoparticle Chain Waveguides”. In: *J. Mater. Chem. C* 6.22 (2018), pp. 5857–5864.
- [114] M. B. Oviedo and B. M. Wong. “Real-Time Quantum Dynamics Reveals Complex, Many-Body Interactions in Solvated Nanodroplets”. In: *J. Chem. Theory Comput.* 12.4 (2016), pp. 1862–1871.
- [115] S. I. Allec et al. “Heterogeneous CPU+ GPU-enabled simulations for DFTB molecular dynamics of large chemical and biological systems”. In: *Journal of chemical theory and computation* 15.5 (2019), pp. 2807–2815.
- [116] S. M. Islam and P.-N. Roy. “Performance of the SCC-DFTB Model for Description of Five-Membered Ring Carbohydrate Conformations: Comparison to Force Fields, High-Level Electronic Structure Methods, and Experiment”. In: *J. Chem. Theory Comput.* 8.7 (2012). PMID: 26588973, pp. 2412–2423.
- [117] K. Leong et al. “Energy and Charge Transfer by Donor–Acceptor Pairs Confined in a Metal–Organic Framework: A Spectroscopic and Computational Investigation”. In: *J. Mater. Chem. A* 2 (10 2014), pp. 3389–3398.
- [118] J. M. Rodríguez-Borbón et al. “Field Programmable Gate Arrays for Enhancing the Speed and Energy Efficiency of Quantum Dynamics Simulations”. In: *Journal of Chemical Theory and Computation* 16.4 (2020), pp. 2085–2098. DOI: [10.1021/acs.jctc.9b01284](https://doi.org/10.1021/acs.jctc.9b01284).

- [119] A. Van der Ven et al. “Linking the electronic structure of solids to their thermodynamic and kinetic properties”. In: *Mathematics and computers in simulation* 80.7 (2010), pp. 1393–1410.
- [120] A. Van der Ven et al. “Nondilute diffusion from first principles: Li diffusion in Li_xTiS_2 ”. In: *Physical Review B* 78.10 (2008), p. 104306.
- [121] A. Drozdov et al. “Conventional superconductivity at 203 kelvin at high pressures in the sulfur hydride system”. In: *Nature* 525.7567 (2015), pp. 73–76.
- [122] T. Ishikawa, T. Miyake, and K. Shimizu. “Materials informatics based on evolutionary algorithms: Application to search for superconducting hydrogen compounds”. In: *Physical Review B* 100.17 (2019), p. 174506.
- [123] B. Guigue, A. Marizy, and P. Loubeyre. “Direct synthesis of pure H_3S from S and H elements: No evidence of the cubic superconducting phase up to 160 GPa”. In: *Physical Review B* 95.2 (2017), p. 020104.
- [124] M. Somayazulu et al. “Evidence for superconductivity above 260 K in lanthanum superhydride at megabar pressures”. In: *Physical review letters* 122.2 (2019), p. 027001.
- [125] A. Drozdov et al. “Superconductivity at 250 K in lanthanum hydride under high pressures”. In: *Nature* 569.7757 (2019), pp. 528–531.
- [126] C. R. Weinberger et al. “Ab initio investigations of the phase stability in group IVB and VB transition metal nitrides”. In: *Computational Materials Science* 138 (2017), pp. 333–345.
- [127] X.-X. Yu, C. R. Weinberger, and G. B. Thompson. “Ab initio investigations of the phase stability in group IVB and VB transition metal carbides”. In: *Computational Materials Science* 112 (2016), pp. 318–326.
- [128] A. R. Oganov and C. W. Glass. “Crystal structure prediction using ab initio evolutionary techniques: Principles and applications”. In: *The Journal of chemical physics* 124.24 (2006), p. 244704.
- [129] A. R. Oganov et al. “Evolutionary crystal structure prediction as a method for the discovery of minerals and materials”. In: *Reviews in Mineralogy and Geochemistry* 71.1 (2010), pp. 271–298.
- [130] Y. Wang et al. “Crystal structure prediction via particle-swarm optimization”. In: *Physical Review B* 82.9 (2010), p. 094116.

- [131] Y. Wang et al. “CALYPSO: A method for crystal structure prediction”. In: *Computer Physics Communications* 183.10 (2012), pp. 2063–2070.
- [132] D.-H. Yoon and I. E. Reimanis. “A review on the joining of SiC for high-temperature applications”. In: *Journal of the Korean Ceramic Society* 57.3 (2020), pp. 246–270.
- [133] X. She et al. “Review of silicon carbide power devices and their applications”. In: *IEEE Transactions on Industrial Electronics* 64.10 (2017), pp. 8193–8205.
- [134] R. Madar. “Silicon carbide in contention”. In: *Nature* 430.7003 (2004), pp. 974–975.
- [135] J. Casady and R. W. Johnson. “Status of silicon carbide (SiC) as a wide-bandgap semiconductor for high-temperature applications: A review”. In: *Solid-State Electronics* 39.10 (1996), pp. 1409–1422.
- [136] C. Wöll. “The chemistry and physics of zinc oxide surfaces”. In: *Progress in surface science* 82.2-3 (2007), pp. 55–120.
- [137] M. Elstner et al. “Self-Consistent-Charge Density-Functional Tight-Binding Method for Simulations of Complex Materials Properties”. In: *Phys. Rev. B* 58 (11 Sept. 1998), pp. 7260–7268.
- [138] A. Santoro and A. D. Mighell. “Properties of crystal lattices: the derivative lattices and their determination”. In: *Acta Crystallographica Section A* 28.3 (May 1972), pp. 284–287.
- [139] A. Santoro and A. D. Mighell. “Coincidence-site lattices”. In: *Acta Crystallographica Section A* 29.2 (Mar. 1973), pp. 169–175.
- [140] G. L. W. Hart and R. W. Forcade. “Algorithm for generating derivative structures”. In: *Phys. Rev. B* 77 (22 June 2008), p. 224115.
- [141] L. Ferreira, S.-H. Wei, and A. Zunger. “Stability, electronic structure, and phase diagrams of novel inter-semiconductor compounds”. In: *The International Journal of Supercomputing Applications* 5.1 (1991), pp. 34–56.
- [142] G. L. Hart, L. J. Nelson, and R. W. Forcade. “Generating derivative structures at a fixed concentration”. In: *Computational Materials Science* 59 (2012), pp. 101–107.
- [143] B. Puchala and A. Van der Ven. “Thermodynamics of the Zr-O system from first-principles calculations”. In: *Phys. Rev. B* 88 (9 Sept. 2013), p. 094108.
- [144] G. Kresse and J. Hafner. “Ab initio molecular dynamics for liquid metals”. In: *Phys. Rev. B* 47 (1 Jan. 1993), pp. 558–561.

- [145] G. Kresse and J. Furthmüller. “Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set”. In: *Phys. Rev. B* 54 (16 Oct. 1996), pp. 11169–11186.
- [146] P. E. Blöchl. “Projector augmented-wave method”. In: *Phys. Rev. B* 50 (24 Dec. 1994), pp. 17953–17979.
- [147] G. Kresse and D. Joubert. “From ultrasoft pseudopotentials to the projector augmented-wave method”. In: *Phys. Rev. B* 59 (3 Jan. 1999), pp. 1758–1775.
- [148] J. P. Perdew, K. Burke, and M. Ernzerhof. “Generalized Gradient Approximation Made Simple”. In: *Phys. Rev. Lett.* 77 (18 Oct. 1996), pp. 3865–3868.
- [149] S. Liu, E. Martínez, and J. LLorca. “Prediction of the Al-rich part of the Al-Cu phase diagram using cluster expansion and statistical mechanics”. In: *Acta Materialia* 195 (2020), pp. 317–326.
- [150] D. Porezag et al. “Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon”. In: *Physical Review B* 51.19 (1995), p. 12947.
- [151] G. Seifert, D. Porezag, and T. Frauenheim. “Calculations of molecules, clusters, and solids with a simplified LCAO-DFT-LDA scheme”. In: *International journal of quantum chemistry* 58.2 (1996), pp. 185–192.
- [152] T. Frauenheim et al. “A self-consistent charge density-functional based tight-binding method for predictive materials simulations in physics, chemistry and biology”. In: *physica status solidi (b)* 217.1 (2000), pp. 41–62.
- [153] T. Frauenheim et al. “Atomistic simulations of complex materials: ground-state and excited-state properties”. In: *Journal of Physics: Condensed Matter* 14.11 (2002), p. 3015.
- [154] A. F. Oliveira et al. “Density-functional based tight-binding: an approximate DFT method”. In: *Journal of the Brazilian Chemical Society* 20.7 (2009), pp. 1193–1205.
- [155] G. Seifert and J.-O. Joswig. “Density-functional tight binding— An approximate density-functional theory method”. In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 2.3 (2012), pp. 456–465.
- [156] M. Gaus, Q. Cui, and M. Elstner. “DFTB3: extension of the self-consistent-charge density-functional tight-binding method (SCC-DFTB)”. In: *Journal of chemical theory and computation* 7.4 (2011), pp. 931–948.

- [157] M. Bezi Javan. “Optical properties of SiC nanocages: ab initio study”. In: *Applied Physics A* 113.1 (2013), pp. 105–113.
- [158] R. Khalil, F. Hussain, and N. A. Niaz. “Studies of Molecular Dynamics and Electronic Structure in Cubic-SiC by Using Density Functional Tight Binding Approach”. In: *Journal of Materials and Physical Sciences* 2.2 (2021), pp. 88–94.
- [159] O. Arayawut, T. Kerdcharoen, and C. Wongchoosuk. “Structures, Electronic Properties, and Gas Permeability of 3D Pillared Silicon Carbide Nanostructures”. In: *Nanomaterials* 12.11 (2022), p. 1869.
- [160] M. Hellström et al. “An SCC-DFTB repulsive potential for various ZnO polymorphs and the ZnO–water system”. In: *The Journal of Physical Chemistry C* 117.33 (2013), pp. 17004–17015.
- [161] V. N. Tuoc et al. “Density functional theory based tight binding study on theoretical prediction of low-density nanoporous phases ZnO semiconductor materials”. In: *Journal of Physics: Conference Series*. Vol. 726. 1. IOP Publishing. 2016, p. 012022.
- [162] A. W. Huran et al. “Efficient automatized density-functional tight-binding parametrizations: application to group IV elements”. In: *Journal of Chemical Theory and Computation* 14.6 (2018), pp. 2947–2954.
- [163] C. Köhler et al. “Theoretical investigation of carbon defects and diffusion in α -quartz”. In: *Physical Review B* 64.8 (2001), p. 085333.
- [164] N. H. Moreira et al. “Toward an accurate density-functional tight-binding description of zinc-containing compounds”. In: *Journal of chemical theory and computation* 5.3 (2009), pp. 605–614.
- [165] V. S. Naumov et al. “Structural, electronic, and thermodynamic properties of TiO₂/organic clusters: performance of DFTB method with different parameter sets”. In: *International Journal of Quantum Chemistry* 121.2 (2021), e26427.
- [166] T. A. Niehaus et al. “Towards a simplified description of thermoelectric materials: accuracy of approximate density functional theory for phonon dispersions”. In: *Journal of Physics: Condensed Matter* 31.39 (2019), p. 395901.
- [167] S. Wang et al. “New ab initio based pair potential for accurate simulation of phase transitions in ZnO”. In: *The Journal of Physical Chemistry C* 118.20 (2014), pp. 11050–11061.
- [168] M. P. Molepo and D. P. Joubert. “Computational study of the structural phases of ZnO”. In: *Physical Review B* 84.9 (2011), p. 094110.

- [169] P. Erhart, A. Klein, and K. Albe. “First-principles study of the structure and stability of oxygen defects in zinc oxide”. In: *Physical Review B* 72.8 (2005), p. 085213.
- [170] C. G. Van de Walle. “Hydrogen as a cause of doping in zinc oxide”. In: *Physical review letters* 85.5 (2000), p. 1012.
- [171] E.-C. Lee et al. “Compensation mechanism for N acceptors in ZnO”. In: *Physical Review B* 64.8 (2001), p. 085120.
- [172] P. Erhart and K. Albe. “First-principles study of migration mechanisms and diffusion of oxygen in zinc oxide”. In: *Physical Review B* 73.11 (2006), p. 115207.
- [173] T. R. Paudel and W. R. Lambrecht. “First-principles calculation of the O vacancy in ZnO: A self-consistent gap-corrected approach”. In: *Physical Review B* 77.20 (2008), p. 205202.
- [174] S. S. R. K. C. Yamijala, Z. A. Ali, and B. M. Wong. “Acceleration vs Accuracy: Influence of Basis Set Quality on the Mechanism and Dynamics Predicted by Ab Initio Molecular Dynamics”. In: *The Journal of Physical Chemistry C* 123.41 (Sept. 2019), pp. 25113–25120. DOI: [10.1021/acs.jpcc.9b03554](https://doi.org/10.1021/acs.jpcc.9b03554).
- [175] M. Bonn et al. “Phonon- versus Electron-Mediated Desorption and Oxidation of CO on Ru(0001)”. In: *Science* 285 (1999), 1042–1045.
- [176] J. Park, L. Baker, and G. Somorjai. “Role of Hot Electrons and Metal-Oxide Interfaces in Surface Chemistry and Catalytic Reactions”. In: *Chem. Rev* 115 (2015), 2781–2817.
- [177] J. Gavnholt et al. “Hot-Electron-Assisted Femtochemistry at Surfaces: A Time-Dependent Density Functional Theory Approach”. In: *Phys. Rev. B: Condens. Matter Mater. Phys* (2009), p. 79. DOI: [10.1103/physrevb.79.195405](https://doi.org/10.1103/physrevb.79.195405).
- [178] J. Wintterlin, R. Schuster, and G. Ertl. “Existence of a “Hot” Atom Mechanism for the Dissociation of O₂ on Pt(111)”. In: *Phys. Rev. Lett* 77 (1996), pp. 123–126.
- [179] M. Blanco-Rey et al. “Electronic Friction Dominates Hydrogen Hot-Atom Relaxation on Pd(100)”. In: *Phys. Rev. Lett* 112 (2014), p. 103203.
- [180] R. Pétuya et al. “Scattering of Atomic Hydrogen Off a H-Covered W(110) Surface: Hot-Atom versus Eley–Rideal Abstraction Dynamics”. In: *J. Phys. Chem. C* 119 (2015), pp. 3171–3179.
- [181] M. Brongersma, N. Halas, and P. Nordlander. “Plasmon-Induced Hot Carrier Science and Technology”. In: *Nat. Nanotechnol* 10 (2015), pp. 25–34.

- [182] G. Liu et al. “Promoting Active Species Generation by Plasmon-Induced Hot-Electron Excitation for Efficient Electrocatalytic Oxygen Evolution”. In: *J. Am. Chem. Soc.* 138 (2016), pp. 9128–9136.
- [183] H. Sakamoto et al. “Hot-Electron-Induced Highly Efficient O₂ Activation by Pt Nanoparticles Supported on Ta₂O₅ Driven by Visible Light”. In: *J. Am. Chem. Soc.* 137 (2015), pp. 9324–9332.
- [184] T. Avanesian and P. Christopher. “Adsorbate Specificity in Hot Electron Driven Photochemistry on Catalytic Metal Surfaces”. In: *J. Phys. Chem. C* 118 (2014), pp. 28017–28031.
- [185] J. Park et al. “Hot-Electron-Mediated Surface Chemistry: Toward Electronic Control of Catalytic Activity”. In: *Acc. Chem. Res.* 48 (2015), 2475–2483.
- [186] N. Bork, V. Loukonen, and H. Vehkamäki. “Reactions and Reaction Rate of Atmospheric SO₂ and O₃⁻ (H₂O)_n Collisions via Molecular Dynamics Simulations”. In: *J. Phys. Chem. A* 117 (2013), pp. 3143–3148.
- [187] N. Tsona et al. “A Closure Study of the Reaction between Sulfur Dioxide and the Sulfate Radical Ion from First-Principles Molecular Dynamics Simulations”. In: *J. Phys. Chem. A* 120 (2016), pp. 1046–1050.
- [188] R. Olsen et al. “Comparison of Methods for Finding Saddle Points without Knowledge of the Final States”. In: *J. Chem. Phys.* 121 (2004), pp. 9776–9792.
- [189] G. Henkelman, G. Jóhannesson, and H. Jónsson. “Methods for Finding Saddle Points and Minimum Energy Paths”. In: *Progress in Theoretical Chemistry and Physics*, pp. 269–302.
- [190] X. Zhou, L. Zhang, and B. Jiang. “Hot-Atom-Mediated Dynamical Displacement of CO Adsorbed on Cu(111) by Incident H Atoms: An Ab Initio Molecular Dynamics Study”. In: *J. Phys. Chem. C* 122 (2018), pp. 15485–15493.
- [191] Z.-T. Wang et al. “Probing Equilibrium of Molecular and Deprotonated Water on TiO₂(110)”. In: *Proceedings of the National Academy of Sciences 2017*. Vol. 114, pp. 1801–1805.
- [192] W. Lin, K. Stocker, and G. Schatz. “Mechanisms of Hydrogen-Assisted CO₂ Reduction on Nickel”. In: *J. Am. Chem. Soc.* 139 (2017), pp. 4663–4666.
- [193] A. Ashwell et al. “Hydrogenation of CO to Methanol on Ni(110) through Subsurface Hydrogen”. In: *J. Am. Chem. Soc.* 139 (2017), pp. 17582–17589.

- [194] W. Lin and G. Schatz. “Mechanisms of Formaldehyde and C₂ Formation from Methy-
lene Reacting with CO₂ Adsorbed on Ni(110)”. In: *J. Phys. Chem. C* 122 (2018),
pp. 13827–13833.
- [195] A. Groß. “Ab Initio Molecular Dynamics Simulations of the O₂ /Pt(1 1 1) Interac-
tion”. en. In: *Catal. Today* 260 (2016), pp. 60–65.
- [196] X. Zhou et al. “Ab Initio Molecular Dynamics Study of Dissociative Chemisorption
and Scattering of CO₂ on Ni(100): Reactivity, Energy Transfer, Steering Dynamics,
and Lattice Effects”. In: *J. Phys. Chem. C* 121 (2017), pp. 5594–5602.
- [197] L. Li and X. Zeng. “Direct Simulation Evidence of Generation of Oxygen Vacancies
at the Golden Cage Au₁₆ and TiO₂ (110) Interface for CO Oxidation”. In: *J. Am.
Chem. Soc* 136 (2014), pp. 15857–15860.
- [198] L. Li, H. Li, and X. Zeng. “Structure Transition of Au₁₈ from Pyramidal to a Hollow-
Cage during Soft-Landing onto a TiO₂(110) Surface”. In: *Chem. Commun* 51 (2015),
pp. 9535–9538.
- [199] B. Hirshberg and R. Benny Gerber. “Formation of Carbonic Acid in Impact of CO₂
on Ice and Water”. In: *J. Phys. Chem. Lett* 7 (2016), pp. 2905–2909.
- [200] J. VandeVondele et al. “Quickstep: Fast and Accurate Density Functional Calcula-
tions Using a Mixed Gaussian and Plane Waves Approach”. In: *Comput. Phys.
Commun* 167 (2005), pp. 103–128.
- [201] V. Blum et al. “Ab Initio Molecular Simulations with Numeric Atom-Centered Or-
bitals”. In: *Comput. Phys. Commun* 180 (2009), pp. 2175–2196.
- [202] G. Kresse and J. Furthmüller. “Efficiency of Ab-Initio Total Energy Calculations for
Metals and Semiconductors Using a Plane-Wave Basis Set”. In: *Comput. Mater. Sci*
6 (1996), 15–50.
- [203] Y. Zhao and D. Truhlar. “Density Functional Calculations of E₂ and S_N2 Reactions:
Effects of the Choice of Density Functional, Basis Set, and Self-Consistent Itera-
tions”. In: *J. Chem. Theory Comput* 6 (2010), 1104–1108. DOI: [10.1021/ct100082z](https://doi.org/10.1021/ct100082z).
- [204] B. Narendrapurapu et al. “Investigating the Effects of Basis Set on Metal–Metal and
Metal–Ligand Bond Distances in Stable Transition Metal Carbonyls: Performance of
Correlation Consistent Basis Sets with 35 Density Functionals”. In: *J. Chem. Theory
Comput* 9 (2013), 2930–2938.
- [205] T. Gavin Williams and A. Wilson. “Importance of the Quality of Metal and Ligand
Basis Sets in Transition Metal Species”. In: *J. Chem. Phys* 129 (2008), p. 054108.

- [206] X. Xu and D. Truhlar. “Accuracy of Effective Core Potentials and Basis Sets for Density Functional Calculations, Including Relativistic Effects, As Illustrated by Calculations on Arsenic Compounds”. In: *J. Chem. Theory Comput* 7 (2011), 2766–2779.
- [207] R. Steele. “Multiple-Timestep Ab Initio Molecular Dynamics Using an Atomic Basis Set Partitioning”. In: *J. Phys. Chem. A* 119 (2015), 12119–12130.
- [208] J. Li and S. Iyengar. “Ab Initio Molecular Dynamics Using Recursive, Spatially Separated, Overlapping Model Subsystems Mixed within an ONIOM-Based Fragmentation Energy Extrapolation Technique”. In: *J. Chem. Theory Comput* 11 (2015), 3978–3991.
- [209] C. Haycraft, J. Li, and S. Iyengar. ““On-the-Fly”, Born-Oppenheimer and Car-Parrinello-Type Dynamics with Coupled Cluster Accuracy through Fragment Based Electronic Structure”. In: *J. Chem. Theory Comput* 13 (2017), 1887–1901.
- [210] R. Steele, R. DiStasio Jr, and M. Head-Gordon. “Non-Covalent Interactions with Dual-Basis Methods: Pairings for Augmented Basis Sets”. In: *J. Chem. Theory Comput* 5 (2009), 1560–1572.
- [211] R. Steele and M. Head-Gordon. “Dual-Basis Self-Consistent Field Methods: 6-31G* Calculations with a Minimal 6-4G Primary Basis”. In: *Molecular Physics* (2007), 2455–2473. DOI: [10.1080/00268970701519754](https://doi.org/10.1080/00268970701519754).
- [212] R. Steele, M. Head-Gordon, and J. Tully. “Ab Initio Molecular Dynamics with Dual Basis Set Methods”. In: *J. Phys. Chem. A* 114 (2010), 11853–11860.
- [213] T. Ricard and S. Iyengar. “Efficiently Capturing Weak Interactions in Ab Initio Molecular Dynamics with on-the-Fly Basis Set Extrapolation”. In: *J. Chem. Theory Comput* 14 (2018), 5535–5552.
- [214] J. Rodriguez et al. “Water Gas Shift Reaction on Cu and Au Nanoparticles Supported on CeO₂(111) and ZnO(000): Intrinsic Activity and Importance of Support Interactions”. In: *Angew. Chem. Int. Ed Engl* 119 (2007), 1351–1354.
- [215] J. Tabatabaei, B. Sakakini, and K. Waugh. “On the Mechanism of Methanol Synthesis and the Water-Gas Shift Reaction on ZnO”. In: *Catal. Letters* 110 (2006), 77–84.
- [216] H. Nakatsuji et al. “Theoretical Study of the Chemisorption and Surface Reaction of HCOOH on a ZnO(10 $\bar{1}$ 0) Surface”. In: *J. Phys. Chem* 100 (1996), 694–700. DOI: [10.1021/jp9504581](https://doi.org/10.1021/jp9504581).

- [217] D. Wang et al. “The Effect of Zinc Oxide in Raney Copper Catalysts on Methanol Synthesis, Water Gas Shift, and Methanol Steam Reforming Reaction”. In: *Studies in Surface Science and Catalysis*. 1996, 1379–1387.
- [218] C. Álvarez Galván et al. “Reverse Water-Gas Shift Reaction at the Cu/ZnO Interface: Influence of the Cu/Zn Ratio on Structure-Activity Correlations”. In: *Appl. Catal. B* 195 (2016), 104–111.
- [219] J. Perdew, K. Burke, and M. Ernzerhof. “Generalized Gradient Approximation Made Simple [Phys”. In: *Rev. Lett* 77 (1996), p. 3865.
- [220] J. VandeVondele and J. Hutter. “Gaussian Basis Sets for Accurate Calculations on Molecular Systems in Gas and Condensed Phases”. In: *J. Chem. Phys* 127 (2007), p. 114105.
- [221] S. Goedecker, M. Teter, and J. Hutter. “Separable Dual-Space Gaussian Pseudopotentials”. In: *Phys. Rev. B: Condens. Matter Mater. Phys* 54 (1996), 1703–1710.
- [222] M. Krack. “Pseudopotentials for H to Kr Optimized for Gradient-Corrected Exchange-Correlation Functionals”. In: *Theor. Chem. Acc* 114 (2005), 145–152.
- [223] H. Shi et al. “Adsorption and Diffusion of CO on Clean and CO₂-Precovered ZnO(1010̄)”. In: *J. Phys. Chem. C* 122 (2018), 8919–8924.
- [224] S. Grimme, S. Ehrlich, and L. Goerigk. “Effect of the Damping Function in Dispersion Corrected Density Functional Theory”. In: *J. Comput. Chem* 32 (2011), 1456–1465.
- [225] B. Meyer and D. Marx. “First-Principles Study of CO Adsorption on ZnO Surfaces”. In: *J. Phys. Condens. Matter* 15 (2003), 89– 94.
- [226] C. Lian, Z. A. Ali, and B. M. Wong. “Charge Density Wave Hampers Exciton Condensation in 1T – TiSe₂.” In: *Phys. Rev. B* 100 (20 Nov. 2019), p. 205423. DOI: [10.1103/physrevb.100.205423](https://doi.org/10.1103/physrevb.100.205423).
- [227] C. Lian et al. “Ultrafast charge ordering by self-amplified exciton–phonon dynamics in TiSe₂”. In: *Nature Communications* 11.1 (Jan. 2020). DOI: [10.1038/s41467-019-13672-7](https://doi.org/10.1038/s41467-019-13672-7).
- [228] D. Snoke and G. M. Kavoulakis. “Bose–Einstein condensation of excitons in Cu₂O: progress over 30 years”. In: *Reports on Progress in Physics* 77.11 (Oct. 2014), p. 116501. DOI: [10.1088/0034-4885/77/11/116501](https://doi.org/10.1088/0034-4885/77/11/116501).

- [229] K. Seki et al. “Excitonic Bose-Einstein condensation in Ta_2NiSe_5 above room temperature”. In: *Physical Review B* 90.15 (Oct. 2014). DOI: [10.1103/physrevb.90.155116](https://doi.org/10.1103/physrevb.90.155116).
- [230] D. W. Snoke, J. P. Wolfe, and A. Mysyrowicz. “Evidence for Bose-Einstein condensation of excitons in Cu_2O ”. In: *Physical Review B* 41.16 (June 1990), pp. 11171–11184. DOI: [10.1103/physrevb.41.11171](https://doi.org/10.1103/physrevb.41.11171).
- [231] K. Johnsen and G. M. Kavoulakis. “Probing Bose-Einstein Condensation of Excitons with Electromagnetic Radiation”. In: *Physical Review Letters* 86.5 (Jan. 2001), pp. 858–861. DOI: [10.1103/physrevlett.86.858](https://doi.org/10.1103/physrevlett.86.858).
- [232] D. W. Snoke, J. P. Wolfe, and A. Mysyrowicz. “Evidence for Bose-Einstein condensation of a two-component exciton gas”. In: *Physical Review Letters* 64.21 (May 1990), pp. 2543–2546. DOI: [10.1103/physrevlett.64.2543](https://doi.org/10.1103/physrevlett.64.2543).
- [233] R. Casella. “On the possibility of observing a Bose-Einstein condensation of excitons in CdS and CdSe”. In: *Journal of Physics and Chemistry of Solids* 24.1 (Jan. 1963), pp. 19–26. DOI: [10.1016/0022-3697\(63\)90037-4](https://doi.org/10.1016/0022-3697(63)90037-4).
- [234] J. L. Lin and J. P. Wolfe. “Bose-Einstein condensation of paraexcitons in stressed Cu_2O ”. In: *Physical Review Letters* 71.8 (Aug. 1993), pp. 1222–1225. DOI: [10.1103/physrevlett.71.1222](https://doi.org/10.1103/physrevlett.71.1222).
- [235] M. Combescot, O. Betbeder-Matibet, and R. Combescot. “Bose-Einstein Condensation in Semiconductors: The Key Role of Dark Excitons”. In: *Physical Review Letters* 99.17 (Oct. 2007). DOI: [10.1103/physrevlett.99.176403](https://doi.org/10.1103/physrevlett.99.176403).
- [236] G. M. Kavoulakis and A. Mysyrowicz. “Auger decay, spin exchange, and their connection to Bose-Einstein condensation of excitons in Cu_2O ”. In: *Physical Review B* 61.24 (June 2000), pp. 16619–16622. DOI: [10.1103/physrevb.61.16619](https://doi.org/10.1103/physrevb.61.16619).
- [237] P. Cudazzo et al. “Strong Charge-Transfer Excitonic Effects and the Bose-Einstein Exciton Condensate in Graphane”. In: *Physical Review Letters* 104.22 (June 2010). DOI: [10.1103/physrevlett.104.226804](https://doi.org/10.1103/physrevlett.104.226804).
- [238] Z. Jiang et al. “Realizing an intrinsic excitonic insulator by decoupling exciton binding energy from the minimum band gap”. In: *Physical Review B* 98.8 (Aug. 2018). DOI: [10.1103/physrevb.98.081408](https://doi.org/10.1103/physrevb.98.081408).
- [239] Z. Jiang et al. “Half-Excitonic Insulator: A Single-Spin Bose-Einstein Condensate”. In: *Physical Review Letters* 122.23 (June 2019). DOI: [10.1103/physrevlett.122.236402](https://doi.org/10.1103/physrevlett.122.236402).

- [240] L. V. Butov et al. “Towards Bose–Einstein condensation of excitons in potential traps”. In: *Nature* 417.6884 (May 2002), pp. 47–52. DOI: [10.1038/417047a](https://doi.org/10.1038/417047a).
- [241] J. M. Blatt, K. W. Böer, and W. Brandt. “Bose-Einstein Condensation of Excitons”. In: *Physical Review* 126.5 (June 1962), pp. 1691–1692. DOI: [10.1103/physrev.126.1691](https://doi.org/10.1103/physrev.126.1691).
- [242] J. P. Eisenstein and A. H. MacDonald. “Bose–Einstein condensation of excitons in bilayer electron systems”. In: *Nature* 432.7018 (Dec. 2004), pp. 691–694. DOI: [10.1038/nature03081](https://doi.org/10.1038/nature03081).
- [243] A. Mysyrowicz, D. W. Snoke, and J. P. Wolfe. “Progress on Bose-Einstein Condensation of Excitons”. In: *physica status solidi (b)* 159.1 (May 1990), pp. 387–401. DOI: [10.1002/pssb.2221590145](https://doi.org/10.1002/pssb.2221590145).
- [244] H. Deng et al. “Condensation of Semiconductor Microcavity Exciton Polaritons”. In: *Science* 298.5591 (Oct. 2002), pp. 199–202. DOI: [10.1126/science.1074464](https://doi.org/10.1126/science.1074464).
- [245] D. Porras et al. “Polariton dynamics and Bose-Einstein condensation in semiconductor microcavities”. In: *Physical Review B* 66.8 (Aug. 2002). DOI: [10.1103/physrevb.66.085304](https://doi.org/10.1103/physrevb.66.085304).
- [246] M. Richard et al. “Experimental evidence for nonequilibrium Bose condensation of exciton polaritons”. In: *Physical Review B* 72.20 (Nov. 2005). DOI: [10.1103/physrevb.72.201301](https://doi.org/10.1103/physrevb.72.201301).
- [247] F. P. Laussy et al. “Effects of Bose-Einstein condensation of exciton polaritons in microcavities on the polarization of emitted light”. In: *Physical Review B* 73.3 (Jan. 2006). DOI: [10.1103/physrevb.73.035315](https://doi.org/10.1103/physrevb.73.035315).
- [248] M. Wouters and I. Carusotto. “Excitations in a Nonequilibrium Bose-Einstein Condensate of Exciton Polaritons”. In: *Physical Review Letters* 99.14 (Oct. 2007). DOI: [10.1103/physrevlett.99.140402](https://doi.org/10.1103/physrevlett.99.140402).
- [249] J. Kasprzak et al. “Second-Order Time Correlations within a Polariton Bose-Einstein Condensate in a CdTe Microcavity”. In: *Physical Review Letters* 100.6 (Feb. 2008). DOI: [10.1103/physrevlett.100.067402](https://doi.org/10.1103/physrevlett.100.067402).
- [250] S. Utsunomiya et al. “Observation of Bogoliubov excitations in exciton-polariton condensates”. In: *Nature Physics* 4.9 (Aug. 2008), pp. 700–705. DOI: [10.1038/nphys1034](https://doi.org/10.1038/nphys1034).
- [251] K. G. Lagoudakis et al. “Quantized vortices in an exciton–polariton condensate”. In: *Nature Physics* 4.9 (Aug. 2008), pp. 706–710. DOI: [10.1038/nphys1051](https://doi.org/10.1038/nphys1051).

- [252] H. Deng, H. Haug, and Y. Yamamoto. “Exciton-polariton Bose-Einstein condensation”. In: *Reviews of Modern Physics* 82.2 (May 2010), pp. 1489–1537. DOI: [10.1103/revmodphys.82.1489](https://doi.org/10.1103/revmodphys.82.1489).
- [253] J. Kasprzak et al. “Bose-Einstein condensation of exciton polaritons”. In: *Nature* 443.7110 (Sept. 2006), pp. 409–414. DOI: [10.1038/nature05131](https://doi.org/10.1038/nature05131).
- [254] C. Rüegg et al. “Bose-Einstein condensation of the triplet states in the magnetic insulator TlCuCl_3 ”. In: *Nature* 423.6935 (May 2003), pp. 62–65. DOI: [10.1038/nature01617](https://doi.org/10.1038/nature01617).
- [255] T. Radu et al. “Bose-Einstein Condensation of Magnons in Cs_2CuCl_4 ”. In: *Physical Review Letters* 95.12 (Sept. 2005). DOI: [10.1103/physrevlett.95.127202](https://doi.org/10.1103/physrevlett.95.127202).
- [256] S. O. Demokritov et al. “Bose-Einstein condensation of quasi-equilibrium magnons at room temperature under pumping”. In: *Nature* 443.7110 (Sept. 2006), pp. 430–433. DOI: [10.1038/nature05117](https://doi.org/10.1038/nature05117).
- [257] T. Nikuni et al. “Bose-Einstein Condensation of Dilute Magnons in TlCuCl_3 ”. In: *Physical Review Letters* 84.25 (June 2000), pp. 5868–5871. DOI: [10.1103/physrevlett.84.5868](https://doi.org/10.1103/physrevlett.84.5868).
- [258] Y. Tserkovnyak et al. “Bose-Einstein condensation of magnons pumped by the bulk spin Seebeck effect”. In: *Physical Review B* 93.10 (Mar. 2016). DOI: [10.1103/physrevb.93.100402](https://doi.org/10.1103/physrevb.93.100402).
- [259] A. Kogar et al. “Signatures of exciton condensation in a transition metal dichalcogenide”. In: *Science* 358.6368 (Dec. 2017), pp. 1314–1317. DOI: [10.1126/science.aam6432](https://doi.org/10.1126/science.aam6432).
- [260] E. Hanamura and H. Haug. “Will a bose-condensed exciton gas be superfluid?” In: *Solid State Communications* 15.9 (Nov. 1974), pp. 1567–1570. DOI: [10.1016/0038-1098\(74\)90940-5](https://doi.org/10.1016/0038-1098(74)90940-5).
- [261] H. Haug and E. Hanamura. “Derivation of the two-fluid model for Bose-condensed excitons”. In: *Physical Review B* 11.9 (May 1975), pp. 3317–3328. DOI: [10.1103/physrevb.11.3317](https://doi.org/10.1103/physrevb.11.3317).
- [262] W. Kohn. “Excitonic Phases”. In: *Physical Review Letters* 19.8 (Aug. 1967), pp. 439–442. DOI: [10.1103/physrevlett.19.439](https://doi.org/10.1103/physrevlett.19.439).
- [263] D. Jérôme, T. M. Rice, and W. Kohn. “Excitonic Insulator”. In: *Physical Review* 158.2 (June 1967), pp. 462–475. DOI: [10.1103/physrev.158.462](https://doi.org/10.1103/physrev.158.462).

- [264] B. I. Halperin and T. M. Rice. “Possible Anomalies at a Semimetal-Semiconductor Transition”. In: *Reviews of Modern Physics* 40.4 (Oct. 1968), pp. 755–766. DOI: [10.1103/revmodphys.40.755](https://doi.org/10.1103/revmodphys.40.755).
- [265] R. F. Egerton. “Electron energy-loss spectroscopy in the TEM”. In: *Reports on Progress in Physics* 72.1 (Dec. 2008), p. 016502. DOI: [10.1088/0034-4885/72/1/016502](https://doi.org/10.1088/0034-4885/72/1/016502).
- [266] L. J. P. Ament et al. “Resonant inelastic x-ray scattering studies of elementary excitations”. In: *Reviews of Modern Physics* 83.2 (June 2011), pp. 705–767. DOI: [10.1103/revmodphys.83.705](https://doi.org/10.1103/revmodphys.83.705).
- [267] F. J. D. Salvo, D. E. Moncton, and J. V. Waszczak. “Electronic properties and superlattice formation in the semimetal TiSe_2 ”. In: *Physical Review B* 14.10 (Nov. 1976), pp. 4321–4328. DOI: [10.1103/physrevb.14.4321](https://doi.org/10.1103/physrevb.14.4321).
- [268] C. Monney et al. “Spontaneous exciton condensation in 1T – TiSe_2 : BCS-like approach”. In: *Physical Review B* 79.4 (Jan. 2009). DOI: [10.1103/physrevb.79.045116](https://doi.org/10.1103/physrevb.79.045116).
- [269] X. Y. Cui et al. “Direct evidence of band modification and suppression of superstructure in TiSe_2 upon Fe intercalation: An angle-resolved photoemission study”. In: *Physical Review B* 73.8 (Feb. 2006). DOI: [10.1103/physrevb.73.085111](https://doi.org/10.1103/physrevb.73.085111).
- [270] E. Morosan et al. “Superconductivity in Cu_xTiSe_2 ”. In: *Nature Physics* 2.8 (July 2006), pp. 544–550. DOI: [10.1038/nphys360](https://doi.org/10.1038/nphys360).
- [271] G. Li et al. “Semimetal-to-Semimetal Charge Density Wave Transition in 1T – TiSe_2 ”. In: *Physical Review Letters* 99.2 (July 2007). DOI: [10.1103/physrevlett.99.027404](https://doi.org/10.1103/physrevlett.99.027404).
- [272] S. Y. Li et al. “Single-Gap s -Wave Superconductivity near the Charge-Density-Wave Quantum Critical Point in Cu_xTiSe_2 ”. In: *Physical Review Letters* 99.10 (Sept. 2007). DOI: [10.1103/physrevlett.99.107001](https://doi.org/10.1103/physrevlett.99.107001).
- [273] H. Barath et al. “Quantum and Classical Mode Softening Near the Charge-Density-Wave-Superconductor Transition of Cu_xTiSe_2 ”. In: *Physical Review Letters* 100.10 (Mar. 2008). DOI: [10.1103/physrevlett.100.106402](https://doi.org/10.1103/physrevlett.100.106402).
- [274] A. F. Kusmartseva et al. “Pressure Induced Superconductivity in Pristine 1T – TiSe_2 ”. In: *Physical Review Letters* 103.23 (Nov. 2009). DOI: [10.1103/physrevlett.103.236401](https://doi.org/10.1103/physrevlett.103.236401).

- [275] J. Jeong et al. “Electronic structure study of Cu-doped 1T – TiSe₂ by angle-resolved photoemission spectroscopy”. In: *Physica C: Superconductivity and its Applications* 470 (Dec. 2010), S648–S650. DOI: [10.1016/j.physc.2009.11.123](https://doi.org/10.1016/j.physc.2009.11.123).
- [276] M. Zaberchik et al. “Possible evidence of a two-gap structure for the Cu_xTiSe₂ superconductor”. In: *Physical Review B* 81.22 (June 2010). DOI: [10.1103/physrevb.81.220505](https://doi.org/10.1103/physrevb.81.220505).
- [277] A. D. Hillier et al. “Probing the superconducting ground state near the charge density wave phase transition in Cu_{0.06}TiSe₂”. In: *Physical Review B* 81.9 (Mar. 2010). DOI: [10.1103/physrevb.81.092507](https://doi.org/10.1103/physrevb.81.092507).
- [278] N. Giang et al. “Superconductivity at 2.3 K in the misfit compound (PbSe)_{1.16}(TiSe₂)₂”. In: *Physical Review B* 82.2 (July 2010). DOI: [10.1103/physrevb.82.024503](https://doi.org/10.1103/physrevb.82.024503).
- [279] E. Morosan et al. “Multiple electronic transitions and superconductivity in Pd_xTiSe₂”. In: *Physical Review B* 81.9 (Mar. 2010). DOI: [10.1103/physrevb.81.094524](https://doi.org/10.1103/physrevb.81.094524).
- [280] M. Iavarone et al. “Evolution of the charge density wave state in Cu_xTiSe₂”. In: *Physical Review B* 85.15 (Apr. 2012). DOI: [10.1103/physrevb.85.155103](https://doi.org/10.1103/physrevb.85.155103).
- [281] J. Kačmarčík et al. “Heat capacity of single-crystal Cu_xTiSe₂ superconductors”. In: *Physical Review B* 88.2 (July 2013). DOI: [10.1103/physrevb.88.020507](https://doi.org/10.1103/physrevb.88.020507).
- [282] P. Husaníková et al. “Magnetization properties and vortex phase diagram of Cu_xTiSe₂ single crystals”. In: *Physical Review B* 88.17 (Nov. 2013). DOI: [10.1103/physrevb.88.174501](https://doi.org/10.1103/physrevb.88.174501).
- [283] R. Ganesh et al. “Theoretical Prediction of a Time-Reversal Broken Chiral Superconducting Phase Driven by Electronic Correlations in a TiSe₂ Single Layer”. In: *Physical Review Letters* 113.17 (Oct. 2014). DOI: [10.1103/physrevlett.113.177001](https://doi.org/10.1103/physrevlett.113.177001).
- [284] Y. I. Joe et al. “Emergence of charge density wave domain walls above the superconducting dome in 1T – TiSe₂”. In: *Nature Physics* 10.6 (Apr. 2014), pp. 421–425. DOI: [10.1038/nphys2935](https://doi.org/10.1038/nphys2935).
- [285] K. Luna et al. “Point-contact tunneling spectroscopy measurement of Cu_xTiSe₂: Disorder-enhanced Coulomb effects”. In: *Physical Review B* 91.9 (Mar. 2015). DOI: [10.1103/physrevb.91.094509](https://doi.org/10.1103/physrevb.91.094509).
- [286] T. Das and K. Dolui. “Superconducting dome in MoS₂ and TiSe₂ generated by quasiparticle-phonon coupling”. In: *Physical Review B* 91.9 (Mar. 2015). DOI: [10.1103/physrevb.91.094510](https://doi.org/10.1103/physrevb.91.094510).

- [287] Z. Medvecká et al. “Observation of a transverse Meissner effect in Cu_xTiSe_2 single crystals”. In: *Physical Review B* 93.10 (Mar. 2016). DOI: [10.1103/physrevb.93.100501](https://doi.org/10.1103/physrevb.93.100501).
- [288] M. J. Wei et al. “Manipulating charge density wave order in monolayer 1T – TiSe_2 by strain and charge doping: A first-principles investigation”. In: *Physical Review B* 96.16 (Oct. 2017). DOI: [10.1103/physrevb.96.165404](https://doi.org/10.1103/physrevb.96.165404).
- [289] A. Kogar et al. “Observation of a Charge Density Wave Incommensuration Near the Superconducting Dome in Cu_xTiSe_2 ”. In: *Physical Review Letters* 118.2 (Jan. 2017). DOI: [10.1103/physrevlett.118.027002](https://doi.org/10.1103/physrevlett.118.027002).
- [290] S. Yan et al. “Influence of Domain Walls in the Incommensurate Charge Density Wave State of Cu Intercalated 1T – TiSe_2 ”. In: *Physical Review Letters* 118.10 (Mar. 2017). DOI: [10.1103/physrevlett.118.106405](https://doi.org/10.1103/physrevlett.118.106405).
- [291] A. Banerjee, A. Garg, and A. Ghosal. “Emergent superconductivity upon disordering a charge density wave ground state”. In: *Physical Review B* 98.10 (Sept. 2018). DOI: [10.1103/physrevb.98.104206](https://doi.org/10.1103/physrevb.98.104206).
- [292] B. Hildebrand et al. “Local Real-Space View of the Achiral 1T – TiSe_2 2x2x2 Charge Density Wave”. In: *Physical Review Letters* 120.13 (Mar. 2018). DOI: [10.1103/physrevlett.120.136404](https://doi.org/10.1103/physrevlett.120.136404).
- [293] Q. Yao et al. “Charge Transfer Effects in Naturally Occurring van der Waals Heterostructures $(\text{PbSe})_{1.16}(\text{TiSe}_2)_m$ ($m = 1, 2$)”. In: *Physical Review Letters* 120.10 (Mar. 2018). DOI: [10.1103/physrevlett.120.106401](https://doi.org/10.1103/physrevlett.120.106401).
- [294] K. Rossnagel, L. Kipp, and M. Skibowski. “Charge-density-wave phase transition in 1T – TiSe_2 : Excitonic insulator versus band-type Jahn-Teller mechanism”. In: *Physical Review B* 65.23 (May 2002). DOI: [10.1103/physrevb.65.235101](https://doi.org/10.1103/physrevb.65.235101).
- [295] H. Cercellier et al. “Evidence for an Excitonic Insulator Phase in 1T – TiSe_2 ”. In: *Physical Review Letters* 99.14 (Oct. 2007). DOI: [10.1103/physrevlett.99.146403](https://doi.org/10.1103/physrevlett.99.146403).
- [296] N. Stoffel et al. “Direct evidence for d-band involvement in the TiSe_2 phase transition”. In: *Solid State Communications* 41.1 (Jan. 1982), pp. 53–56. DOI: [10.1016/0038-1098\(82\)90248-4](https://doi.org/10.1016/0038-1098(82)90248-4).
- [297] O. Anderson, R. Manzke, and M. Skibowski. “Three-Dimensional and Relativistic Effects in Layered 1T – TiSe_2 ”. In: *Physical Review Letters* 55.20 (Nov. 1985), pp. 2188–2191. DOI: [10.1103/physrevlett.55.2188](https://doi.org/10.1103/physrevlett.55.2188).

- [298] T. Pillo et al. “Photoemission of bands above the Fermi level: The excitonic insulator phase transition in 1T – TiSe₂”. In: *Physical Review B* 61.23 (June 2000), pp. 16213–16222. DOI: [10.1103/physrevb.61.16213](https://doi.org/10.1103/physrevb.61.16213).
- [299] T. E. Kidd et al. “Electron-Hole Coupling and the Charge Density Wave Transition in TiSe₂”. In: *Physical Review Letters* 88.22 (May 2002). DOI: [10.1103/physrevlett.88.226402](https://doi.org/10.1103/physrevlett.88.226402).
- [300] C. Monney et al. “Temperature-dependent photoemission on 1T – TiSe₂: Interpretation within the exciton condensate phase model”. In: *Physical Review B* 81.15 (Apr. 2010). DOI: [10.1103/physrevb.81.155104](https://doi.org/10.1103/physrevb.81.155104).
- [301] C Monney et al. “Probing the exciton condensate phase in 1T – TiSe₂ with photoemission”. In: *New Journal of Physics* 12.12 (Dec. 2010), p. 125019. DOI: [10.1088/1367-2630/12/12/125019](https://doi.org/10.1088/1367-2630/12/12/125019).
- [302] J. van Wezel, P. Nahai-Williamson, and S. S. Saxena. “An alternative interpretation of recent ARPES measurements on TiSe₂”. In: *EPL (Europhysics Letters)* 89.4 (Feb. 2010), p. 47004. DOI: [10.1209/0295-5075/89/47004](https://doi.org/10.1209/0295-5075/89/47004).
- [303] C. Monney et al. “Exciton Condensation Driving the Periodic Lattice Distortion of 1T – TiSe₂”. In: *Physical Review Letters* 106.10 (Mar. 2011). DOI: [10.1103/physrevlett.106.106404](https://doi.org/10.1103/physrevlett.106.106404).
- [304] M. M. May et al. “Charge-Density-Wave Phase of 1T – TiSe₂: The Influence of Conduction Band Population”. In: *Physical Review Letters* 107.17 (Oct. 2011). DOI: [10.1103/physrevlett.107.176405](https://doi.org/10.1103/physrevlett.107.176405).
- [305] C. Monney et al. “Electron-hole fluctuation phase in 1T – TiSe₂”. In: *Physical Review B* 85.23 (June 2012). DOI: [10.1103/physrevb.85.235150](https://doi.org/10.1103/physrevb.85.235150).
- [306] C Monney et al. “Electron-hole instability in 1T – TiSe₂”. In: *New Journal of Physics* 14.7 (July 2012), p. 075026. DOI: [10.1088/1367-2630/14/7/075026](https://doi.org/10.1088/1367-2630/14/7/075026).
- [307] M. Cazzaniga et al. “*Ab initio* many-body effects in TiSe₂: A possible excitonic insulator scenario from GW band-shape renormalization”. In: *Physical Review B* 85.19 (May 2012). DOI: [10.1103/physrevb.85.195111](https://doi.org/10.1103/physrevb.85.195111).
- [308] B. Zenker et al. “Chiral charge order in 1T – TiSe₂: Importance of lattice degrees of freedom”. In: *Physical Review B* 88.7 (Aug. 2013). DOI: [10.1103/physrevb.88.075138](https://doi.org/10.1103/physrevb.88.075138).

- [309] S. Koley et al. “Preformed excitons, orbital selectivity, and charge density wave order in 1T – TiSe₂”. In: *Physical Review B* 90.11 (Sept. 2014). DOI: [10.1103/physrevb.90.115146](https://doi.org/10.1103/physrevb.90.115146).
- [310] H. Watanabe, K. Seki, and S. Yunoki. “Charge-density wave induced by combined electron-electron and electron-phonon interactions in 1T – TiSe₂: A variational Monte Carlo study”. In: *Physical Review B* 91.20 (May 2015). DOI: [10.1103/physrevb.91.205135](https://doi.org/10.1103/physrevb.91.205135).
- [311] J.-P. Peng et al. “Molecular beam epitaxy growth and scanning tunneling microscopy study of TiSe₂ ultrathin films”. In: *Physical Review B* 91.12 (Mar. 2015). DOI: [10.1103/physrevb.91.121113](https://doi.org/10.1103/physrevb.91.121113).
- [312] G. Monney et al. “Impact of Electron-Hole Correlations on the 1T – TiSe₂ Electronic Structure”. In: *Physical Review Letters* 114.8 (Feb. 2015). DOI: [10.1103/physrevlett.114.086402](https://doi.org/10.1103/physrevlett.114.086402).
- [313] K. Sugawara et al. “Unconventional Charge-Density-Wave Transition in Monolayer 1T – TiSe₂”. In: *ACS Nano* 10.1 (Dec. 2015), pp. 1341–1345. DOI: [10.1021/acsnano.5b06727](https://doi.org/10.1021/acsnano.5b06727).
- [314] B. Hildebrand et al. “Short-range phase coherence and origin of the 1T – TiSe₂ charge density wave”. In: *Physical Review B* 93.12 (Mar. 2016). DOI: [10.1103/physrevb.93.125140](https://doi.org/10.1103/physrevb.93.125140).
- [315] A. Novello et al. “Stripe and Short Range Order in the Charge Density Wave of 1T – Cu_xTiSe₂”. In: *Physical Review Letters* 118.1 (Jan. 2017). DOI: [10.1103/physrevlett.118.017002](https://doi.org/10.1103/physrevlett.118.017002).
- [316] D. Pasquier and O. V. Yazyev. “Excitonic effects in two-dimensional TiSe₂ from hybrid density functional theory”. In: *Physical Review B* 98.23 (Dec. 2018). DOI: [10.1103/physrevb.98.235106](https://doi.org/10.1103/physrevb.98.235106).
- [317] N. V. Baranov et al. “Ni intercalation of titanium diselenide: effect on the lattice, specific heat and magnetic properties”. In: *Journal of Physics: Condensed Matter* 16.50 (Dec. 2004), pp. 9243–9258. DOI: [10.1088/0953-8984/16/50/014](https://doi.org/10.1088/0953-8984/16/50/014).
- [318] C. Chen et al. “Reproduction of the Charge Density Wave Phase Diagram in 1T – TiSe₂ Exposes its Excitonic Character”. In: *Physical Review Letters* 121.22 (Nov. 2018). DOI: [10.1103/physrevlett.121.226602](https://doi.org/10.1103/physrevlett.121.226602).
- [319] T. Rohwer et al. “Collapse of long-range charge order tracked by time-resolved photoemission at high momenta”. In: *Nature* 471.7339 (Mar. 2011), pp. 490–493. DOI: [10.1038/nature09829](https://doi.org/10.1038/nature09829).

- [320] H. P. Hughes. “Structural distortion in TiSe_2 and related materials—a possible Jahn-Teller effect?” In: *Journal of Physics C: Solid State Physics* 10.11 (June 1977), pp. L319–L323. DOI: [10.1088/0022-3719/10/11/009](https://doi.org/10.1088/0022-3719/10/11/009).
- [321] N. Wakabayashi et al. “Phonons and charge density waves in $1\text{T} - \text{TiSe}_2$ ”. In: *Solid State Communications* 28.11 (Dec. 1978), pp. 923–926. DOI: [10.1016/0038-1098\(78\)90112-6](https://doi.org/10.1016/0038-1098(78)90112-6).
- [322] J. H. Gaby et al. “Origin of the structural transition in TiSe_2 ”. In: *Solid State Communications* 39.11 (Sept. 1981), pp. 1167–1170. DOI: [10.1016/0038-1098\(81\)91106-6](https://doi.org/10.1016/0038-1098(81)91106-6).
- [323] K. Motizuki et al. “Role of electron-lattice interaction in lattice dynamics and lattice instability of $1\text{T} - \text{TiSe}_2$ ”. In: *Solid State Communications* 40.11 (Dec. 1981), pp. 995–998. DOI: [10.1016/0038-1098\(81\)90052-1](https://doi.org/10.1016/0038-1098(81)90052-1).
- [324] N. Suzuki, A. Yamamoto, and K. Motizuki. “Electron-lattice interaction and the CDW state of $1\text{T} - \text{TiSe}_2$ ”. In: *Solid State Communications* 49.11 (Mar. 1984), pp. 1039–1043. DOI: [10.1016/0038-1098\(84\)90418-6](https://doi.org/10.1016/0038-1098(84)90418-6).
- [325] J. M. Lopez-Castillo et al. “Phonon-drag effect in $\text{TiSe}_{2-x}\text{S}_x$ mixed compounds”. In: *Physical Review B* 36.8 (Sept. 1987), pp. 4249–4253. DOI: [10.1103/physrevb.36.4249](https://doi.org/10.1103/physrevb.36.4249).
- [326] M. Holt et al. “X-Ray Studies of Phonon Softening in TiSe_2 ”. In: *Physical Review Letters* 86.17 (Apr. 2001), pp. 3799–3802. DOI: [10.1103/physrevlett.86.3799](https://doi.org/10.1103/physrevlett.86.3799).
- [327] A. Bussmann-Holder and A. R. Bishop. “Suppression of charge-density formation in TiSe_2 by Cu doping”. In: *Physical Review B* 79.2 (Jan. 2009). DOI: [10.1103/physrevb.79.024302](https://doi.org/10.1103/physrevb.79.024302).
- [328] J. van Wezel, P. Nahai-Williamson, and S. S. Saxena. “Exciton-phonon-driven charge density wave in TiSe_2 ”. In: *Physical Review B* 81.16 (Apr. 2010). DOI: [10.1103/physrevb.81.165109](https://doi.org/10.1103/physrevb.81.165109).
- [329] K. Rossnagel. “Suppression and emergence of charge-density waves at the surfaces of layered $1\text{T} - \text{TiSe}_2$ and $1\text{T} - \text{TaS}_2$ by *in situ* Rb deposition”. In: *New Journal of Physics* 12.12 (Dec. 2010), p. 125018. DOI: [10.1088/1367-2630/12/12/125018](https://doi.org/10.1088/1367-2630/12/12/125018).
- [330] M. Calandra and F. Mauri. “Charge-Density Wave and Superconducting Dome in TiSe_2 from Electron-Phonon Interaction”. In: *Physical Review Letters* 106.19 (May 2011). DOI: [10.1103/physrevlett.106.196406](https://doi.org/10.1103/physrevlett.106.196406).

- [331] F. Weber et al. “Electron-Phonon Coupling and the Soft Phonon Mode in TiSe_2 ”. In: *Physical Review Letters* 107.26 (Dec. 2011). DOI: [10.1103/physrevlett.107.266401](https://doi.org/10.1103/physrevlett.107.266401).
- [332] Z. Zhu, Y. Cheng, and U. Schwingenschlögl. “Origin of the charge density wave in $1\text{T} - \text{TiSe}_2$ ”. In: *Physical Review B* 85.24 (June 2012). DOI: [10.1103/physrevb.85.245133](https://doi.org/10.1103/physrevb.85.245133).
- [333] V. Olevano et al. “Comment on “Charge-Density Wave and Superconducting Dome in TiSe_2 from Electron-Phonon Interaction””. In: *Physical Review Letters* 112.4 (Jan. 2014). DOI: [10.1103/physrevlett.112.049701](https://doi.org/10.1103/physrevlett.112.049701).
- [334] B.-T. Wang et al. “First-principles study of superconductivity in the two- and three-dimensional forms of PbTiSe_2 : Suppressed charge density wave in $1\text{T} - \text{TiSe}_2$ ”. In: *Physical Review B* 98.1 (July 2018). DOI: [10.1103/physrevb.98.014514](https://doi.org/10.1103/physrevb.98.014514).
- [335] T. Kaneko, Y. Ohta, and S. Yunoki. “Exciton-phonon cooperative mechanism of the triple- \mathbf{q} charge-density-wave and antiferroelectric electron polarization in TiSe_2 ”. In: *Physical Review B* 97.15 (Apr. 2018). DOI: [10.1103/physrevb.97.155131](https://doi.org/10.1103/physrevb.97.155131).
- [336] J. J. Mortensen, L. B. Hansen, and K. W. Jacobsen. “Real-space grid implementation of the projector augmented wave method”. In: *Physical Review B* 71.3 (Jan. 2005). DOI: [10.1103/physrevb.71.035109](https://doi.org/10.1103/physrevb.71.035109).
- [337] J. Enkovaara et al. “Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method”. In: *Journal of Physics: Condensed Matter* 22.25 (June 2010), p. 253202. DOI: [10.1088/0953-8984/22/25/253202](https://doi.org/10.1088/0953-8984/22/25/253202).
- [338] A. H. Larsen et al. “The atomic simulation environment—a Python library for working with atoms”. In: *Journal of Physics: Condensed Matter* 29.27 (June 2017), p. 273002. DOI: [10.1088/1361-648x/aa680e](https://doi.org/10.1088/1361-648x/aa680e).
- [339] P. E. Blöchl. “Projector augmented-wave method”. In: *Physical Review B* 50.24 (Dec. 1994), pp. 17953–17979. DOI: [10.1103/physrevb.50.17953](https://doi.org/10.1103/physrevb.50.17953).
- [340] J. P. Perdew, K. Burke, and M. Ernzerhof. “Generalized Gradient Approximation Made Simple”. In: *Physical Review Letters* 77.18 (Oct. 1996), pp. 3865–3868. DOI: [10.1103/physrevlett.77.3865](https://doi.org/10.1103/physrevlett.77.3865).
- [341] H. J. Monkhorst and J. D. Pack. “Special points for Brillouin-zone integrations”. In: *Physical Review B* 13.12 (June 1976), pp. 5188–5192. DOI: [10.1103/physrevb.13.5188](https://doi.org/10.1103/physrevb.13.5188).

- [342] S. Sharma et al. “Bootstrap Approximation for the Exchange-Correlation Kernel of Time-Dependent Density-Functional Theory”. In: *Physical Review Letters* 107.18 (Oct. 2011). DOI: [10.1103/physrevlett.107.186401](https://doi.org/10.1103/physrevlett.107.186401).
- [343] Z. Yuan and S. Gao. “Linear response approach to collective electronic excitations of solids and surfaces”. In: *Computer Physics Communications* 180.3 (Mar. 2009), pp. 466–473. DOI: [10.1016/j.cpc.2008.10.019](https://doi.org/10.1016/j.cpc.2008.10.019).
- [344] J. Yan, K. W. Jacobsen, and K. S. Thygesen. “First-principles study of surface plasmons on Ag(111) and H/Ag(111)”. In: *Physical Review B* 84.23 (Dec. 2011). DOI: [10.1103/physrevb.84.235430](https://doi.org/10.1103/physrevb.84.235430).
- [345] C. Lian et al. “Integrated Plasmonics: Broadband Dirac Plasmons in Borophene”. In: *Physical Review Letters* 125.11 (Sept. 2020). DOI: [10.1103/physrevlett.125.116802](https://doi.org/10.1103/physrevlett.125.116802).
- [346] B. M. Wong and T. H. Hsieh. “Optoelectronic and Excitonic Properties of Oligoacenes: Substantial Improvements from Range-Separated Time-Dependent Density Functional Theory”. In: *Journal of Chemical Theory and Computation* 6.12 (Nov. 2010), pp. 3704–3712. DOI: [10.1021/ct100529s](https://doi.org/10.1021/ct100529s).
- [347] M. E. Foster and B. M. Wong. “Nonempirically Tuned Range-Separated DFT Accurately Predicts Both Fundamental and Excitation Gaps in DNA and RNA Nucleobases”. In: *Journal of Chemical Theory and Computation* 8.8 (July 2012), pp. 2682–2687. DOI: [10.1021/ct300420f](https://doi.org/10.1021/ct300420f).
- [348] B. M. Wong, M. Piacenza, and F. D. Sala. “Absorption and fluorescence properties of oligothiophene biomarkers from long-range-corrected time-dependent density functional theory”. In: *Physical Chemistry Chemical Physics* 11.22 (2009), p. 4498. DOI: [10.1039/b901743g](https://doi.org/10.1039/b901743g).
- [349] B. M. Wong and J. G. Cordaro. “Coumarin dyes for dye-sensitized solar cells: A long-range-corrected density functional study”. In: *The Journal of Chemical Physics* 129.21 (Dec. 2008), p. 214703. DOI: [10.1063/1.3025924](https://doi.org/10.1063/1.3025924).
- [350] B. M. Wong. “Optoelectronic Properties of Carbon Nanorings: Excitonic Effects from Time-Dependent Density Functional Theory”. In: *The Journal of Physical Chemistry C* 113.52 (Dec. 2009), pp. 21921–21927. DOI: [10.1021/jp9074674](https://doi.org/10.1021/jp9074674).
- [351] A. E. Raeber and B. M. Wong. “The Importance of Short- and Long-Range Exchange on Various Excited State Properties of DNA Monomers, Stacked Complexes, and Watson–Crick Pairs”. In: *Journal of Chemical Theory and Computation* 11.5 (Apr. 2015), pp. 2199–2209. DOI: [10.1021/acs.jctc.5b00105](https://doi.org/10.1021/acs.jctc.5b00105).

- [352] W. Ku, T. Berlijn, and C.-C. Lee. “Unfolding First-Principles Band Structures”. In: *Physical Review Letters* 104.21 (May 2010). DOI: [10.1103/physrevlett.104.216401](https://doi.org/10.1103/physrevlett.104.216401).
- [353] V. Popescu and A. Zunger. “Extracting \mathbf{E} versus \vec{k} effective band structure from supercell calculations on alloys and impurities”. In: *Physical Review B* 85.8 (Feb. 2012). DOI: [10.1103/physrevb.85.085201](https://doi.org/10.1103/physrevb.85.085201).
- [354] C. Lian and S. Meng. “Dirac cone pairs in silicene induced by interface Si-Ag hybridization: A first-principles effective band study”. In: *Physical Review B* 95.24 (June 2017). DOI: [10.1103/physrevb.95.245409](https://doi.org/10.1103/physrevb.95.245409).
- [355] S. Baroni et al. “Phonons and related crystal properties from density-functional perturbation theory”. In: *Reviews of Modern Physics* 73.2 (July 2001), pp. 515–562. DOI: [10.1103/revmodphys.73.515](https://doi.org/10.1103/revmodphys.73.515).
- [356] F. Giustino. “Electron-phonon interactions from first principles”. In: *Reviews of Modern Physics* 89.1 (Feb. 2017). DOI: [10.1103/revmodphys.89.015003](https://doi.org/10.1103/revmodphys.89.015003).
- [357] S. Ponc e et al. “Verification of first-principles codes: Comparison of total energies, phonon frequencies, electron–phonon coupling and zero-point motion correction to the gap between ABINIT and QE/Yambo”. In: *Computational Materials Science* 83 (Feb. 2014), pp. 341–348. DOI: [10.1016/j.commatsci.2013.11.031](https://doi.org/10.1016/j.commatsci.2013.11.031).
- [358] S. Ponc e et al. “Temperature dependence of the electronic structure of semiconductors and insulators”. In: *The Journal of Chemical Physics* 143.10 (Sept. 2015), p. 102813. DOI: [10.1063/1.4927081](https://doi.org/10.1063/1.4927081).
- [359] S. Ponc e et al. “EPW: Electron–phonon coupling, transport and superconducting properties using maximally localized Wannier functions”. In: *Computer Physics Communications* 209 (Dec. 2016), pp. 116–133. DOI: [10.1016/j.cpc.2016.07.028](https://doi.org/10.1016/j.cpc.2016.07.028).
- [360] G. Vignale and W. Kohn. “Current-Dependent Exchange-Correlation Potential for Dynamical Linear Response Theory”. In: *Physical Review Letters* 77.10 (Sept. 1996), pp. 2037–2040. DOI: [10.1103/physrevlett.77.2037](https://doi.org/10.1103/physrevlett.77.2037).
- [361] H. O. Wijewardane and C. A. Ullrich. “Time-Dependent Kohn-Sham Theory with Memory”. In: *Physical Review Letters* 95.8 (Aug. 2005). DOI: [10.1103/physrevlett.95.086401](https://doi.org/10.1103/physrevlett.95.086401).
- [362] C. A. Ullrich and G. Vignale. “Time-dependent current-density-functional theory for the linear response of weakly disordered systems”. In: *Physical Review B* 65.24 (May 2002). DOI: [10.1103/physrevb.65.245102](https://doi.org/10.1103/physrevb.65.245102).

- [363] C. Ullrich and G. Vignale. “Erratum: Time-dependent current-density-functional theory for the linear response of weakly disordered systems Phys. Rev. B **65**, 245102 (2002)”. In: *Physical Review B* 70.23 (Dec. 2004). DOI: [10.1103/physrevb.70.239903](https://doi.org/10.1103/physrevb.70.239903).
- [364] C. A. Ullrich. “Time-dependent density-functional theory beyond the adiabatic approximation: Insights from a two-electron model system”. In: *The Journal of Chemical Physics* 125.23 (Dec. 2006), p. 234108. DOI: [10.1063/1.2406069](https://doi.org/10.1063/1.2406069).
- [365] L. Lacombe and N. T. Maitra. “Density-Matrix Coupled Time-Dependent Exchange-Correlation Functional Approximations”. In: *Journal of Chemical Theory and Computation* 15.3 (Jan. 2019), pp. 1672–1678. DOI: [10.1021/acs.jctc.8b01159](https://doi.org/10.1021/acs.jctc.8b01159).
- [366] C. A. Ullrich and K. Burke. “Excitation energies from time-dependent density-functional theory beyond the adiabatic approximation”. In: *The Journal of Chemical Physics* 121.1 (2004), p. 28. DOI: [10.1063/1.1756865](https://doi.org/10.1063/1.1756865).
- [367] G. Vignale, C. A. Ullrich, and S. Conti. “Time-Dependent Density Functional Theory Beyond the Adiabatic Local Density Approximation”. In: *Physical Review Letters* 79.24 (Dec. 1997), pp. 4878–4881. DOI: [10.1103/physrevlett.79.4878](https://doi.org/10.1103/physrevlett.79.4878).
- [368] C. A. Ullrich and G. Vignale. “Theory of the Linewidth of Intersubband Plasmons in Quantum Wells”. In: *Physical Review Letters* 87.3 (July 2001). DOI: [10.1103/physrevlett.87.037402](https://doi.org/10.1103/physrevlett.87.037402).
- [369] C. A. Ullrich and G. Vignale. “Collective intersubband transitions in quantum wells: A comparative density-functional study”. In: *Physical Review B* 58.23 (Dec. 1998), pp. 15756–15765. DOI: [10.1103/physrevb.58.15756](https://doi.org/10.1103/physrevb.58.15756).
- [370] J. A. Berger. “Fully Parameter-Free Calculation of Optical Spectra for Insulators, Semiconductors, and Metals from a Simple Polarization Functional”. In: *Physical Review Letters* 115.13 (Sept. 2015). DOI: [10.1103/physrevlett.115.137402](https://doi.org/10.1103/physrevlett.115.137402).
- [371] K. Hanasaki et al. “Implementation of Real-Time TDDFT for Periodic Systems in the Open-Source PySCF Software Package”. In: *N/A* (). (Writing Manuscript).
- [372] E. Runge and E. K. U. Gross. “Density-Functional Theory for Time-Dependent Systems”. In: *Physical Review Letters* 52.12 (Mar. 1984), pp. 997–1000. DOI: [10.1103/physrevlett.52.997](https://doi.org/10.1103/physrevlett.52.997).
- [373] R. D. King-Smith and D. Vanderbilt. “Theory of polarization of crystalline solids”. In: *Physical Review B* 47.3 (Jan. 1993), pp. 1651–1654. DOI: [10.1103/physrevb.47.1651](https://doi.org/10.1103/physrevb.47.1651).

- [374] C. Lian et al. “Ultrafast charge ordering by self-amplified exciton–phonon dynamics in TiSe₂”. In: *Nature Communications* 11.1 (Jan. 2020). (b) *ibid.*, “Illustrating the nature of ultrafast charge density wave dynamics in TiSe₂: Interplay of exciton electron-phonon coupling”, arXiv.1901.00610v2 cond-mat.mes-hall, -. DOI: [10.1038/s41467-019-13672-7](https://doi.org/10.1038/s41467-019-13672-7).
- [375] J. Sun et al. “Real-Time Exciton Dynamics with Time-Dependent Density-Functional Theory”. In: *Physical Review Letters* 127.7 (Aug. 2021). DOI: [10.1103/physrevlett.127.077401](https://doi.org/10.1103/physrevlett.127.077401).
- [376] K. Yabana et al. “Real-time, real-space implementation of the linear response time-dependent density-functional theory”. In: *physica status solidi (b)* 243.5 (Apr. 2006), pp. 1121–1138. DOI: [10.1002/pssb.200642005](https://doi.org/10.1002/pssb.200642005).
- [377] M. Noda et al. “SALMON: Scalable Ab-initio Light–Matter simulator for Optics and Nanoscience”. In: *Computer Physics Communications* 235 (Feb. 2019), pp. 356–365. DOI: [10.1016/j.cpc.2018.09.018](https://doi.org/10.1016/j.cpc.2018.09.018).
- [378] J. K. Dewhurst et al. *An efficient algorithm for time propagation within time-dependent density functional theory*. 2014. DOI: [10.48550/ARXIV.1412.0996](https://doi.org/10.48550/ARXIV.1412.0996).
- [379] C. Pemmaraju et al. “Velocity-gauge real-time TDDFT within a numerical atomic orbital basis set”. In: *Computer Physics Communications* 226 (May 2018), pp. 30–38. DOI: [10.1016/j.cpc.2018.01.013](https://doi.org/10.1016/j.cpc.2018.01.013).
- [380] J. VandeVondele and J. Hutter. “Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases”. In: *The Journal of Chemical Physics* 127.11 (Sept. 2007), p. 114105. DOI: [10.1063/1.2770708](https://doi.org/10.1063/1.2770708).
- [381] Q. Sun et al. “PySCF: the Python-based simulations of chemistry framework”. In: *WIREs Computational Molecular Science* 8.1 (2018), e1340. DOI: <https://doi.org/10.1002/wcms.1340>.
- [382] J. L. Whitten. “Coulombic potential energy integrals and approximations”. In: *The Journal of Chemical Physics* 58.10 (May 1973), pp. 4496–4501. DOI: [10.1063/1.1679012](https://doi.org/10.1063/1.1679012).
- [383] F. Weigend. “Accurate Coulomb-fitting basis sets for H to Rn”. In: *Physical Chemistry Chemical Physics* 8.9 (2006), p. 1057. DOI: [10.1039/b515623h](https://doi.org/10.1039/b515623h).
- [384] Q. Sun et al. “Gaussian and plane-wave mixed density fitting for periodic systems”. In: *The Journal of Chemical Physics* 147.16 (Oct. 2017), p. 164119. DOI: [10.1063/1.4998644](https://doi.org/10.1063/1.4998644).

- [385] G. Lippert, J. Hutter, and M. Parrinello. “A hybrid Gaussian and plane wave density functional scheme”. In: *Molecular Physics* 92.3 (Oct. 1997), pp. 477–487. DOI: [10.1080/00268979709482119](https://doi.org/10.1080/00268979709482119).
- [386] Z. Merali. “Computational science: ...Error”. In: *Nature* 467.7317 (Oct. 2010), pp. 775–777. DOI: [10.1038/467775a](https://doi.org/10.1038/467775a).
- [387] M. Thiele, E. K. U. Gross, and S. Kümmel. “Adiabatic Approximation in Nonperturbative Time-Dependent Density-Functional Theory”. In: *Physical Review Letters* 100.15 (Apr. 2008). DOI: [10.1103/physrevlett.100.153004](https://doi.org/10.1103/physrevlett.100.153004).
- [388] W. Kohn and L. J. Sham. “Self-Consistent Equations Including Exchange and Correlation Effects”. In: *Physical Review* 140.4A (Nov. 1965), A1133–A1138. DOI: [10.1103/physrev.140.a1133](https://doi.org/10.1103/physrev.140.a1133).
- [389] J. P. Perdew and Y. Wang. “Accurate and simple analytic representation of the electron-gas correlation energy”. In: *Physical Review B* 45.23 (June 1992), pp. 13244–13249. DOI: [10.1103/physrevb.45.13244](https://doi.org/10.1103/physrevb.45.13244).
- [390] J. P. Perdew, K. Burke, and M. Ernzerhof. “Generalized Gradient Approximation Made Simple”. In: *Physical Review Letters* 77.18 (Oct. 1996), pp. 3865–3868. DOI: [10.1103/physrevlett.77.3865](https://doi.org/10.1103/physrevlett.77.3865).
- [391] J. Crank and P. Nicolson. “A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type”. In: *Mathematical Proceedings of the Cambridge Philosophical Society* 43.1 (Jan. 1947), pp. 50–67. DOI: [10.1017/s0305004100023197](https://doi.org/10.1017/s0305004100023197).
- [392] P. Pulay. “Convergence acceleration of iterative sequences. the case of scf iteration”. In: *Chemical Physics Letters* 73.2 (July 1980), pp. 393–398. DOI: [10.1016/0009-2614\(80\)80396-4](https://doi.org/10.1016/0009-2614(80)80396-4).
- [393] S. Goedecker, M. Teter, and J. Hutter. “Separable dual-space Gaussian pseudopotentials”. In: *Physical Review B* 54.3 (July 1996), pp. 1703–1710. DOI: [10.1103/physrevb.54.1703](https://doi.org/10.1103/physrevb.54.1703).
- [394] U. N. Morzan et al. “Electron dynamics in complex environments with real-time time dependent density functional theory in a QM-MM framework”. In: *The Journal of Chemical Physics* 140.16 (Apr. 2014), p. 164105. DOI: [10.1063/1.4871688](https://doi.org/10.1063/1.4871688).
- [395] K. Yabana et al. “Time-dependent density functional theory for strong electromagnetic fields in crystalline solids”. In: *Physical Review B* 85.4 (Jan. 2012). DOI: [10.1103/physrevb.85.045134](https://doi.org/10.1103/physrevb.85.045134).

- [396] S. H. Vosko, L. Wilk, and M. Nusair. “Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis”. In: *Canadian Journal of Physics* 58.8 (Aug. 1980), pp. 1200–1211. DOI: [10.1139/p80-159](https://doi.org/10.1139/p80-159).
- [397] T. H. Dunning. “Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen”. In: *The Journal of Chemical Physics* 90.2 (Jan. 1989), pp. 1007–1023. DOI: [10.1063/1.456153](https://doi.org/10.1063/1.456153).
- [398] W. H. Press et al. “Numerical recipes in C”. In: *Press Syndicate of the University of Cambridge, New York* 24.78 (1999), pp. 398–447.
- [399] From Eqs. (8.18) and (8.21), it is clear that the dipole spectrum vanishes at $\omega = 0$ while that of the dipole velocity does not necessarily. The latter can take finite value at $\omega = 0$ because of numerical errors arising from the finite-range Fourier transformation.
- [400] L. Liu, Y. P. Feng, and Z. X. Shen. “Structural and electronic properties of *h*-BN”. In: *Physical Review B* 68.10 (Sept. 2003). DOI: [10.1103/physrevb.68.104102](https://doi.org/10.1103/physrevb.68.104102).
- [401] H. J. Monkhorst and J. D. Pack. “Special points for Brillouin-zone integrations”. In: *Physical Review B* 13.12 (June 1976), pp. 5188–5192. DOI: [10.1103/physrevb.13.5188](https://doi.org/10.1103/physrevb.13.5188).
- [402] M. D. Hanwell et al. “Avogadro: an advanced semantic chemical editor, visualization, and analysis platform”. In: *Journal of Cheminformatics* 4.1 (Aug. 2012). DOI: [10.1186/1758-2946-4-17](https://doi.org/10.1186/1758-2946-4-17).
- [403] B. Aradi, B. Hourahine, and T. Frauenheim. “DFTB, a Sparse Matrix-Based Implementation of the DFTB Method”. In: *The Journal of Physical Chemistry A* 111.26 (June 2007), pp. 5678–5684. DOI: [10.1021/jp070186p](https://doi.org/10.1021/jp070186p).
- [404] M. Gaus, A. Goez, and M. Elstner. “Parametrization and Benchmark of DFTB3 for Organic Molecules”. In: *Journal of Chemical Theory and Computation* 9.1 (Nov. 2012), pp. 338–354. DOI: [10.1021/ct300849w](https://doi.org/10.1021/ct300849w).
- [405] X. Lu et al. “Parametrization of DFTB3/3OB for Magnesium and Zinc for Chemical and Biological Applications”. In: *The Journal of Physical Chemistry B* 119.3 (Sept. 2014), pp. 1062–1082. DOI: [10.1021/jp506557r](https://doi.org/10.1021/jp506557r).
- [406] C. Lian et al. “Indirect but Efficient: Laser-Excited Electrons Can Drive Ultrafast Polarization Switching in Ferroelectric Materials”. In: *The Journal of Physical Chemistry Letters* 10.12 (June 2019), pp. 3402–3407. DOI: [10.1021/acs.jpcllett.9b01046](https://doi.org/10.1021/acs.jpcllett.9b01046).

- [407] M. Dawber, K. M. Rabe, and J. F. Scott. “Physics of thin-film ferroelectric oxides”. In: *Reviews of Modern Physics* 77.4 (Oct. 2005), p. 1083. DOI: [10.1103/RevModPhys.77.1083](https://doi.org/10.1103/RevModPhys.77.1083).
- [408] C. A. P. D. Araujo et al. “Ferroelectric memories”. In: *Ferroelectrics* 104.1 (Apr. 1990), p. 241. DOI: [10.1080/00150199008223827](https://doi.org/10.1080/00150199008223827).
- [409] A. Chanthbouala et al. “A ferroelectric memristor”. In: *Nature Materials* 11.10 (Sept. 2012), p. 860. DOI: [10.1038/nmat3415](https://doi.org/10.1038/nmat3415).
- [410] S. H. Jo et al. “Nanoscale Memristor Device as Synapse in Neuromorphic Systems”. In: *Nano Letters* 10.4 (Apr. 2010), p. 1297. DOI: [10.1021/nl904092h](https://doi.org/10.1021/nl904092h).
- [411] M.-K. Kim and J.-S. Lee. “Ferroelectric Analog Synaptic Transistors”. In: *Nano Letters* (Feb. 2019), acs.nanolett.9b00180. DOI: [10.1021/acs.nanolett.9b00180](https://doi.org/10.1021/acs.nanolett.9b00180).
- [412] X. Zhou et al. “Recent theoretical progress in the development of perovskite photo-voltaic materials”. In: *Journal of Energy Chemistry* 27.3 (May 2018), p. 637. DOI: [10.1016/j.jechem.2017.10.010](https://doi.org/10.1016/j.jechem.2017.10.010).
- [413] W. Li et al. “Control of Charge Recombination in Perovskites by Oxidation State of Halide Vacancy”. In: *J. Am. Chem. Soc.* 140 (2018), p. 15753. DOI: [10.1021/jacs.8b08448](https://doi.org/10.1021/jacs.8b08448).
- [414] S. Ghosh, D. D. Sante, and A. Stroppa. “Strain Tuning of Ferroelectric Polarization in Hybrid Organic Inorganic Perovskite Compounds”. In: *The Journal of Physical Chemistry Letters* 6.22 (Nov. 2015), pp. 4553–4559. DOI: [10.1021/acs.jpcllett.5b01806](https://doi.org/10.1021/acs.jpcllett.5b01806).
- [415] A. Stroppa et al. “Ferroelectric Polarization of CH₃NH₃PbI₃: A Detailed Study Based on Density Functional Theory and Symmetry Mode Analysis”. In: *The Journal of Physical Chemistry Letters* 6.12 (June 2015), p. 2223. DOI: [10.1021/acs.jpcllett.5b00542](https://doi.org/10.1021/acs.jpcllett.5b00542).
- [416] H.-W. Chen et al. “Emergence of Hysteresis and Transient Ferroelectric Response in Organo-Lead Halide Perovskite Solar Cells”. In: *The Journal of Physical Chemistry Letters* 6.1 (Dec. 2014), p. 164. DOI: [10.1021/jz502429u](https://doi.org/10.1021/jz502429u).
- [417] J. Wei et al. “Hysteresis Analysis Based on the Ferroelectric Effect in Hybrid Perovskite Solar Cells”. In: *The Journal of Physical Chemistry Letters* 5.21 (Oct. 2014), p. 3937. DOI: [10.1021/jz502111u](https://doi.org/10.1021/jz502111u).

- [418] H.-S. Kim et al. “Ferroelectric Polarization in CH₃NH₃PbI₃ Perovskite”. In: *The Journal of Physical Chemistry Letters* 6.9 (Apr. 2015), p. 1729. DOI: [10.1021/acs.jpcllett.5b00695](https://doi.org/10.1021/acs.jpcllett.5b00695).
- [419] C.-J. Tong et al. “Role of Methylammonium Orientation in Ion Diffusion and Current–Voltage Hysteresis in the CH₃NH₃PbI₃ Perovskite”. In: *ACS Energy Letters* 2.9 (Aug. 2017), p. 1997. DOI: [10.1021/acsenergylett.7b00659](https://doi.org/10.1021/acsenergylett.7b00659).
- [420] C. Ji et al. “Bandgap Narrowing of Lead-Free Perovskite-Type Hybrids for Visible-Light-Absorbing Ferroelectric Semiconductors”. In: *The Journal of Physical Chemistry Letters* 8.9 (Apr. 2017), p. 2012. DOI: [10.1021/acs.jpcllett.7b00673](https://doi.org/10.1021/acs.jpcllett.7b00673).
- [421] J. Jankowska, R. Long, and O. V. Prezhdo. “Quantum Dynamics of Photogenerated Charge Carriers in Hybrid Perovskites: Dopants, Grain Boundaries, Electric Order, and Other Realistic Aspects”. In: *ACS Energy Letters* 2.7 (June 2017), p. 1588. DOI: [10.1021/acsenergylett.7b00198](https://doi.org/10.1021/acsenergylett.7b00198).
- [422] J. Jankowska and O. V. Prezhdo. “Ferroelectric Alignment of Organic Cations Inhibits Nonradiative Electron–Hole Recombination in Hybrid Perovskites: Ab Initio Nonadiabatic Molecular Dynamics”. In: *The Journal of Physical Chemistry Letters* 8.4 (Feb. 2017), p. 812. DOI: [10.1021/acs.jpcllett.7b00008](https://doi.org/10.1021/acs.jpcllett.7b00008).
- [423] W. Li et al. “Hole Trapping by Iodine Interstitial Defects Decreases Free Carrier Losses in Perovskite Solar Cells: A Time-Domain Ab Initio Study”. In: *ACS Energy Letters* 2.6 (May 2017), p. 1270. DOI: [10.1021/acsenergylett.7b00183](https://doi.org/10.1021/acsenergylett.7b00183).
- [424] A. Kakekhani and S. Ismail-Beigi. “Ferroelectric-Based Catalysis: Switchable Surface Chemistry”. In: *ACS Catalysis* 5.8 (June 2015), p. 4537. DOI: [10.1021/acscatal.5b00507](https://doi.org/10.1021/acscatal.5b00507).
- [425] K. M. Bal and E. C. Neyts. “Overcoming Old Scaling Relations and Establishing New Correlations in Catalytic Surface Chemistry: Combined Effect of Charging and Doping”. In: *The Journal of Physical Chemistry C* 123.10 (Feb. 2019), p. 6141. DOI: [10.1021/acs.jpcc.9b01216](https://doi.org/10.1021/acs.jpcc.9b01216).
- [426] W. J. Merz. “Domain Formation and Domain Wall Motions in Ferroelectric BaTiO₃ Single Crystals”. In: *Physical Review* 95.3 (Aug. 1954), p. 690. DOI: [10.1103/PhysRev.95.690](https://doi.org/10.1103/PhysRev.95.690).
- [427] P. K. Larsen et al. “Nanosecond Switching of Thin Ferroelectric Films”. In: *Appl. Phys. Lett.* 59 (1991), p. 611. DOI: [10.1063/1.105402](https://doi.org/10.1063/1.105402).

- [428] S. Fahy and R. Merlin. “Reversal of Ferroelectric Domains by Ultrashort Optical Pulses”. In: *Phys. Rev. Lett.* 73 (1994), p. 1122. DOI: [10.1103/PhysRevLett.73.1122](https://doi.org/10.1103/PhysRevLett.73.1122).
- [429] J. Li et al. “Ultrafast Polarization Switching in Thin-Film Ferroelectrics”. In: *Appl. Phys. Lett.* 84 (2004), p. 1174. DOI: [10.1063/1.1644917](https://doi.org/10.1063/1.1644917).
- [430] K. Takahashi, N. Kida, and M. Tonouchi. “Terahertz Radiation by an Ultrafast Spontaneous Polarization Modulation of Multiferroic BiFeO₃ Thin Films”. In: *Physical Review Letters* 96.11 (Mar. 2006), p. 117402. DOI: [10.1103/PhysRevLett.96.117402](https://doi.org/10.1103/PhysRevLett.96.117402).
- [431] A. Cavalleri et al. “Tracking the motion of charges in a terahertz light field by femtosecond X-ray diffraction”. In: *Nature* 442.7103 (Aug. 2006), p. 664. DOI: [10.1038/nature05041](https://doi.org/10.1038/nature05041).
- [432] T. Qi et al. “Collective Coherent Control: Synchronization of Polarization in Ferroelectric PbTiO₃ by Shaped THz Fields”. In: *Physical Review Letters* 102.24 (June 2009), p. 247603. DOI: [10.1103/PhysRevLett.102.247603](https://doi.org/10.1103/PhysRevLett.102.247603).
- [433] C. Hauf et al. “Soft-mode driven polarity reversal in ferroelectrics mapped by ultrafast x-ray diffraction”. In: *Structural Dynamics* 5.2 (Mar. 2018), p. 024501. DOI: [10.1063/1.5026494](https://doi.org/10.1063/1.5026494).
- [434] A. von Hoegen et al. “Probing the interatomic potential of solids with strong-field nonlinear phononics”. In: *Nature* 555.7694 (Feb. 2018), p. 79. DOI: [10.1038/nature25484](https://doi.org/10.1038/nature25484).
- [435] F. Chen et al. “Ultrafast terahertz-field-driven ionic response in ferroelectric BaTiO₃”. In: *Physical Review B* 94.18 (Nov. 2016), p. 180104. DOI: [10.1103/PhysRevB.94.180104](https://doi.org/10.1103/PhysRevB.94.180104).
- [436] M. T. Islam et al. “High resolution melting curve analysis enables rapid and reliable detection of G6PD variants in heterozygous females”. In: *BMC Genetics* 19.1 (Aug. 2018), p. 58. DOI: [10.1186/s12863-018-0664-1](https://doi.org/10.1186/s12863-018-0664-1).
- [437] A. Subedi. “Proposal for ultrafast switching of ferroelectrics using midinfrared pulses”. In: *Physical Review B* 92.21 (Dec. 2015), p. 214303. DOI: [10.1103/PhysRevB.92.214303](https://doi.org/10.1103/PhysRevB.92.214303).
- [438] R. Mankowsky et al. “Ultrafast Reversal of the Ferroelectric Polarization”. In: *Physical Review Letters* 118.19 (May 2017), p. 197601. DOI: [10.1103/PhysRevLett.118.197601](https://doi.org/10.1103/PhysRevLett.118.197601).

- [439] D. Daranciang et al. “Ultrafast Photovoltaic Response in Ferroelectric Nanolayers”. In: *Phys. Rev. Lett.* 108 (2012), p. 087601. DOI: [10.1103/PhysRevLett.108.087601](https://doi.org/10.1103/PhysRevLett.108.087601).
- [440] M. Lejman et al. “Giant ultrafast photo-induced shear strain in ferroelectric BiFeO₃”. In: *Nature Communications* 5.1 (July 2014), p. 4301. DOI: [10.1038/ncomms5301](https://doi.org/10.1038/ncomms5301).
- [441] Y. Ahn et al. “Photoinduced Domain Pattern Transformation in Ferroelectric-Dielectric Superlattices”. In: *Physical Review Letters* 119.5 (July 2017), p. 057601. DOI: [10.1103/PhysRevLett.119.057601](https://doi.org/10.1103/PhysRevLett.119.057601).
- [442] H. Akamatsu et al. “Light-Activated Gigahertz Ferroelectric Domain Dynamics”. In: *Physical Review Letters* 120.9 (Feb. 2018), p. 096101. DOI: [10.1103/PhysRevLett.120.096101](https://doi.org/10.1103/PhysRevLett.120.096101).
- [443] R. Shinde and A. K. Singh. “Nonlinear Polarization and Low-Dissipation Ultrafast Optical Switching in Phosphorene”. In: *J. Phys. Chem. C* 122 (2018), p. 19146. DOI: [10.1021/acs.jpcc.8b04134](https://doi.org/10.1021/acs.jpcc.8b04134).
- [444] C. v. Korff Schmising et al. “Coupled Ultrafast Lattice and Polarization Dynamics in Ferroelectric Nanolayers”. In: *Physical Review Letters* 98.25 (June 2007), p. 257601. DOI: [10.1103/PhysRevLett.98.257601](https://doi.org/10.1103/PhysRevLett.98.257601).
- [445] D. S. Rana et al. “Understanding the Nature of Ultrafast Polarization Dynamics of Ferroelectric Memory in the Multiferroic BiFeO₃”. In: *Advanced Materials* 21.28 (July 2009), p. 2881. DOI: [10.1002/adma.200802094](https://doi.org/10.1002/adma.200802094).
- [446] Y.-H. Kuo et al. “Ultrafast light-induced symmetry changes in single BaTiO₃ nanowires”. In: *Journal of Materials Chemistry C* 5.6 (2017), p. 1522. DOI: [10.1039/c6tc04448d](https://doi.org/10.1039/c6tc04448d).
- [447] A. Rubano et al. “Ultrafast modification of the polarity at LaAlO₃/SrTiO₃ interfaces”. In: *Physical Review B* 97.3 (Jan. 2018), p. 035438. DOI: [10.1103/PhysRevB.97.035438](https://doi.org/10.1103/PhysRevB.97.035438).
- [448] M. Först et al. “Displacive lattice excitation through nonlinear phononics viewed by femtosecond X-ray diffraction”. In: *Solid State Communications* 169 (Sept. 2013), p. 24. DOI: [10.1016/j.ssc.2013.06.024](https://doi.org/10.1016/j.ssc.2013.06.024).
- [449] M. Porer et al. “Ultrafast Relaxation Dynamics of the Antiferrodistortive Phase in Ca Doped SrTiO₃”. In: *Physical Review Letters* 121.5 (Sept. 2018), p. 055701. DOI: [10.1103/PhysRevLett.121.055701](https://doi.org/10.1103/PhysRevLett.121.055701).
- [450] Y. Qi et al. “Ultrafast Electric Field Pulse Control of Giant Temperature Change in Ferroelectrics”. In: *Physical Review Letters* 120.5 (Jan. 2018), p. 055901. DOI: [10.1103/PhysRevLett.120.055901](https://doi.org/10.1103/PhysRevLett.120.055901).

- [451] S. Meng and E. Kaxiras. “Real-time, local basis-set implementation of time-dependent density functional theory for excited state dynamics simulations”. In: *The Journal of Chemical Physics* 129.5 (Aug. 2008), p. 054110. DOI: [10.1063/1.2960628](https://doi.org/10.1063/1.2960628).
- [452] C. Lian et al. “Momentum-resolved TDDFT algorithm in atomic basis for real time tracking of electronic excitation”. In: *The Journal of Chemical Physics* 149.15 (Oct. 2018), p. 154104. DOI: [10.1063/1.5036543](https://doi.org/10.1063/1.5036543).
- [453] C. Lian et al. “Photoexcitation in Solids: First-Principles Quantum Simulations by Real-Time TDDFT”. In: *Advanced Theory and Simulations* 1.8 (July 2018), p. 1800055. DOI: [10.1002/adts.201800055](https://doi.org/10.1002/adts.201800055).
- [454] E. Runge and E. K. U. Gross. “Density-Functional Theory for Time-Dependent Systems”. In: *Physical Review Letters* 52.12 (Mar. 1984), p. 997. DOI: [10.1103/PhysRevLett.52.997](https://doi.org/10.1103/PhysRevLett.52.997).
- [455] G. F. Bertsch et al. “Real-space, real-time method for the dielectric function”. In: *Physical Review B* 62.12 (Sept. 2000), p. 7998. DOI: [10.1103/PhysRevB.62.7998](https://doi.org/10.1103/PhysRevB.62.7998).
- [456] Z. Wang, S.-S. Li, and L.-W. Wang. “Efficient Real-Time Time-Dependent Density Functional Theory Method and its Application to a Collision of an Ion with a 2D Material”. In: *Physical Review Letters* 114.6 (Feb. 2015), p. 063004. DOI: [10.1103/PhysRevLett.114.063004](https://doi.org/10.1103/PhysRevLett.114.063004).
- [457] P. Giannozzi et al. “QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials”. In: *Journal of Physics: Condensed Matter* 21.39 (Sept. 2009), p. 395502. DOI: [10.1088/0953-8984/21/39/395502](https://doi.org/10.1088/0953-8984/21/39/395502).
- [458] P. Giannozzi et al. “Advanced capabilities for materials modelling with Quantum ESPRESSO”. In: *Journal of Physics: Condensed Matter* 29.46 (Oct. 2017), p. 465901. DOI: [10.1088/1361-648X/aa8f79](https://doi.org/10.1088/1361-648X/aa8f79).
- [459] A. D. Corso. “Pseudopotentials periodic table: From H to Pu”. In: *Computational Materials Science* 95 (Dec. 2014), p. 337. DOI: [10.1016/j.commatsci.2014.07.043](https://doi.org/10.1016/j.commatsci.2014.07.043).
- [460] R. Dronskowski and P. E. Blöchl. “Crystal orbital Hamilton populations (COHP): energy-resolved visualization of chemical bonding in solids based on density-functional calculations”. In: *The Journal of Physical Chemistry* 97.33 (Aug. 1993), p. 8617. DOI: [10.1021/j100135a014](https://doi.org/10.1021/j100135a014).
- [461] V. L. Deringer, A. L. Tchougreeff, and R. Dronskowski. “Crystal Orbital Hamilton Population (COHP) Analysis As Projected from Plane-Wave Basis Sets”. In: *The Journal of Physical Chemistry A* 115.21 (June 2011), p. 5461. DOI: [10.1021/jp202489s](https://doi.org/10.1021/jp202489s).

- [462] S. Maintz et al. “LOBSTER: A tool to extract chemical bonding from plane-wave based DFT”. In: *Journal of Computational Chemistry* 37.11 (Apr. 2016), pp. 1030–1035. DOI: [10.1002/jcc.24300](https://doi.org/10.1002/jcc.24300).
- [463] D. I. Bilc et al. “Hybrid exchange-correlation functional for accurate prediction of the electronic and structural properties of ferroelectric oxides”. In: *Physical Review B* 77.16 (Apr. 2008), p. 165107. DOI: [10.1103/PhysRevB.77.165107](https://doi.org/10.1103/PhysRevB.77.165107).
- [464] Z. Sun et al. “Exploring a Lead-free Semiconducting Hybrid Ferroelectric with a Zero-Dimensional Perovskite-like Structure”. In: *Angewandte Chemie International Edition* 55.39 (Aug. 2016), p. 11854. DOI: [10.1002/anie.201606079](https://doi.org/10.1002/anie.201606079).
- [465] V. Polinger, P. Garcia-Fernandez, and I. Bersuker. “Pseudo Jahn–Teller origin of ferroelectric instability in BaTiO₃ type perovskites: The Green’s function approach and beyond”. In: *Physica B: Condensed Matter* 457 (Jan. 2015), p. 296. DOI: [10.1016/j.physb.2014.09.048](https://doi.org/10.1016/j.physb.2014.09.048).
- [466] K. Chang et al. “Discovery of robust in-plane ferroelectricity in atomic-thick SnTe”. In: *Science* 353.6296 (July 2016), p. 274. DOI: [10.1126/science.aad8609](https://doi.org/10.1126/science.aad8609).
- [467] S. S. R. K. C. Yamijala, R. Shinde, and B. M. Wong. “Real-time degradation dynamics of hydrated per- and polyfluoroalkyl substances (PFASs) in the presence of excess electrons”. In: *Physical Chemistry Chemical Physics* 22.13 (2020), pp. 6804–6808. DOI: [10.1039/c9cp06797c](https://doi.org/10.1039/c9cp06797c).
- [468] S. S. Yamijala et al. “Photo-induced degradation of PFASs: Excited-state mechanisms from real-time time-dependent density functional theory”. In: *Journal of Hazardous Materials* 423 (Feb. 2022), p. 127026. DOI: [10.1016/j.jhazmat.2021.127026](https://doi.org/10.1016/j.jhazmat.2021.127026).
- [469] E. M. Sunderland et al. “A review of the pathways of human exposure to poly- and perfluoroalkyl substances (PFASs) and present understanding of health effects”. In: *Journal of Exposure Science & Environmental Epidemiology* 29.2 (Nov. 2018), pp. 131–147. DOI: [10.1038/s41370-018-0094-1](https://doi.org/10.1038/s41370-018-0094-1).
- [470] C. Lau et al. “Perfluoroalkyl Acids: A Review of Monitoring and Toxicological Findings”. In: *Toxicological Sciences* 99.2 (May 2007), pp. 366–394. DOI: [10.1093/toxsci/kfm128](https://doi.org/10.1093/toxsci/kfm128).
- [471] M. Kotthoff et al. “Perfluoroalkyl and polyfluoroalkyl substances in consumer products”. In: *Environmental Science and Pollution Research* 22.19 (Feb. 2015), pp. 14546–14559. DOI: [10.1007/s11356-015-4202-7](https://doi.org/10.1007/s11356-015-4202-7).

- [472] J. S. Boone et al. “Per- and polyfluoroalkyl substances in source and treated drinking waters of the United States”. In: *Science of The Total Environment* 653 (Feb. 2019), pp. 359–369. DOI: [10.1016/j.scitotenv.2018.10.245](https://doi.org/10.1016/j.scitotenv.2018.10.245).
- [473] J.-C. Liou et al. “Investigating the biodegradability of perfluorooctanoic acid”. In: *Chemosphere* 80.2 (June 2010), pp. 176–183. DOI: [10.1016/j.chemosphere.2010.03.009](https://doi.org/10.1016/j.chemosphere.2010.03.009).
- [474] Y. Su et al. “Potential-Driven Electron Transfer Lowers the Dissociation Energy of the C–F Bond and Facilitates Reductive Defluorination of Perfluorooctane Sulfonate (PFOS)”. In: *ACS Applied Materials & Interfaces* 11.37 (Aug. 2019), pp. 33913–33922. DOI: [10.1021/acsami.9b10449](https://doi.org/10.1021/acsami.9b10449).
- [475] Y. Gu et al. “Efficient Reductive Decomposition of Perfluorooctanesulfonate in a High Photon Flux UV/Sulfite System”. In: *Environmental Science & Technology* 50.19 (Sept. 2016), pp. 10554–10561. DOI: [10.1021/acs.est.6b03261](https://doi.org/10.1021/acs.est.6b03261).
- [476] L. Duan et al. “Efficient Photocatalytic PFOA Degradation over Boron Nitride”. In: *Environmental Science & Technology Letters* 7.8 (June 2020), pp. 613–619. DOI: [10.1021/acs.estlett.0c00434](https://doi.org/10.1021/acs.estlett.0c00434).
- [477] W. Wang et al. “Photocatalytic reductive defluorination of perfluorooctanoic acid in water under visible light irradiation: the role of electron donor”. In: *Environmental Science: Water Research & Technology* 6.6 (2020), pp. 1638–1648. DOI: [10.1039/d0ew00205d](https://doi.org/10.1039/d0ew00205d).
- [478] O. C. Olatunde, A. T. Kuvarega, and D. C. Onwudiwe. “Photo enhanced degradation of polyfluoroalkyl and perfluoroalkyl substances”. In: *Helvion* 6.12 (Dec. 2020), e05614. DOI: [10.1016/j.helivion.2020.e05614](https://doi.org/10.1016/j.helivion.2020.e05614).
- [479] B. Xu et al. “Photocatalytic removal of perfluoroalkyl substances from water and wastewater: Mechanism, kinetics and controlling factors”. In: *Chemosphere* 189 (Dec. 2017), pp. 717–729. DOI: [10.1016/j.chemosphere.2017.09.110](https://doi.org/10.1016/j.chemosphere.2017.09.110).
- [480] M. Ohno et al. “Photochemical decomposition of perfluorooctanoic acid mediated by iron in strongly acidic conditions”. In: *Journal of Hazardous Materials* 268 (Mar. 2014), pp. 150–155. DOI: [10.1016/j.jhazmat.2013.12.059](https://doi.org/10.1016/j.jhazmat.2013.12.059).
- [481] J. J. Goings, P. J. Lestrangle, and X. Li. “Real-time time-dependent electronic structure theory”. In: *WIREs Computational Molecular Science* 8.1 (Sept. 2017). DOI: [10.1002/wcms.1341](https://doi.org/10.1002/wcms.1341).

- [482] M. R. Provorse and C. M. Isborn. “Electron dynamics with real-time time-dependent density functional theory”. In: *International Journal of Quantum Chemistry* 116.10 (Feb. 2016), pp. 739–749. DOI: [10.1002/qua.25096](https://doi.org/10.1002/qua.25096).
- [483] A. Raza et al. “A Machine Learning Approach for Predicting Defluorination of Per- and Polyfluoroalkyl Substances (PFAS) for Their Efficient Treatment and Removal”. In: *Environmental Science & Technology Letters* 6.10 (Sept. 2019), pp. 624–629. DOI: [10.1021/acs.estlett.9b00476](https://doi.org/10.1021/acs.estlett.9b00476).
- [484] J. J. Mortensen, L. B. Hansen, and K. W. Jacobsen. “Real-space grid implementation of the projector augmented wave method”. In: *Physical Review B* 71.3 (Jan. 2005). DOI: [10.1103/physrevb.71.035109](https://doi.org/10.1103/physrevb.71.035109).
- [485] A. Ojanperä et al. “Nonadiabatic Ehrenfest molecular dynamics within the projector augmented-wave method”. In: *The Journal of Chemical Physics* 136.14 (Apr. 2012), p. 144103. DOI: [10.1063/1.3700800](https://doi.org/10.1063/1.3700800).
- [486] GPAW. <https://wiki.fysik.dtu.dk/gpaw/documentation/tddft/timepropagation.html>.
- [487] C. Hu et al. “Photocatalysis Enhanced by External Fields”. In: *Angewandte Chemie International Edition* 60.30 (Jan. 2021), pp. 16309–16328. DOI: [10.1002/anie.202009518](https://doi.org/10.1002/anie.202009518).
- [488] M. Trtica et al. “Femtosecond laser-assisted surface modification of tungsten with 1015 W/cm² intensity in air and vacuum ambience”. In: *Applied Surface Science* 464 (Jan. 2019), pp. 99–107. DOI: [10.1016/j.apsusc.2018.09.074](https://doi.org/10.1016/j.apsusc.2018.09.074).
- [489] N. V. Ilawe, M. B. Oviedo, and B. M. Wong. “Real-Time Quantum Dynamics of Long-Range Electronic Excitation Transfer in Plasmonic Nanoantennas”. In: *Journal of Chemical Theory and Computation* 13.8 (July 2017), pp. 3442–3454. DOI: [10.1021/acs.jctc.7b00423](https://doi.org/10.1021/acs.jctc.7b00423).
- [490] N. V. Ilawe, M. B. Oviedo, and B. M. Wong. “Effect of quantum tunneling on the efficiency of excitation energy transfer in plasmonic nanoparticle chain waveguides”. In: *Journal of Materials Chemistry C* 6.22 (2018), pp. 5857–5864. DOI: [10.1039/c8tc01466c](https://doi.org/10.1039/c8tc01466c).
- [491] J. M. Rodríguez-Borbón et al. “Field Programmable Gate Arrays for Enhancing the Speed and Energy Efficiency of Quantum Dynamics Simulations”. In: *Journal of Chemical Theory and Computation* 16.4 (Mar. 2020), pp. 2085–2098. DOI: [10.1021/acs.jctc.9b01284](https://doi.org/10.1021/acs.jctc.9b01284).

- [492] C. Lian et al. “Indirect but Efficient: Laser-Excited Electrons Can Drive Ultrafast Polarization Switching in Ferroelectric Materials”. In: *The Journal of Physical Chemistry Letters* 10.12 (June 2019), pp. 3402–3407. DOI: [10.1021/acs.jpcllett.9b01046](https://doi.org/10.1021/acs.jpcllett.9b01046).
- [493] L. Xia et al. “Orientation Effect in the Low-Energy Electron Attachment to the Apolar Carbon Tetrafluoride Molecule”. In: *Angewandte Chemie International Edition* 52.3 (Nov. 2012), pp. 1013–1016. DOI: [10.1002/anie.201206948](https://doi.org/10.1002/anie.201206948).
- [494] L. McInnes et al. “UMAP: Uniform Manifold Approximation and Projection”. In: *Journal of Open Source Software* 3.29 (Sept. 2018), p. 861. DOI: [10.21105/joss.00861](https://doi.org/10.21105/joss.00861).
- [495] F. Pedregosa et al. “Scikit-learn: Machine learning in Python”. In: *Journal of Machine Learning Research* 12 (Oct. 2011), pp. 2825–2830.
- [496] K. R. Shahapure and C. Nicholas. “Cluster Quality Analysis Using Silhouette Score”. In: *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*. IEEE, Oct. 2020. DOI: [10.1109/dsaa49011.2020.00096](https://doi.org/10.1109/dsaa49011.2020.00096).
- [497] Kaya and Bilge. “Deep Metric Learning: A Survey”. In: *Symmetry* 11.9 (Sept. 2019), p. 1066. DOI: [10.3390/sym11091066](https://doi.org/10.3390/sym11091066).
- [498] K. Sohn. “Improved Deep Metric Learning with Multi-class N-pair Loss Objective”. In: *Advances in Neural Information Processing Systems*. Ed. by D. Lee et al. Vol. 29. Curran Associates, Inc., 2016, pp. 1857–1865.
- [499] J. V. Davis et al. “Information-theoretic metric learning”. In: *Proceedings of the 24th international conference on Machine learning - ICML '07*. ACM Press, 2007. DOI: [10.1145/1273496.1273523](https://doi.org/10.1145/1273496.1273523).
- [500] G. S. Na, H. Chang, and H. W. Kim. “Machine-guided representation for accurate graph-based molecular machine learning”. In: *Physical Chemistry Chemical Physics* 22.33 (2020), pp. 18526–18535. DOI: [10.1039/d0cp02709j](https://doi.org/10.1039/d0cp02709j).
- [501] G. M. Morris, R. Huey, and A. J. Olson. “Using AutoDock for Ligand-Receptor Docking”. In: *Current Protocols in Bioinformatics* 24.1 (Dec. 2008). DOI: [10.1002/0471250953.bi0814s24](https://doi.org/10.1002/0471250953.bi0814s24).
- [502] S. Salentin et al. “PLIP: fully automated protein–ligand interaction profiler”. In: *Nucleic Acids Research* 43.W1 (Apr. 2015), W443–W447. DOI: [10.1093/nar/gkv315](https://doi.org/10.1093/nar/gkv315).

- [503] O. Trott and A. J. Olson. “AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading”. In: *Journal of Computational Chemistry* (2009), pp. 455–461. DOI: [10.1002/jcc.21334](https://doi.org/10.1002/jcc.21334).
- [504] N. M. O'Boyle et al. “Open Babel: An open chemical toolbox”. In: *Journal of Cheminformatics* 3.1 (Oct. 2011). DOI: [10.1186/1758-2946-3-33](https://doi.org/10.1186/1758-2946-3-33).
- [505] A. Volz, C. Fonatsch, and A. Ziegler. “Regional mapping of the gene for autosomal dominant spinocerebellar ataxia (SCA1) by localizing the closely linked D6S89 locus to 6p24.2→p23.05”. In: *Cytogenetic and Genome Research* 60.1 (1992), pp. 37–39. DOI: [10.1159/000133291](https://doi.org/10.1159/000133291).
- [506] S. Leysen et al. “A Structural Study of the Cytoplasmic Chaperone Effect of 14-3-3 Proteins on Ataxin-1”. In: *Journal of Molecular Biology* 433.19 (Sept. 2021), p. 167174. DOI: [10.1016/j.jmb.2021.167174](https://doi.org/10.1016/j.jmb.2021.167174).
- [507] C. Sreedhar, N. Kasiviswanath, and P. C. Reddy. “Clustering large datasets using K-means modified inter and intra clustering (KM-I2C) in Hadoop”. In: *Journal of Big Data* 4.1 (Sept. 2017). DOI: [10.1186/s40537-017-0087-2](https://doi.org/10.1186/s40537-017-0087-2).
- [508] J. Schnabel, K. Weber, and M. Hatzfeld. “Protein–protein interactions between keratin polypeptides expressed in the yeast two-hybrid system”. In: *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1403.2 (June 1998), pp. 158–168. DOI: [10.1016/s0167-4889\(98\)00036-6](https://doi.org/10.1016/s0167-4889(98)00036-6).
- [509] M. Cruz-Montegudo et al. “Activity cliffs in drug discovery: Dr Jekyll or Mr Hyde?” In: *Drug Discovery Today* 19.8 (Aug. 2014), pp. 1069–1080. DOI: [10.1016/j.drudis.2014.02.003](https://doi.org/10.1016/j.drudis.2014.02.003).
- [510] D. Stumpfe et al. “Recent Progress in Understanding Activity Cliffs and Their Utility in Medicinal Chemistry”. In: *Journal of Medicinal Chemistry* 57.1 (Sept. 2013), pp. 18–28. DOI: [10.1021/jm401120g](https://doi.org/10.1021/jm401120g).
- [511] H. Hu and J. Bajorath. “Systematic Exploration of Activity Cliffs Containing Privileged Substructures”. In: *Molecular Pharmaceutics* 17.3 (Jan. 2020), pp. 979–989. DOI: [10.1021/acs.molpharmaceut.9b01236](https://doi.org/10.1021/acs.molpharmaceut.9b01236).
- [512] S. Ong et al. “Python Materials Genomics (pymatgen): A Robust”. In: *Open-Source Python Library for Materials Analysis. Comput. Mater. Sci* 68 (2013), pp. 314–319.
- [513] V. Fiorentini and M. Methfessel. “Extracting Convergent Surface Energies from Slab Calculations”. In: *J. Phys. Condens. Matter* 8 (1996), pp. 6525–6529.
- [514] M. Frisch et al. “Gaussian 16”. In: *Revision B.01* (2016).

- [515] G. Grüner. “The dynamics of spin-density waves”. In: *Reviews of Modern Physics* 66.1 (Jan. 1994), pp. 1–24. DOI: [10.1103/revmodphys.66.1](https://doi.org/10.1103/revmodphys.66.1).
- [516] A. Ojanperä et al. “Nonadiabatic Ehrenfest molecular dynamics within the projector augmented-wave method”. In: *The Journal of Chemical Physics* 136.14 (Apr. 2012), p. 144103. DOI: [10.1063/1.3700800](https://doi.org/10.1063/1.3700800).

Appendix A

The Diamine Cation is not a

Chemical Example where DFT fails

A.1 DMP Coords

Table A.1: Reference Cartesian coordinates (in Å) for the MP2-optimized DMP-L⁺ molecular cation.

N	0.8669040000	1.0704380000	0.0000000000
C	0.1612010000	0.6287860000	1.1938500000
C	0.1612010000	-0.9247560000	1.2290130000
N	-0.4616440000	-1.3789290000	0.0000000000
C	0.1612010000	-0.9247560000	-1.2290130000
C	0.1612010000	0.6287860000	-1.1938500000
C	1.0464460000	2.5339080000	0.0000000000
C	-1.8315630000	-1.8536180000	0.0000000000
H	-0.8886810000	0.9965970000	1.2408420000
H	0.6910660000	0.9796830000	2.0924050000
H	-0.3948770000	-1.3161980000	2.0904220000
H	1.2045650000	-1.2762920000	1.2365510000
H	0.6910660000	0.9796830000	-2.0924050000
H	-0.8886810000	0.9965970000	-1.2408420000
H	1.2045650000	-1.2762920000	-1.2365510000
H	-0.3948770000	-1.3161980000	-2.0904220000
H	1.6186930000	2.8236570000	0.8921360000
H	0.0779740000	3.0737480000	0.0000000000
H	1.6186930000	2.8236570000	-0.8921360000
H	-2.0156980000	-2.4438030000	-0.9065480000
H	-2.5030430000	-0.9717060000	0.0000000000
H	-2.0156980000	-2.4438030000	0.9065480000

Table A.2: Reference Cartesian coordinates (in Å) for the MP2-optimized DMP-D⁺ molecular cation.

N	-0.5345895389	1.3087803821	-0.0000000006
C	-0.6076389847	0.5258335438	1.1961899695
C	0.6076396600	-0.5258341301	1.1961895184
N	0.5345917563	-1.3087816692	-0.0000000008
C	0.6076396610	-0.5258341309	-1.1961895206
C	-0.6076389837	0.5258335428	-1.1961899716
C	0.3752547937	2.4554774414	0.0000000013
C	-0.3752553891	-2.4554764973	0.0000000006
H	-1.5467752809	-0.0494318097	1.2197931686
H	-0.5248208759	1.1610168822	2.0895997652
H	0.5248206257	-1.1610168488	2.0895996511
H	1.5467759018	0.0494313515	1.2197940032
H	-0.5248208747	1.1610168827	-2.0895997661
H	-1.5467752807	-0.0494318095	-1.2197931687
H	1.5467759027	0.0494313505	-1.2197940087
H	0.5248206252	-1.1610168511	-2.0895996516
H	0.1877028870	3.0620743686	0.8958603262
H	0.1877028889	3.0620743723	-0.8958603212
H	1.4356271683	2.1334602394	0.0000000015
H	-0.1877045140	-3.0620740720	-0.8958601190
H	-1.4356276341	-2.1334584711	0.0000000005
H	-0.1877045138	-3.0620740676	0.8958601228

Table A.3: Reference Cartesian coordinates (in Å) for the MP2-optimized transition state of the diamine molecular cation.

N	0.5169376764	1.1624721594	-0.0043853521
C	-0.0451819170	0.6525896513	1.1954852164
C	0.2162493170	-0.9379387657	1.2083443371
N	-0.3585603333	-1.4297639692	0.0041178155
C	0.2090314973	-0.9437533629	-1.2055308074
C	-0.0528041300	0.6474273750	-1.1981536662
C	1.2053913627	2.4578724505	-0.0095705479
C	-1.7689762346	-1.7788119017	0.0093772545
H	-1.1434665994	0.7816116871	1.2867889588
H	0.4504144731	1.0826153945	2.0777503663
H	-0.2657739082	-1.3862848409	2.0870216334
H	1.3034248099	-1.0878007401	1.2161733426
H	0.4369808769	1.0733476541	-2.0856411190
H	-1.1517443039	0.7752599973	-1.2828594262
H	1.2962575979	-1.0927655848	-1.2191034966
H	-0.2780782203	-1.3957409619	-2.0794962237
H	1.8456444047	2.5244848737	0.8812971522
H	0.5007534668	3.3079744178	-0.0101019721
H	1.8413753440	2.5197755019	-0.9038513631
H	-2.0076296657	-2.3591927147	-0.8907664415
H	-2.3812719858	-0.8530646307	0.0113746009
H	-2.0010665285	-2.3584946900	0.9117297381

Table A.4: Reference Cartesian coordinates (in Å) for the CCSD-optimized DMP-L⁺ molecular cation.

N	0.8492470000	1.0785750000	0.0000000000
C	0.1517520000	0.6387710000	1.1999580000
C	0.1517520000	-0.9190230000	1.2370970000
N	-0.4671290000	-1.3826540000	0.0000000000
C	0.1517520000	-0.9190230000	-1.2370970000
C	0.1517520000	0.6387710000	-1.1999580000
C	1.0988920000	2.5319720000	0.0000000000
C	-1.8380220000	-1.8789370000	0.0000000000
H	-0.9001810000	0.9999440000	1.2543460000
H	0.6867250000	0.9953650000	2.0940890000
H	-0.4120010000	-1.3090020000	2.0956900000
H	1.1937630000	-1.2766870000	1.2561590000
H	0.6867250000	0.9953650000	-2.0940890000
H	-0.9001810000	0.9999440000	-1.2543460000
H	1.1937630000	-1.2766870000	-1.2561590000
H	-0.4120010000	-1.3090020000	-2.0956900000
H	1.6855120000	2.7954320000	0.8933120000
H	0.1578220000	3.1203250000	0.0000000000
H	1.6855120000	2.7954320000	-0.8933120000
H	-2.0115770000	-2.4750280000	-0.9068480000
H	-2.5243930000	-1.0070060000	0.0000000000
H	-2.0115770000	-2.4750280000	0.9068480000

Table A.5: Reference Cartesian coordinates (in Å) for the CCSD-optimized DMP-D⁺ molecular cation.

N	0.2368380000	1.3948160000	0.0000000000
C	-0.2368380000	0.7727630000	1.1975510000
C	0.2368380000	-0.7727630000	1.1975510000
N	-0.2368380000	-1.3948160000	0.0000000000
C	0.2368380000	-0.7727630000	-1.1975510000
C	-0.2368380000	0.7727630000	-1.1975510000
C	1.6067020000	1.9205960000	0.0000000000
C	-1.6067020000	-1.9205960000	0.0000000000
H	-1.3384270000	0.7822810000	1.2287040000
H	0.1733680000	1.2649960000	2.0912800000
H	-0.1733680000	-1.2649960000	2.0912800000
H	1.3384270000	-0.7822810000	1.2287040000
H	0.1733680000	1.2649960000	-2.0912800000
H	-1.3384270000	0.7822810000	-1.2287040000
H	1.3384270000	-0.7822810000	-1.2287040000
H	-0.1733680000	-1.2649960000	-2.0912800000
H	1.7548980000	2.5393440000	0.8966320000
H	1.7548980000	2.5393440000	-0.8966320000
H	2.3550490000	1.1017830000	0.0000000000
H	-1.7548980000	-2.5393440000	0.8966320000
H	-1.7548980000	-2.5393440000	-0.8966320000
H	-2.3550490000	-1.1017830000	0.0000000000

Table A.6: Reference Cartesian coordinates (in Å) for the CCSD-optimized transition state of the diamine molecular cation.

N	0.6158602154	1.1338056095	-0.0078384890
C	0.0222013504	0.6417989195	1.2037815406
C	0.2031723820	-0.9330104953	1.2273619047
N	-0.3867270858	-1.4374022491	0.0089457548
C	0.1958489089	-0.9461387102	-1.2183074413
C	0.0149043271	0.6288482298	-1.2104989320
C	1.1698597856	2.4929227602	-0.0167553229
C	-1.8022556033	-1.7984516066	0.0151541513
H	-1.0667362027	0.8469400541	1.2847998802
H	0.5339765420	1.0688375637	2.0794892256
H	-0.3034891298	-1.3669072394	2.1008794022
H	1.2788744902	-1.1605953845	1.2438856424
H	0.5213270809	1.0464709553	-2.0938243751
H	-1.0745253827	0.8330695380	-1.2871295679
H	1.2714498571	-1.1738621298	-1.2387681730
H	-0.3159679063	-1.3893957778	-2.0840856952
H	1.7999951737	2.6294023607	0.8755362601
H	0.3769644107	3.2657865968	-0.0183697606
H	1.7943719501	2.6199133624	-0.9143821126
H	-2.0339073493	-2.3868172270	-0.8837006922
H	-2.4191921027	-0.8750508197	0.0128501283
H	-2.0280987115	-2.3783453106	0.9209766715

Appendix B

Harnessing Semi-Supervised

Machine Learning to

Automatically Predict Bioactivities

of Per- and Polyfluoroalkyl

Substances (PFASs)

B.1 Data Collection

Our selection of proteins focuses on targets described previously in Cheng et al. [83], and we screened the targets based on the availability of X-ray structures for our docking simulations (see Table B.1).

Table B.1: Selected targets (ion channels, enzymes, and receptors).

Target	PubChem AID	Description	PDB structure
CYP2C9	PCBA-883	Cytochrome p450 enzyme	1OG5
CYP3A4	PCBA-884	Cytochrome p450 enzyme	4D7D
CYP2D6	PCBA-891	Cytochrome p450 enzyme	5TFT
ATXN	PCBA-651635	Ataxin-2 protein	3KTR
K18	PCBA-1468	Protein-protein interaction	5HJC
VP16	PCBA-2546	Transcription factor	6T4X
ROR gamma	PCBA-2551	Nuclear receptor ROR gamma	6T4X

The bioactivity data from bioassays against 26 targets were extracted from Ref. [83].

B.2 Discussion of Unsupervised Learning Methods

For the first unsupervised learning step, we used *dimension reduction* methods to project the high-dimensional fingerprint input data onto a 2-dimensional space. Two different dimension reduction methods, (i) Principal Component (PC) Analysis (PCA) followed by t-Distributed Stochastic Neighbor Embedding (t-SNE), i.e., PC t-SNE, and (ii) UMAP (Uniform Manifold Approximation and Projection for Dimension Reduction) [494], were used on our fingerprint data. While t-SNE is a widely-used dimension reduction technique

for many types of data analysis, prior studies have shown that UMAP exhibits better clustering performance, especially for large datasets [494]. Thus, we decided to compare the clustering performance of both techniques.

After the dimension reduction procedure, we used the scikit-learn [494] library to execute three different *clustering methods*: k-means, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), and Hierarchical DBSCAN (HDBSCAN). While k-means clustering is relatively more time-efficient, DBSCAN and HDBSCAN can efficiently handle outliers and noisy datasets. Since each method has clear advantages, we used all three techniques and evaluated their performances. By combining dimension reduction and clustering methods, we classified/grouped various PFAS structures based on their similarities.

Two different dimension reduction methods, Principal Component Analysis (PCA) followed by t-Distributed Stochastic Neighbor Embedding (t-SNE), and Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP) [494] were used on the fingerprint data we generated. Three different clustering methods –k-means, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), and Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) –were used after dimension reduction. Our work utilized the PCA/PC, t-SNE, k-means, DBSCAN, and HDBSCAN in the scikit-learn [495] library.

B.3 Model Selection

The final models were selected based on the best Silhouette score, which analyzes the distances of each data point to its cluster and neighboring clusters. In short, a higher Silhouette score guarantees better performance in clustering. The Silhouette score $s(i)$ of a data x_i can be calculated by the following expression:

$$s(i) = \frac{b(i) - w(i)}{\max\{b(i), w(i)\}} \quad \text{with} \quad b(i) = \min_k\{B(i, k)\} \tag{B.1}$$

where $w(i)$ is the average distance from the i^{th} point to the other points in its cluster, and $B(i, k)$ is the average distance from the i^{th} point to points in another cluster k . The Silhouette score was also used as an essential statistical method to optimize hyperparameters [496].

B.4 Brief Discussion of Semi-Supervised Metric Learning

Deep metric learning (DML) is a widely used machine learning technique to investigate mixed data distributions before building prediction models. DML approaches have been accurately used in state-of-the-art computer vision technologies [497–499], and recently, Na et al. reported their usefulness in cheminformatics [500]. The central concept of DML is to make a prediction problem easy by creating a new vector representation and separating the data well depending on their target values (cf. Figure 4.2 in the main text).

The metric-learning algorithms compute/learn Mahalanobis distances, which are given by the expression:

$$D(x, x') = \sqrt{(Lx - Lx')^T(Lx - Lx')} \tag{B.2}$$

Based on the semi-supervised data, the metric learning problem is generally formulated as an optimization of Mahalanobis distances. That is, the metric-learning algorithm seeks to find the parameters of a distance function that optimizes some objective function measuring the agreement with the training data. In other words, we used metric-learning with a semi-supervised learning algorithm to learn a distance metric that places molecules with similar bioactivities close together and molecules with opposite bioactivities far away.

B.5 Molecular Docking Calculations

After identifying certain substructures that play a critical role in determining bioactivities, we simulated their interaction in the binding pocket of the targets to further probe the fundamental interactions giving rise to their enhanced bioactivity. To this end, AutoDock4 [92, 501, 502] and Autodock Vina (v1.5.6) [503] were used to carry out the molecular docking procedure and visualize the docked conformations of the ligand-protein complexes.

The structures of the proteins for the docking analysis were obtained from the RCSB PDB database (<https://www.rcsb.org/>). The other three-dimensional structures were constructed from SMILES using Open Babel (v2.4.0) [504]. Autodock4 was used to process the ligand and ligand interaction conformation analysis in the binding affinity analysis. The docking pockets were based on the global search of the protein to determine the optimal binding sites [501].

AutoDock4 uses a semiempirical free energy-force field scoring function to perform a quick and accurate evaluation of the binding energies of ligands to proteins using a two-

step approach. First, the intramolecular energetics of the transition from the unbound state to the bound form of the protein-ligand complex is estimated, and the intermolecular energetics of the bound complex are subsequently evaluated.

Table B.2: Optimized hyperparameters and Silhouette scores for various unsupervised machine learning models.

Dimension Reduction		Clustering			Silhouette Score
Model	Hyperparameters	Model	Hyperparameters	n of Clusters	
PC t-SNE	$n_{components} = 60$ perplexity= 40 $n_{iterations} = 5000$	K-means	$n_{clusters} = 70$	70	0.469
	$n_{components} = 110$ perplexity= 30 $n_{iterations} = 1000$	DBSCAN	eps= 0.07 $min_{samples} = 3$	137	0.379
	$n_{components} = 50$ perplexity= 30 $n_{iterations} = 1000$	HDBSCAN	eps= 0.09 $min_{samples} = 3$	63	0.414
UMAP	$n_{neighbor} = 20$	K-means	$n_{clusters} = 10$	10	0.577
	$n_{neighbor} = 30$	DBSCAN	eps= 0.03 $min_{samples} = 3$	77	0.456
	$n_{neighbor} = 30$	HDBSCAN	eps= 0.07 $min_{samples} = 3$	59	0.469

B.6 Unsupervised Machine Learning Results

In this section, we compare the performance of each machine learning algorithm in classifying/clustering the PFAS molecules based on their bioactivity (see Table B.2 for

their parameters and settings.) The parameters for each model are optimized for the best performance (i.e., the highest Silhouette score). As mentioned in the methods section, we tested six unsupervised learning models by combining two different dimension reduction methods (PC t-SNE and UMAP) with three different clustering methods (k-means, DBSCAN, and HDBSCAN). Table B.2 shows that the combination of k-means clustering and the UMAP model exhibited better performance (with a Silhouette score of 0.577) than the other machine-learning methods.

Figure B.1 displays the distribution based on the molecular structures of C3F6 datasets. First, the molecular structures were converted into constant-length arrays by ECFP. Next, the vector representations were projected onto a 2-dimensional space using the PC t-SNE dimension reduction methods (molecules that are grouped closer together indicate structural similarity). These molecules were subsequently clustered into ten groups using k-means clustering. We then used the bioactivity data from each target to analyze the molecules with respect to promising substructures leading to bioactivity.

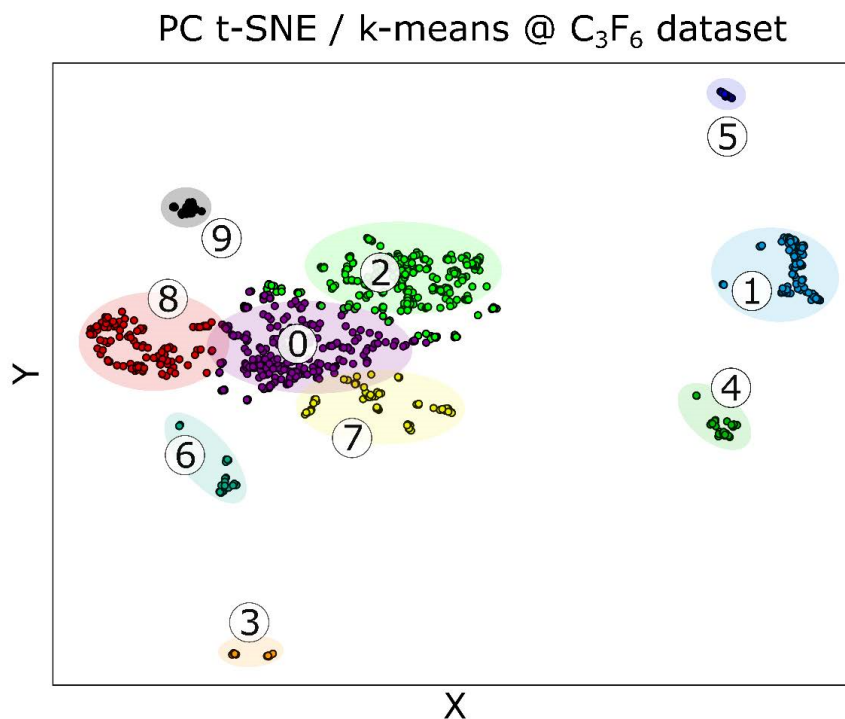
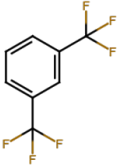
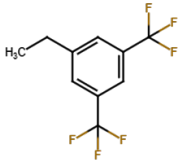


Figure B.1: Clustering of molecules in the C_3F_6 dataset. PFAS molecules are represented on a 2-dimensional latent space using the PC t-SNE algorithm. Each point represents a molecule, and the colors of the points designate clusters classified using k-means clustering.

We found that 90.6%, 76.9%, and 70.0% of molecules in cluster 1 are bioactive on CYP2C9, CYP2D6, and CYP3A4, respectively. CYP2C9, CYP2D6, and CYP3A4 are members of Cytochrome p450 enzymes (Cyps) involved in metabolism processes by oxidizing xenobiotics in the human body. Cyps are major phase-I xenobiotic-metabolizing enzymes induced by many environmental xenobiotics and drugs [96]. Also, 42.9% of molecules in cluster 4 are bioactive on ATXN, a DNA-binding protein [505, 506]. We only show targets with the top 4 true-positive rates of 90.6%, 76.9%, 70.0%, and 42.9%. The true positive rate is the probability that an actual positive (bioactive molecule) will be predicted positive

(bioactive). Table B.3 shows the maximum common substructures of the clusters.

Table B.3: Cluster number, accuracy, and maximum common structure most likely to be found in bioactive molecules toward each target. Most molecules with 1,3-bis(trifluoromethyl) benzene are bioactive toward the CYP enzymes, while molecules with 1-ethyl-3,5-bis(trifluoromethyl) benzene are likely to be bioactive toward ATXN.

Target	Cluster	True-Positive Rate	Maximum Common Substructure
CYP2C9	1	90.6%	
CYP2D6	1	76.9%	
CYP3A4	1	70.0%	1,3-bis(trifluoromethyl)
ATXN	4	42.9%	 1-ethyl-3,5-bis(trifluoromethyl) benzene

Even though our unsupervised learning approaches could automatically classify PFASs into a reasonable number of clusters, it is essential to note that not all 26 targets clustered well (i.e., only ATXN, CYP2C9, CYP3A4, and CYP2D6 showed reasonable accuracy). From this result, we inferred that the bioactivity of PFAS against ATXN and Cyps have the strongest correlation with their molecular structures.

Furthermore, the larger CF dataset (containing 62,043 molecules) did not have a higher clustering performance than the smaller C3F6 dataset (containing 1,012 molecules) since traditional unsupervised learning clustering methods, such as k-means, work better when applied to smaller datasets [507]. Because the CF dataset is 50 times larger than the C3F6 dataset and screened based on less rigorous criteria, we expect it to be more

challenging to handle with unsupervised learning techniques.

For example, Figure B.2a shows the bioactivities of molecules in the CF dataset toward keratin18 (K18), an intermediate filament protein [508]. Our unsupervised machine learning failed to successfully cluster the PFAS molecules based on their bioactivity toward K18, demonstrating a true-positive rate of only 55.9%. Similarly, Figure B.2c shows the bioactivity toward CYP2C9 predicted with unsupervised learning in which only 4.0% of the cluster showed bioactivity. These results highlight the inherent limitation of purely unsupervised learning for molecular structures. Since bioactivities are sensitive to minimal changes in molecular structures, they exhibit a mixed distribution in the molecular structure space. In other words, a pair of similar-structured molecules can have entirely different bioactivities, which are commonly referred to as activity cliffs in cheminformatics [509–511].

To overcome the limitations of PC t-SNE and UMAP, we utilized semi-supervised metric learning to produce a more distinct separation between bioactive and inactive molecules toward K18, which can be seen in Figure B.2b, which gives a true-positive rate of 79.2%. Similarly, Figure B.2c shows the bioactivity toward CYP2C9 predicted by unsupervised learning in which, again, only 4.0% of the cluster showed bioactivity. Finally, Figure B.2d shows the clustering based on semi-supervised metric learning, where 99.7% of a cluster’s molecules are bioactive (i.e., a true positive rate of 99.7%).

It is not surprising that semi-supervised metric learning demonstrated significantly higher performance for PFAS QSAR classifications since it uses partially labeled data. In contrast, an unsupervised learning model takes unlabeled data as input. Furthermore,

the performance difference is more significant with larger data, considering that traditional unsupervised learning methods, including k-means clustering, work better when applied to a smaller dataset [507]. Semi-supervised metric learning uses partially labeled bioactivity data as input data, placing molecules with similar bioactivities closer and opposite bioactivities further away. On the other hand, unsupervised learning tries to predict bioactivities solely based on the chemical structure, making it extremely difficult to complete an appropriate QSAR task.

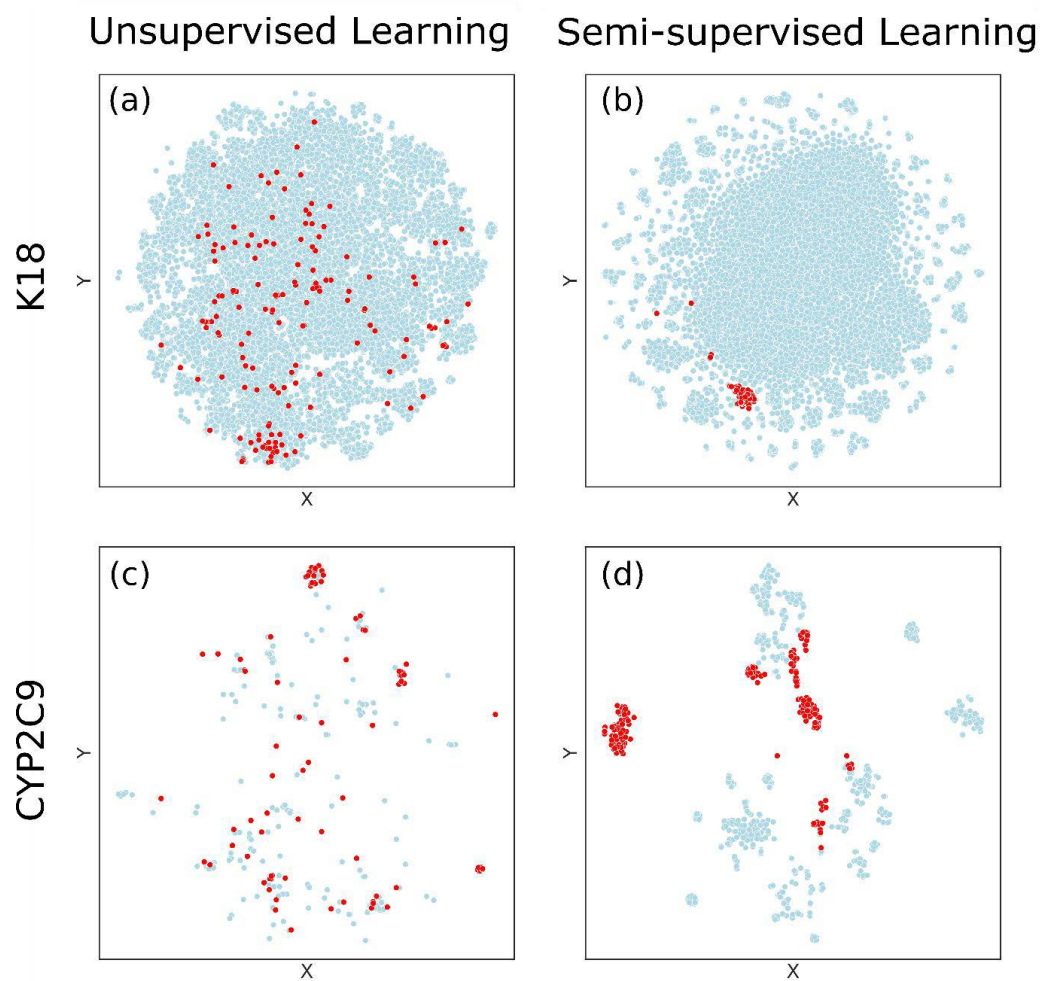
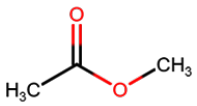
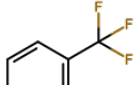
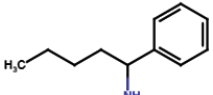
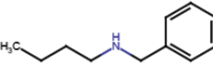
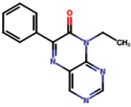
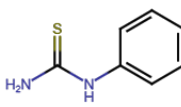
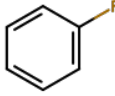
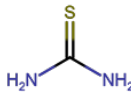
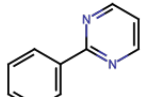
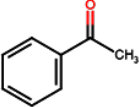
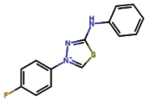
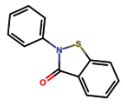


Figure B.2: Clustering of molecules in the CF dataset. Each point represents a molecule that is either bioactive (red) or inactive (blue) towards (a, b) K18 and (c, d) CYP2C9. The molecules are visualized on a 2-dimensional space using (a, c) PC t-SNE (unsupervised) or (b, d) semi-supervised metric learning.

Table B.4: Predicted substructures that induce bioactivity toward each target, based on semi-supervised metric learning.

Substructure	Target	Substructure	Target
 Ester	CYP2C9 CYP3A4	 Trifluoromethylbenzene	CYP3A4 K18
 1-phenylpentan-1-amine	CYP2C9 CYP3A4	 Benzyl(butyl)amine	CYP3A4
 6-phenyl-8-ethyl-7(8H)-pteridinone	CYP2C9 CYP3A4	 Phenylthiocarbamide	VP16 ROR Gamma
 Fluorobenzene	CYP2C9 CYP3A4	 Thiocarbamide	ALDH1A1, VP16 ROR Gamma
 Phenylprimidine	CYP2D6	 Phenylethanone	CYP2D6
 4-benzyl-2-(4-fluorophenyl)-1,2-thiazole	ATXN	 2-phenyl-2,3-dihydro-1,2-benzothiazol-3-one	ATXN

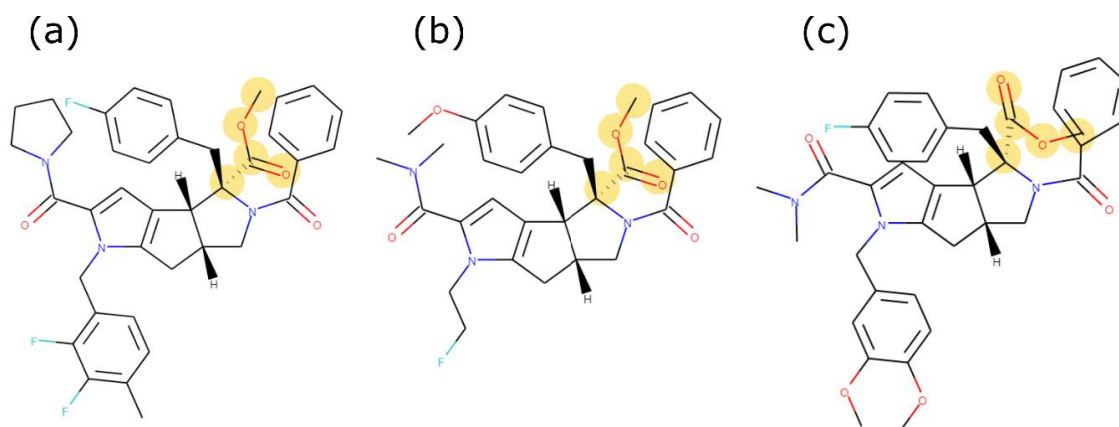


Figure B.3: Representative molecules found in the ChEMBL database that are bioactive on CYP2C9. The ChEMBL IDs of the molecules are (a) CHEMBL1411743, (b) CHEMBL1411201, and (c) CHEMBL1332759. The yellow circles represent the functional groups resulting in a structural alert of each molecule.

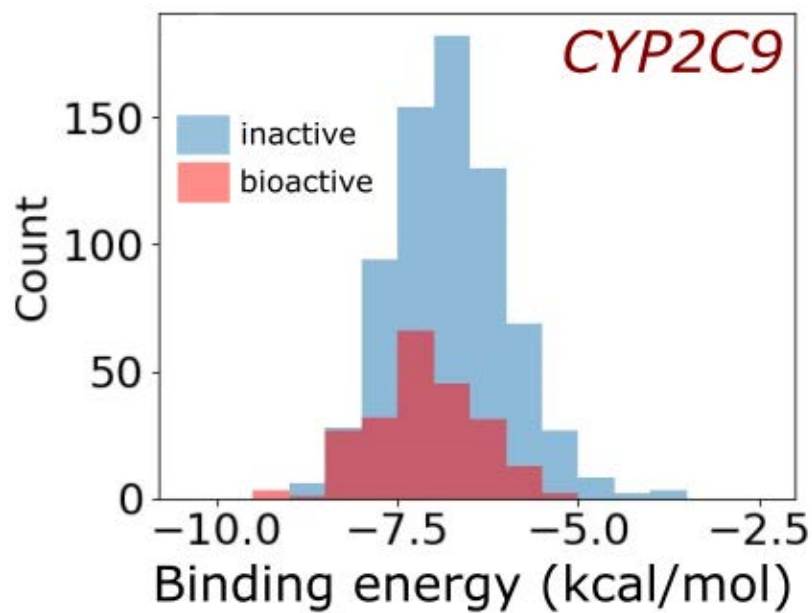


Figure B.4: Histogram of binding energies between molecules in the CF dataset and CYP2C9. The histogram bins are organized by binding energies, where each bin represents the number of bioactive (red) and inactive (blue) molecules.

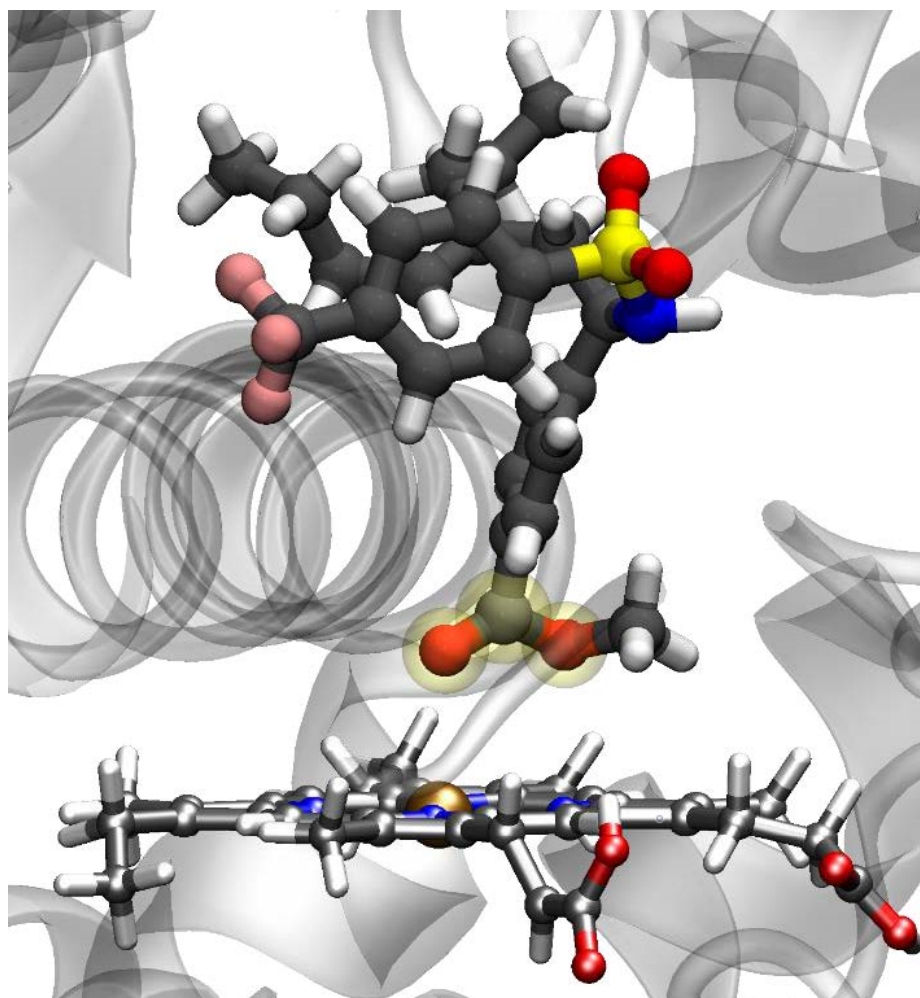


Figure B.5: Interactions between CYP2C9 and an ester-group-containing molecule (methyl 4-[2-propyl-1-([4-trifluoromethyl]phenyl)sulfonylamino)-2-hexen-1-yl]benzoate), as predicted with Autodock. The ester group is highlighted with yellow spheres. The planar structure at the bottom represents the HEME group in CYP2C9. The C, N, O, F, and Fe atoms are colored grey, blue, red, pink, and brown, respectively.

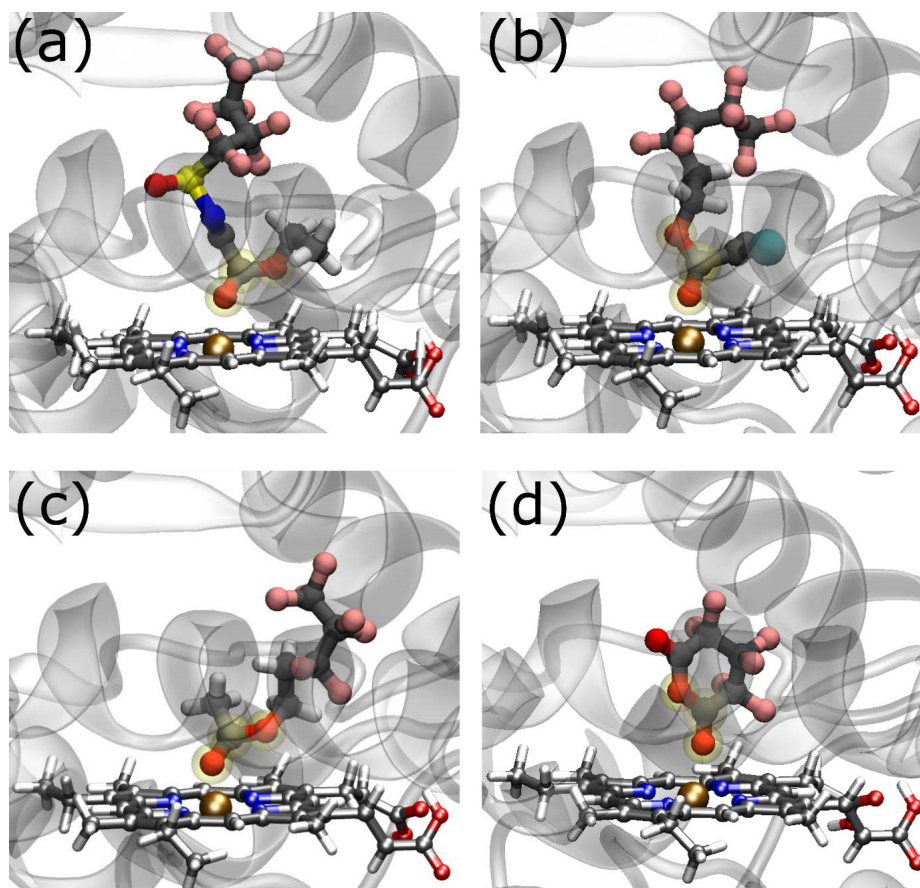


Figure B.6: Interactions between CYP and ester-containing PFAS ligands, (a) ethyl-[(nonafluorobutyl)sulfinyl]iminoacetate, (b) nonafluorohexyl2-chloroacrylate, (c) heptafluoropentyl acetate, and (d) hexafluorooxane-2,6-dione from the OECD list, as predicted with Autodock. The ester groups are highlighted with yellow spheres. The planar structure at the bottom represents the HEME group of CYP2C9. The C, N, O, F, and Fe atoms are colored grey, blue, red, pink, and brown, respectively.

Appendix C

Acceleration vs Accuracy:

Influence of Basis Set

C.1 Additional Computational Details

C.1.1 Generation of the ZnO Slab Along the $(10\bar{1}0)$ Direction

Across the three software packages, we first optimized both the cell parameters and the coordinates of bulk ZnO such that the forces on each atom were less than 0.01 eV/Å. The converged cell parameters obtained with each software package are given in Table C.1. Clearly, the cell parameters obtained with all of the software packages compare reasonably well with differences < 0.01 Å.

Table C.1: Optimized cell parameters of bulk ZnO obtained with each software package.

Software Package	a (Å)	b (Å)	c (Å)
CP2K	3.280	3.280	5.294
FHI-aims	3.288	3.288	5.296
VASP	3.284	3.284	5.293

The corresponding bulk cell parameters and coordinates obtained from each software package were further used to construct a 4-layer 4x4 slab along the $(10\bar{1}0)$ direction using pymatgen [512]. In all of the slab calculations, we have used more than 15 Å of vacuum space to avoid any spurious interactions between the periodic images. For these 4-layer-thick 4x4 slabs, we have fixed the coordinates of the bottom two layers to their respective bulk positions and re-optimized the coordinates of the top two layers such that the forces on each atom were less than 0.01 eV/Å. The same procedure was also used to generate 4-layer-thick 2x1 and 6x3 slabs. For a 6x3 slab, the simulation box dimensions were $\mathbf{a} = 19.68$ Å, $\mathbf{b} = 15.879$ Å, $\mathbf{c} = 32.8$ Å, and the angles were $\alpha = 90^\circ$, $\beta = 120^\circ$, $\gamma =$

90°.

C.1.2 Various Convergence Tests

Layer Convergence

It is important to mention that we used a 4-layer-thick slab in our calculations since we confirmed that the surface energy of the slab converges with four layers (see Table C.2). The layer convergence tests were performed with a 2x1 slab using the FHI-aims package with the Tier-1 basis set and with a 6x6x1 Monkhorst-Pack k-mesh (see Table C.3 for the k-mesh convergence).

The surface energy of a slab was computed as [513]

$$\gamma = \frac{E_{Slab} - NE_{Bulk}}{2A}, \quad (\text{C.1})$$

where E_{Bulk} is the energy of a ZnO pair in the bulk ZnO, E_{Slab} is the total energy of the slab, A is the area of the slab, and N is the number of ZnO pairs in the slab.

Table C.2: Layer convergence test for a 2x1 slab with a 6x6x1 k-mesh (single-point calculations).

Number of Layers	Absolute Energy (eV)	Surface Energy (eV·Å ²)
1	-201853.2929	0.06333
2	-403708.6318	0.06792
3	-605564.1243	0.06810
4	-807419.6268	0.06799
5	-1009275.1263	0.06797
6	-1211130.6217	0.06807
7	-1412986.1206	0.06806

K-point Convergence

Table C.3: K-point convergence test for a 2x1 slab with 4 layers (single-point calculations).

K-mesh	Absolute Energy (eV)	Energy Difference Between Consecutive kmesh Grids (eV)
1x1x1	-807420.1033	
2x2x1	-807419.8309	0.2724
3x3x1	-807419.8536	0.0228
4x4x1	-807419.8554	0.0018
5x5x1	-807419.8556	0.0002
6x6x1	-807419.8556	0.0000
7x7x1	-807419.8556	0.0000
8x8x1	-807419.8556	0.0000
9x9x1	-807419.8556	0.0000
10x10x1	-807419.8556	0.0000

Vacuum Convergence

As shown in Table C.4, the vacuum is converged (within 1 meV accuracy) for all values above 10 Å. However, to ensure convergence, in all of the adsorption energy calculations, we have used more than 15 Å of vacuum space.

Table C.4: Vacuum convergence tests for a 2x1 slab with a 6x6x1 k-mesh and with 4 layers (single-point calculations).

Vacuum Space (\AA)	Energy (eV)	Energy Difference (eV)
10	-807415.007	
11	-807415.007	-0.0003
12	-807415.007	-0.0003
13	-807415.007	-0.0002
14	-807415.008	-0.0002
15	-807415.008	-0.0002
16	-807415.008	-0.0002
17	-807415.008	-0.0002
18	-807415.008	-0.0002
19	-807415.008	-0.0001
20	-807415.008	0.0000

Basis Set Convergence

We have performed additional adsorption energy calculations with the FHI-aims package using the Tier-3 basis (this is the largest basis set available in FHI-aims and is equivalent to a quadruplezeta level basis), and these results were compared against the Tier-1 and Tier-2 basis set results (see Table C.5). Although the difference between the Tier-1 and Tier-2 results is 90 meV, the difference in the adsorption energies between Tier-2 and Tier-3 is less than 5 meV. Thus, one can consider the results obtained with the Tier-2 basis to be already converged.

Table C.5: Single-point adsorption energy of CO on the ZnO surface calculated with three different basis set tiers. In all of these adsorption energy calculations, the geometries of ZnO, CO, and ZnO+CO were fixed at the optimized geometry of ZnO+CO obtained using the Tier-2 basis.

System	Adsorption Energy with the Tier-1 basis, E_{Tier1} (eV)	Adsorption Energy with the Tier-2 basis, E_{Tier2} (eV)	Adsorption Energy with the Tier-3 basis, E_{Tier3} (eV)
Optimized “ZnO Slab + CO” with the Tier-2 basis	-0.6410	-0.5517	-0.5468
Difference in Adsorption Energy with Respect to Tier-2 Basis (eV)	-0.0893	0.0000	0.0049

C.1.3 Generation of Adsorbate Configurations

For each molecule, we considered 14 different configurations on a 4-layer-thick 2x1 slab and optimized all of them to identify the most stable adsorbate configuration. All 56 calculations (we considered 4 different molecules and 14 configurations for each) were performed using the FHI-aims package with the Tier-1 basis set and a 6x6x1 Monkhorst-Pack k-mesh. It is important to mention that of the 14 initial configurations considered for each molecule, we observed that some of them converged to similar final configurations (with a \sim 2-10 meV change in the energy). The most stable configuration obtained for each molecule at this level of theory (FHI-aims, Tier-1) was used as the initial configuration for all of the 4-layer thick 4x4 slab calculations considered in this study (across all software packages). An example pymatgen script for generating a 4-layer-thick slab and all the CO adsorption configurations used in this study is given below (see Table C.6).

Table C.6: Script for generating a 4-layer-thick slab and all the different adsorbate configurations for a CO molecule on a 2x1 slab.

```
1      #!/usr/bin/env python
2      import os
3      import pymatgen as pmg
4      from pymatgen.core.surface import SlabGenerator, generate_all_slabs,
5      Structure, Lattice
6      from pymatgen.io.cif import CifParser, CifWriter
7      from pymatgen import Molecule
8      from pymatgen.io.xyz import XYZ
9
10     #Read molecule from xyz file
11     mol = Molecule.from_file("CO.xyz")
12
13     #change these with the corresponding optimized lattice parameters
14     lattice = Lattice.hexagonal(3.28815084, 5.29625273)
15     zno = Structure(lattice,
16     ["Zn", "Zn", "O", "O"],
17     [[0.66692231, 0.33304453, 0.98286115],
18     [0.33307833, 0.66695668, 0.48263409],
19     [0.66694228, 0.33308454, 0.36198217],
20     [0.33305804, 0.66691602, 0.86220921]])
21
22     slab = SlabGenerator(zno,(1,0,0), 9.3, 20, center_slab=True).get_slab()
23     slab.to(filename="zno_1x1_cell.cif")
24
25     ads_sites = AdsorbateSiteFinder(slab)
26     nAdsConfigs = len(ads_sites.generate_adsorption_structures(mol,repeat=(2,1,1)))
27     for i in range(nAdsConfigs):
28     molcif=str(i)+".cif"
29     slabWithCO = ads_sites.generate_adsorption_structures(mol,repeat=(2,1,1))[i]
30     slabWithCO.make_supercell([1,1,1]) # Thus, generating a 2x1x1 slab
31     slabWithCO.to(filename=molcif)
32
```

C.1.4 Additional Tests for Adsorption Energy Calculations

With the above-mentioned initial configurations, all the slab-plus-adsorbate systems were further optimized (again, by fixing the coordinates of the bottom two layers) until their forces were less than 0.01 eV/Å. To keep all calculations on the same footing, we have also optimized the adsorbates in the same 4x4 supercell (i.e., the same cell parameters were used for the slab, adsorbate, and slab-plus-adsorbate). These optimized structures were then used to calculate the adsorption energies. As mentioned in the main text, in all of these calculations across the software packages, the Γ -point approximation was used for the Brillouin zone integrations. However, we have verified (with FHI-aims) that even with an increment in the k-mesh, the reported adsorption energies (not the absolute energies) were accurate within 1 meV (see Table C.7).

Table C.7: Difference in the adsorption energy (eV) of a CO molecule on a 4x4 ZnO slab with a change in the k-mesh size (single-point calculations).

System	K-mesh	Absolute Energy (eV)	Energy Difference Between Consecutive k-mesh Grids (eV)	Adsorption Energy (eV)	Difference in the Adsorption Energy Between Consecutive k-mesh Grids (eV)
CO	1x1x1	-3081.469821			
4x4 Slab	1x1x1	-6459358.7319			
4x4 Slab	2x2x1	-6459358.8657	0.1338		
4x4 Slab	3x3x1	-6459358.8660	0.0004		
4x4 Slab	4x4x1	-6459358.8660	0.0000		
4x4 Slab	5x5x1	-6459358.8660	0.0000		
4x4 Slab + 1 CO	1x1x1	-6462440.9091		-0.7075	
4x4 Slab + 1 CO	2x2x1	-6462441.0421	0.1330	-0.7066	0.0009
4x4 Slab + 1 CO	3x3x1	-6462441.0424	0.0003	-0.7066	0.0000
4x4 Slab + 1 CO	4x4x1	-6462441.0424	0.0000	-0.7066	0.0000
4x4 Slab + 1 CO	5x5x1	-6462441.0424	0.0000	-0.7066	0.0000

Furthermore, we have performed three sets of additional calculations to check the convergence of our CP2K-NVE results as a function of the number of k points. For

these calculations, we considered three representative geometries from the NVE simulation performed with the TZ basis and with H₂O colliding with the substrate at a 6 eV incident energy (Figure 6.4b of the main text). These three geometries, which correspond to the geometries at the 125, 1000, and 2000 timesteps of the NVE simulation, are given in Figure C.1. At the 125th step, the 4th CO and H₂O interact with each other, and the 4th CO is still bound to its original Zn-site. At the 1000th step, the CO moves away from its original Zn site but is still relatively close to the ZnO surface. Also, by this time step, the H₂O is in its dissociated form on the ZnO surface. At the 2000th step, the CO is relatively far away from the ZnO surface.

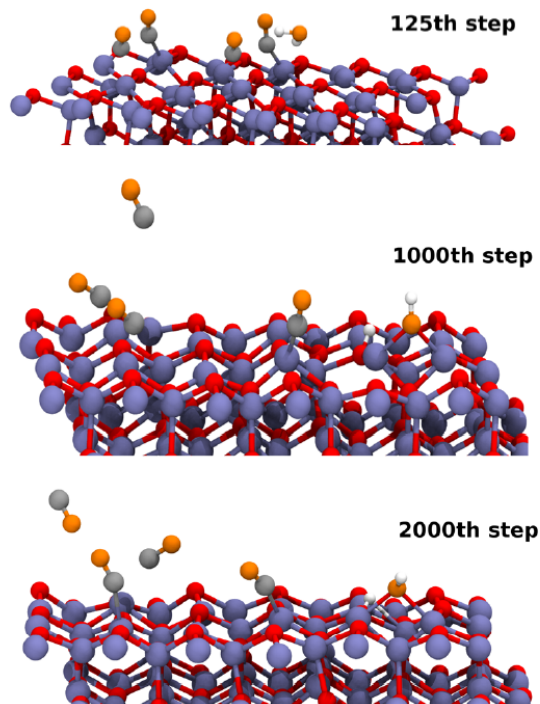


Figure C.1: Geometries at the 125, 1000, and 2000 time steps of the NVE simulation with 6 eV of collision energy using the TZ basis set.

For each of the above three geometries, we have performed k-mesh convergence

tests with 1x1x1, 2x2x1, and 4x4x1 Monkhorst k-grids, and the results are presented in Table C.8. As shown in the last column of Table C.8, the change in the interaction energy of the 4th CO with the rest of the system (three CO molecules and one H₂O molecule on a 4-layer 6x3 ZnO supercell) with a change in the k-mesh is ~ 5 meV for all the three geometries.

Thus, these results clearly suggest that, for this system, performing BOMD calculations at the Γ -point alone is sufficient. Here, the interaction energy is calculated by subtracting the

entire system’s energy ($E_{4\text{CO}+1\text{H}_2\text{O}+6\text{x}3\text{Slab}}$) from the configuration without the 4th CO molecule ($E_{1\text{H}_2\text{O}+6\text{x}3\text{Slab}}$) and the energy of the isolated CO molecule (E_{CO}). All of these calculations were single-point energies and were performed using the TZ basis. Finally, these results also compare reasonably well with our k -mesh convergence tests performed with the FHI-aims package.

Table C.8: Difference in the interaction energy (eV) of the 4th CO molecule with the rest of the system (three CO molecules, one H₂O molecule, and the 6x3 ZnO slab) with a change in the k-mesh size (single-point energy calculations) at three different time steps of the NVE simulation (with **6 eV** of collision energy).

Step Number	Monkhorst K-grid	$E_{(4\text{CO}+1\text{H}_2\text{O}+6\text{x}3\text{Slab})}$ (eV)	$E_{(3\text{CO}+1\text{H}_2\text{O}+6\text{x}3\text{Slab})}$ (eV)	Interaction Energy (eV)	Interaction Energy Difference w.r.t Γ Point Energy (meV)
125	1x1x1	-302924.479	-302334.237	-0.291	0.0
125	2x2x1	-302924.644	-302334.403	-0.289	1.9
125	4x4x1	-302924.644	-302334.407	-0.286	5.1
1000	1x1x1	-302923.091	-302333.104	-0.035	0.0
1000	2x2x1	-302923.250	-302333.264	-0.035	0.1
1000	4x4x1	-302923.248	-302333.261	-0.035	0.0
2000	1x1x1	-302922.033	-302332.058	-0.024	0.0
2000	2x2x1	-302922.194	-302332.218	-0.024	0.3
2000	4x4x1	-302922.190	-302332.215	-0.024	0.1

C.2 Additional Tests with CP2K

C.2.1 Basis Sets

In CP2K, the molecularly optimized basis sets for GTH pseudopotentials of double- and triplezeta quality can be of two types, namely, standard and short-ranged. For example, a double-zeta quality basis can be either DZVP or DZVP-SR. Here, SR represents the short-range nature of the basis set, and these basis sets are designed to speed up the BOMD calculations with minimal changes in accuracy compared to the standard/normal basis sets. Currently, in CP2K, the molecularly optimized basis sets for Zn are always short-ranged; i.e., we have only DZVP-SR and TZV2P-SR basis sets for Zn. For the other elements (O, H, C, and N) at the double-zeta quality, we have both DZVP and DZVP-SR basis sets. However, at the triple-zeta level, we have only the TZV2P basis set. As such, while using the triple-zeta quality basis sets, we are bound to use TZV2P-SR for Zn and TZV2P for all the other elements (without any choice). However, while using double-zeta quality basis sets, we can either use DZVP-SR for all the elements or just DZVP-SR for Zn and DZVP for all the other elements (similar to the triple-zeta case). For completeness, we have considered both of these approaches in our adsorption energy calculations, and the results are given in Table C.9. To compare the results across basis sets, we have also given the difference in the adsorption energy between TZV2P and DZVP-SR ($E_{\text{diff}}^{\text{TD-SR}}$) as well as TZV2P and DZVP ($E_{\text{diff}}^{\text{TD}}$) in the same table. To give a general trend, we have also considered two more molecules, namely, NO and N₂ in our calculations. From Table C.9, it is evident that DZ quality basis sets always yield stronger binding energies than TZ quality basis sets. However, we did not observe any general trends in the accuracy of one

basis over the other between DZVP and DZVP-SR. For example, the adsorption energy of CO predicted with the DZVP-SR basis is closer to the TZV2P basis than that predicted by the DZVP basis, with the reverse being true for the case of H₂.

In general, these CP2K results (and also the FHI-aims results in the main paper) suggest that the barrier heights predicted with a particular basis set cannot be translated to another basis even when the calculations are performed with the same software package and with the same functional. Furthermore, the predicted outcomes of a reaction at one basis set can be completely different from that of the other basis set, and this is evident by a 0.05 to 0.4 eV change in the adsorption energies for most of the adsorbate molecules considered in this study. For any reaction involving CO on ZnO, the current results show that the barrier height of the reaction can be reduced by as much as ~ 0.4 eV (~ 0.1 eV) by using a higher-quality basis set with the CP2K (FHI-aims) software package. Thus, these results strongly suggest the importance of considering larger basis sets in calculations involving barrier heights.

Going back to the results in Table C.9, although using the TZV2P, DZVP, or DZVP-SR basis sets can yield similar results in certain cases, it is difficult to know a priori when they will be similar. As such, it is advisable to compare the binding energies obtained with the DZ quality basis against those calculated with the TZ quality basis before performing any MD/collision calculations with the DZ quality basis (as also mentioned in the main text).

Table C.9: Adsorption energies of molecules calculated using CP2K with 3 different basis sets and the differences in the adsorption energies between basis sets. Here, DZVP, DZVP-SR, and TZV2P correspond to the basis set used for the elements O, H, C, and N. For Zn, only short-ranged basis sets are used in all of the cases.

System (basis-set) [4x4 denotes a 4x4 ZnO slab]	Adsorption Energy (eV)	Adsorption Energy Difference $E_{\text{TZV2P}} - E_{\text{DZVP}}$ (eV)	Adsorption Energy Difference $E_{\text{TZV2P}} - E_{\text{DZVP-SR}}$ (eV)
4x4+CO (DZVP-SR)	-0.688		
4x4+CO (DZVP)	-0.884		
4x4+CO (TZV2P)	-0.504	0.380	0.184
4x4+CO ₂ (DZVP-SR)	-0.200		
4x4+CO ₂ (DZVP)	-0.177		
4x4+CO ₂ (TZV2P)	-0.175	0.002	0.025
4x4+H ₂ (DZVP-SR)	-0.180		
4x4+H ₂ (DZVP)	-0.088		
4x4+H ₂ (TZV2P)	-0.083	0.005	0.097
4x4+H ₂ O (DZVP-SR)	-0.989		
4x4+H ₂ O (DZVP)	-0.990		
4x4+H ₂ O (TZV2P)	-0.961	0.029	0.028
4x4+N ₂ (DZVP-SR)	-0.113		
4x4+N ₂ (DZVP)	-0.117		
4x4+N ₂ (TZV2P)	-0.101	0.016	0.012
4x4+NO (DZVP-SR)	-0.241		
4x4+NO (DZVP)	-0.237		
4x4+NO (TZV2P)	-0.231	0.006	0.010

C.2.2 Dispersion Effect

Table C.10: Adsorption energies of molecules calculated using CP2K with 2 different basis sets and the differences in the adsorption energies between basis sets. The effects of dispersion corrections are also presented.

Basis Set Effect	With Dispersion	Without Dispersion	Difference (eV) between With and Without Dispersion	$E_{\text{TZ}} - E_{\text{DZSR}}$ (eV) with Dispersion	$E_{\text{TZ}} - E_{\text{DZSR}}$ (eV) without Dispersion
ZnO Slab+CO (DZVP-SR)	-0.863	-0.688	-0.175		
ZnO Slab+CO (TZV2P)	-0.676	-0.504	-0.172	0.188	0.184
ZnO Slab+CO ₂ (DZVP-SR)	-0.434	-0.200	-0.233		
ZnO Slab+CO ₂ (TZV2P)	-0.404	-0.175	-0.229	0.030	0.025
ZnO Slab+H ₂ (DZVP-SR)	-0.303	-0.180	-0.123		
ZnO Slab+H ₂ (TZV2P)	-0.209	-0.083	-0.126	0.094	0.097
ZnO Slab+H ₂ O (DZVP-SR)	-1.210	-0.989	-0.221		
ZnO Slab+H ₂ O (TZV2P)	-1.180	-0.961	-0.219	0.030	0.028

In Table C.10, we show the effect of dispersion interactions on the calculated adsorption energies. Here, dispersion interactions are included in our calculations through Grimme’s D3-dispersion correction, as implemented in the CP2K package [224]. We did not study the effects of dispersion with all of the software packages as it is evident from the CP2K results that dispersion brings only minor changes to the difference in adsorption energies with a change in the basis set; i.e., $E_{\text{ads}}(\text{TZ})-E_{\text{ads}}(\text{DZ})$ is approximately zero. Although the inclusion of dispersion interactions changed the adsorption energy of each molecule (by as high as ~ 0.2 eV for CO_2 and H_2O), it did not affect the difference in the adsorption energy when the basis set was changed (last two columns of Table C.10).

C.2.3 Planewave Cutoff Effect on Adsorption Energy (eV)

Table C.11: Adsorption energies of CO and CO_2 molecules on a ZnO slab with various plane-wave cutoff energies. The plane-wave cutoff is converged to 1 meV at the 1000 Ry cutoff.

System	Adsorption Energy at 1000 Ry Cutoff, E_{1000} (eV)	Adsorption Energy at 1500 Ry Cutoff, E_{1500} (eV)	Difference in Adsorption Energy ($E_{1500} - E_{1000}$), (eV)
ZnO Slab + CO (TZV2P)	-0.676	-0.675	0.001
ZnO Slab + CO_2 (TZV2P)	-0.404	-0.403	0.001

C.2.4 Structural Relaxation Effect on Basis-Set Induced Adsorption Energy Changes

In this section, we highlight the major differences in the optimized geometry that we obtained for different basis sets. More importantly, we also show that the basis-set induced adsorption energy differences are primarily dictated by the substrate-adsorbate orientation and the electron density distribution rather than by the individual fragment’s

structural relaxation.

In Figure C.2, we have given the optimized geometries of the CO/CO₂ molecule on a 4x4 ZnO slab, optimized at both the PBE/DZ and PBE/TZ level of theories. In the same figure, we have also highlighted the important bond lengths. Clearly, even for the case of the CO molecule (which we find the adsorption energies to be quite different for different basis sets), the structural differences with different basis sets were minor. The major structural difference that we observed was only in the Zn-C distance, which only differed by only 0.05 Å. Specifically, we find the Zn-C distance to be 2.09 Å with the DZ basis and 2.14 Å for the TZ basis. Clearly, these distances reflect the stronger (weaker) CO-ZnO interaction that we obtain for calculations performed with the DZ (TZ) basis. Similarly, the structural differences were found to be minor even for the CO₂ molecule.

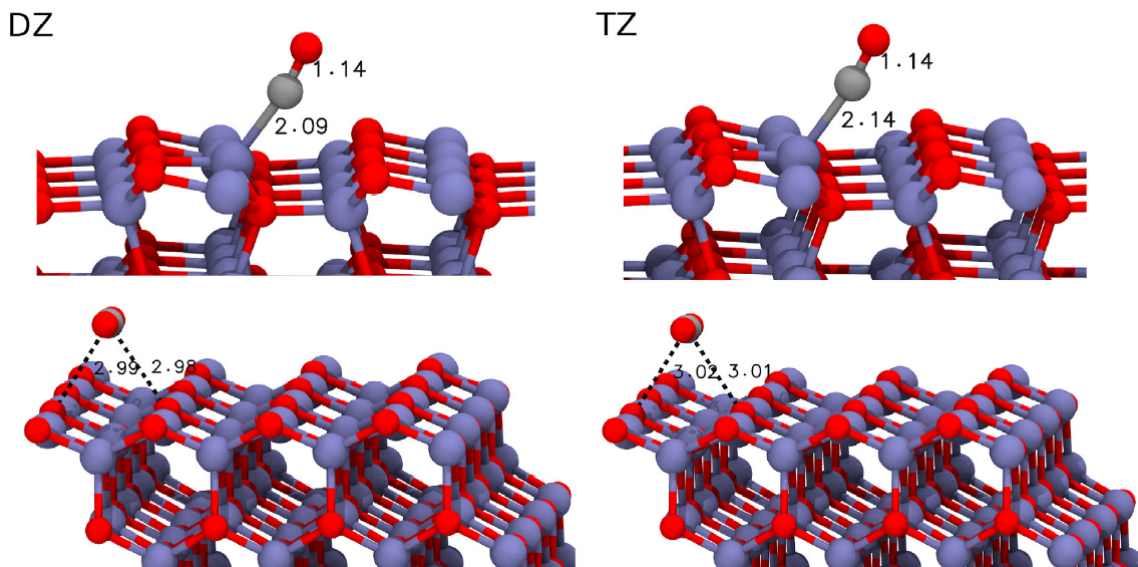


Figure C.2: Differences in the optimized geometries of CO/CO₂ on a 4x4 ZnO slab with different basis sets. The differences are minor; however, it can be noticed that the optimized structures at the DZ basis have shorter adsorbate-surface distances compared to the optimized structures at the TZ basis, reflecting the stronger (weaker) adsorption energies with the DZ (TZ) basis.

Next, we present the basis-set induced adsorption energy changes for a fixed geometry. As shown in Table C.12, the remarkable difference in the adsorption energies of a CO molecule induced by a basis set change was found to be persistent even with the fixed-geometry calculations. For example, when the geometries of the ZnO slab and the CO molecule are fixed at their combined geometry (optimized at the PBE/TZV2P level), we find a difference of 0.177 eV between the adsorption energies calculated with the DZVP-SR and TZV2P basis sets. This energy difference (0.177 eV) is very close to the energy difference (0.184 eV) that we obtained by allowing both the ZnO slab and the CO molecule to completely relax (see Table C.9). Here, we note that we are comparing the difference in the adsorption energies but not the absolute values of the adsorption energies (which will obviously change as we are comparing between the relaxed vs. un-relaxed structures). Similarly, for the case of CO₂, the calculated adsorption energy differences for different basis sets are 0.024 and 0.025 eV for the single-point and completely optimized structures, respectively. These results clearly show that the basis-set induced adsorption energy differences can be understood even without relaxing the individual fragments (as long as we keep the substrate-adsorbate orientation intact). We have used this fact for calculating the electron density difference maps and performing the Natural Bond Order (NBO) analysis (given in the next section).

Table C.12: Differences in the adsorption energies for CO and CO₂ molecules with different basis sets for a fixed nuclear geometry. Here, we used the geometries of a single CO/CO₂ on a 4x4 ZnO slab optimized at the PBE/TZV2P level of theory. Note that the adsorption energy differences in the last column are very close to the adsorption energy differences reported for the completely optimized structures, as shown in Table C.9.

System	Adsorption Energy with DZVP-SR Basis, E_{DZ} (eV)	Adsorption Energy with TZV2P Basis, E_{TZ} (eV)	Difference in Adsorption Energies ($E_{\text{TZ}} - E_{\text{DZ}}$), (eV)
ZnO Slab + CO	-0.856	-0.679	0.177
ZnO Slab + CO ₂	-0.210	-0.186	0.024

C.3 NBO Analysis

To further understand the changes in the CO-ZnO interaction with different basis sets, we have performed a Natural Bond Orbital (NBO) analysis on the interaction of the CO molecule with a Zn₉O₉ cluster (the geometry of this CO + Zn₉O₉ cluster (see Figure C.3) was obtained by cutting the PBE/TZV2P optimized ZnO slab + CO geometry) using the Gaussian 16 software package[514]. Here, we have used the PBE exchange-correlation functional and considered the 6-31g(d) and 6-311g(d,p) basis sets for comparison. In these calculations, as we aim to understand basis set effects, we have not optimized the geometries. From the NBO analysis (see Table C.13), we find that the bond order of the Zn-C bond decreases from the 6-31g(d) basis to the 6-311g(d,p) basis. This decrement in the bond order is a clear indication of the decrease in the bonding electrons between the Zn-C bond when using the triple-zeta basis. Similarly, the electron-density difference maps obtained with the CP2K package suggests a similar behavior. Furthermore, from the natural charge analysis (see Table C.13), we find a greater charge transfer (CT) from the CO-bound Zn atom to the C atom (of the CO) as we change from the DZ basis to the TZ basis. This

increment in the CT from the Zn atom to the C atom leads to a decrement in the Zn-C bond’s covalent character and a lower adsorption energy when employing the TZ basis.

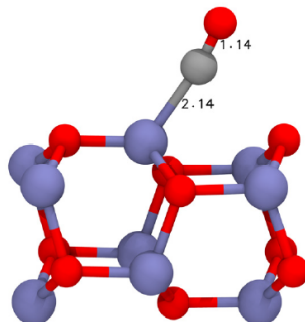


Figure C.3: The geometric configuration of the CO-Zn₉O₉ cluster considered for the NBO analysis. Important bond lengths are shown.

Table C.13: Calculated Wiberg bond indices for the Zn-C bond and natural charges for the Zn and C atoms involved in the bonding with the DZ and TZ basis sets.

System	Basis Set	Natural Charge on Zn Atom (Bound to CO)	Natural Charge on C Atom (of CO)	Wiberg Bond Index (Zn-C Bond)
Zn ₉ O ₉ + CO	6-31g(d) = DZ	0.989	0.439	0.473
	6-311g(d,p) = TZ	1.562	0.234	0.230

C.4 Mulliken Charge Analysis of the 6 eV Collision Trajectories

We have performed a Mulliken charge analysis for the TZ and DZ trajectories of the 6 eV collision in the vicinity of the 4th CO. These analyses are in line with the energy changes presented in the main text. During the first picosecond of the simulation, there is a 0.4 e⁻ charge transfer from the Zn atom (binding to the 4th CO molecule) to the 4th CO molecule. While this additional charge has been transferred back from the CO to the other

Zn atoms of the substrate in the DZ basis set simulation (as it is close to the substrate), during the TZ basis set simulation, this charge was retained on the 4th CO molecule (as the 4th CO is desorbing from the surface). Also, as expected, all the atoms in the vicinity of the 4th CO–Zn bond and in the vicinity of the H₂O impact area had appreciable changes in their charges (see Figure C.4). However, we could not find any appreciable differences in the charges between the TZ and DZ basis set. One notable difference between the TZ and DZ basis simulations is that the transferred charge was localized on the oxygen atom of the CO molecule in the DZ case (see Figure C.5), whereas it is localized on the carbon atom for the TZ basis. We have also verified this result by optimizing a CO molecule with an additional electron with both the DZ and TZ basis. This additional charge on the carbon atom could be a reason for the weak bonding between the Zn and C atom of CO, resulting in the facile dissociation of the CO molecule in the simulations with the TZ basis.

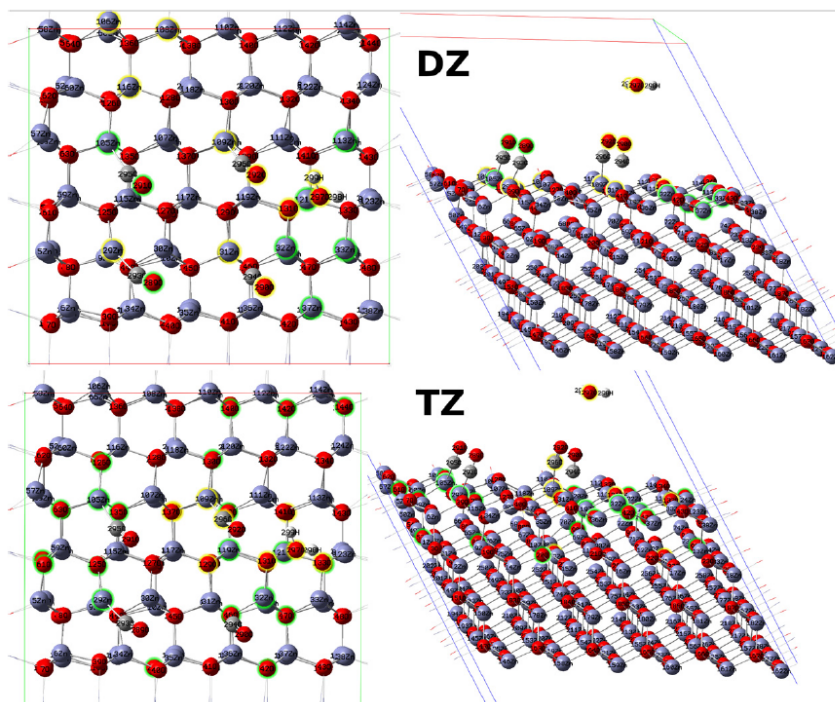


Figure C.4: Top- and side-view of the 6 eV collision structure along with the numbering scheme used. The atoms with a maximum change in the Mulliken charge during the NVE simulation were highlighted with green ($> 0.2 e^-$ change) and yellow ($> 0.25 e^-$ change). The top and bottom panels correspond to the DZ and TZ basis set, respectively.

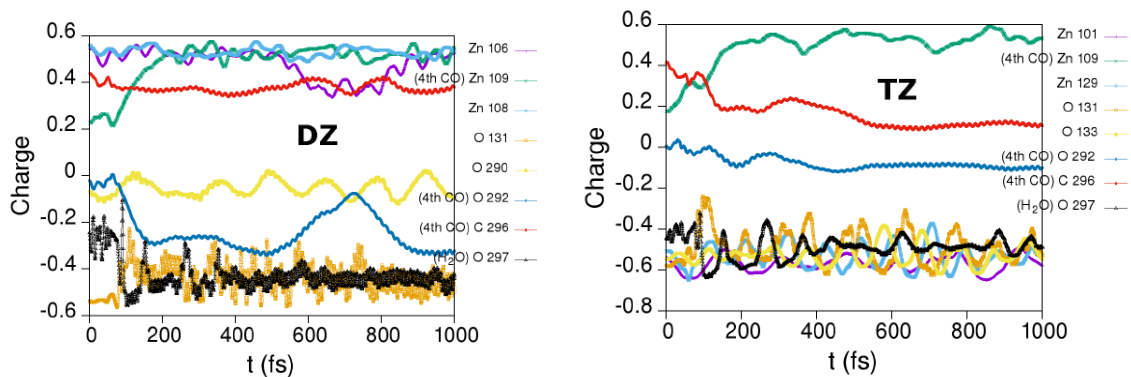


Figure C.5: Changes in the Mulliken charges for a few important atoms during the simulation (6 eV collision). The left and right panels show the changes in the Mulliken charges for the simulations performed with the DZ and TZ basis, respectively. The numbering scheme of the atoms is given in Figure C.4.

C.5 Variation in the Adsorption Energy Across Software Packages

Table C.14 shows the adsorption energies for all the four molecules calculated using three different software packages with sufficiently high-quality basis sets (TZV2P for CP2K, Tier-2 for FHI-aims, and a plane-wave basis with a 500 eV cutoff for VASP. 500 eV is 1.5 times greater than the suggested cutoff of all the elements involved in this work). Since all of these software packages do not use the same form for the dispersion corrections, we did not include these corrections for the results presented in Table C.14. However, as noted earlier, we have studied the effect of dispersion corrections for the case of CP2K. From Table C.14, it is clear that, although all the software packages use the same exchange-correlation functional, the predicted adsorption energies differ by as much as 0.2 eV (see the last three columns of Table C.14), especially for the case of the CO molecule. It is important to mention here that we are not comparing the absolute energies but rather the adsorption energies predicted by the same functional across various software packages *with high-quality basis sets; moreover*, 200 meV is not a small number, considering the fact that room temperature corresponds to a thermal energy of ~ 26 meV.

Table C.14: Adsorption energies of different molecules on a ZnO slab as predicted by the three different software packages. In the last 3 columns, the difference in the adsorption energies predicted by any two software packages is given for comparison.

System	E_{CP2K} (TZV2P)	$E_{FHI-aims}$ (Tier-2)	E_{VASP} (500 eV cutoff)	$E_{CP2K}-E_{VASP}$	$E_{CP2K}-E_{FHIaims}$	$E_{FHI-aims}-E_{VASP}$
ZnO Slab+CO	-0.504	-0.398	-0.300	-0.204	-0.106	-0.098
ZnO Slab+CO ₂	-0.175	-0.172	-0.151	-0.024	-0.003	-0.021
ZnO Slab+H ₂	-0.083	-0.096	0.009	-0.093	0.013	-0.105
ZnO Slab+H ₂ O	-0.961	-0.999	-0.856	-0.105	0.038	-0.143

C.6 Effect of Exchange-Correlation Functionals

To confirm that the effect of basis sets on the BOMD predicted reaction mechanisms (reported in the main text) are not just limited to the PBE-D3 functional (i.e., the PBE functional with Grimme’s dispersion correction) we have also performed additional BOMD simulations with two more functionals, namely, LDA and PBE (without dispersion correction). For these NVE simulations, we have used the same initial conditions (positions and velocities) that we considered earlier for the PBE-D3 calculations. Also, we have only considered the 6 eV collision in the new BOMD calculations since the differences in the results obtained with the DZ and TZ basis were apparent at this collision energy while using the PBE-D3 functional. These calculations were carried out with long simulation times or until the molecules were either desorbed/reflected from the simulation cell along the *c*-direction.

C.6.1 BOMD Simulation Results with PBE Functional (without Grimme’s Dispersion)

The results obtained with the PBE functional for both the DZ and TZ basis are shown in Figure C.6. Once again, we find that the 4th CO molecule desorbs with the TZ basis, whereas it diffuses with the DZ basis. Also, for the high energy collision, all of the CO molecules evolved similarly with both the PBE and PBE-D3 functionals. However, with the PBE functional we find that the 4th CO desorbed from the surface at a quicker rate (< 1 ps) compared to the PBE-D3 functional (~ 3 ps), which is due to the lower binding energy of CO to the ZnO surface in the absence of any dispersion interaction (see Table

C.15). Also, while employing the DZ basis, we observed that the H_2O molecule scatters from the surface with the PBE functional.

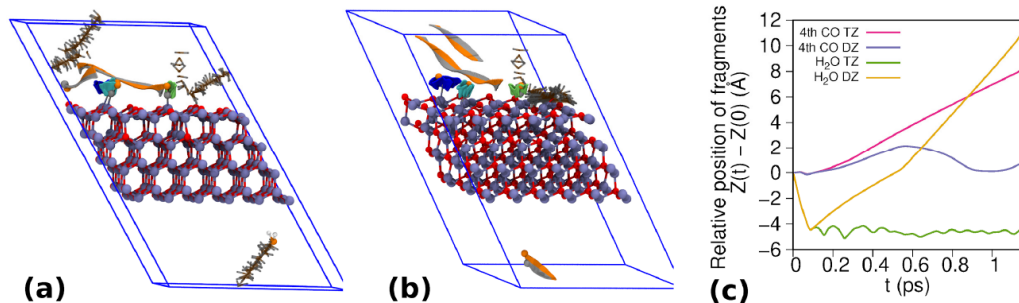


Figure C.6: Panels (a) and (b) show how the CO and H_2O molecules evolve during the NVE simulation with the DZVP and TZV2P basis sets, respectively, for the 6 eV collision. In both panels, the cyan, lime, and blue trajectories show the change in the positions of the 1st, 2nd, and 3rd CO molecules, respectively, for the entire trajectory (the simulation was stopped after the desorption of any molecule from the ZnO surface). The changes in the 4th CO position are shown with the stick model without changing its colors, and the changes in the water molecule are shown with the glass ball-and-stick model. Once again, due to its proximity to the impact region, the 4th CO molecule exhibited significant changes with both basis sets. A difference in the behavior of the 4th CO with the DZ and TZ basis sets is apparent (results are discussed in the main text). The trajectories of the molecules appear discontinuous since they crossed the periodic boundaries at the points of visual discontinuity. In panel (a), the diffusion of the 4th CO and scattering of H_2O were observed (DZ basis); in panel (b), the desorption of the 4th CO and the dissociation of H_2O were observed (TZ basis); (c) Relative change in the positions of the 4th CO and H_2O molecules along the c-direction of the lattice.

Table C.15: Adsorption energy of the CO molecule calculated using CP2K with different basis sets and exchange-correlation functionals.

System	Adsorption Energy	Adsorption Energy	Adsorption Energy
	with DZVP-SR Basis, $E_{\text{DZVP-SR}}$ (eV)	with DZVP Basis, E_{DZVP} (eV)	with TZV2P Basis, E_{TZV2P} (eV)
ZnO Slab + CO (PBE)	-0.688	-0.884	-0.504
ZnO Slab + CO (PBE-D3)	-0.863	-1.060	-0.676
ZnO Slab + CO (LDA)	-1.128	-1.310	-0.890

C.6.2 BOMD Simulation Results with LDA Functional

Next, we present our results with the LDA functional in Figure C.7. Similar to the PBE functional, the results obtained with the TZ and DZ basis are quite different, even with the LDA functional. However, unlike the PBE functional, here the changes primarily occur with the colliding water molecule, which dissociates only with the TZ basis and not with the DZ basis (see Figure C.7). Furthermore, we note that for both the TZ and DZ basis, although we observed a hopping mechanism for the 4th CO molecule from one Zn site to another (much like our 0.6 eV collision result with the PBE-D3 functional), we did not observe any desorption with either of the basis sets. To understand this behavior, we calculated the adsorption energy of CO on the ZnO surface using the LDA functional with various basis sets (see Table C.15) and compared these results with those obtained with the PBE and PBE-D3 functionals.

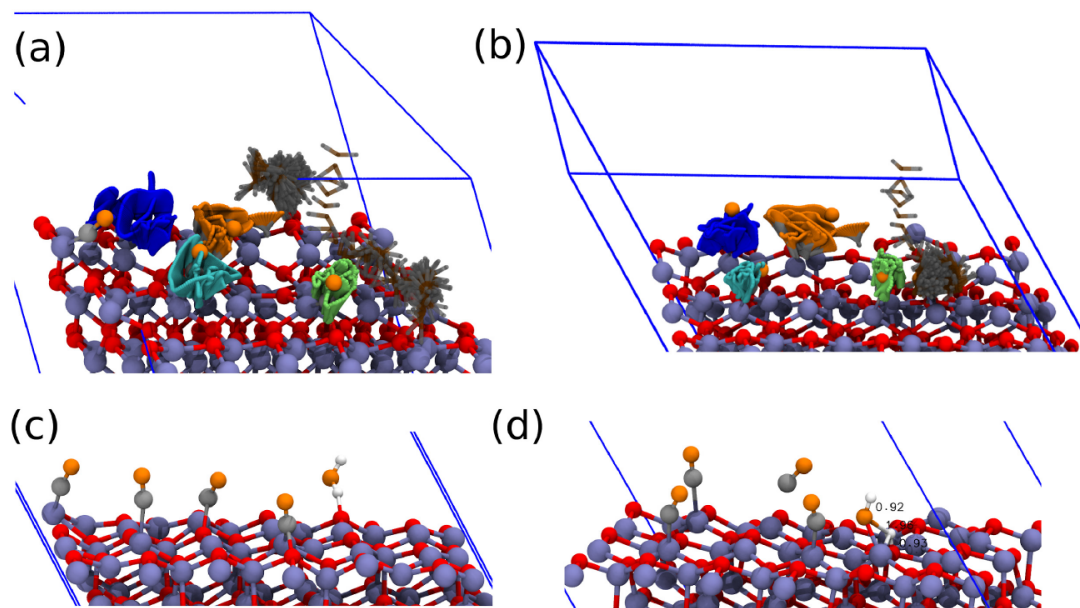


Figure C.7: Panels (a) and (b) show how the CO and H₂O molecules evolve during the NVE simulation with the DZVP and TZV2P basis sets, respectively, for the 6 eV collision using LDA. Panels (c) and (d) show the final frames of the simulations. The dissociation of the H₂O molecule occurs with the TZ basis while diffusion of the H₂O occurs when the DZ basis is used.

From the adsorption energy values calculated with various functionals, it is quite apparent that the LDA functional overbinds the CO molecule compared to both the PBE and PBE-D3 functionals. Specifically, the LDA functional overbinds the CO molecule by ~ 0.25 eV and ~ 0.4 eV compared to the PBE-D3 and plain PBE functional, respectively (see Table C.15). This additional binding energy prohibits the desorption of CO during the 6 eV collision (while using LDA).

As mentioned earlier, this finding is very similar to the case of the 0.6 eV collision while using the PBE-D3 functional, where we calculated the kinetic energy gained by the 4th CO molecule to be on the order of 0.1-0.4 eV during the entire trajectory (see Figure C.8). As long as the CO molecule is close to the surface, this kinetic energy gain is not sufficient

for the CO molecule to escape from the surface, even with the TZV2P basis (${}^{\text{TZV2P}}E_{\text{ads}} = -0.676$ eV). The situation is almost the same with LDA, where the kinetic energy gained by the 4th CO molecule, after a 6 eV collision, is only on the order of 0.1-0.7 eV for both the DZ and TZ basis throughout the trajectory (except for a tiny interval at the beginning of the simulation). Once again, this kinetic energy gain is not sufficient for the 4th CO molecule to escape from the ZnO surface even with the TZV2P basis (${}^{\text{TZV2P}}E_{\text{ads}} = -0.890$ eV), and thus, it is bound to the surface during the simulations with LDA, regardless of the employed basis set (see Figure C.8). Contrary to both of the above cases, the kinetic energy gained by the 4th CO molecule is on the order of 0.5-0.7 eV during the entire trajectory when the PBE functional is used with the 6 eV collision. Here, this kinetic energy gain is sufficient for the 4th CO molecule to escape from the ZnO surface with the TZV2P basis (${}^{\text{TZV2P}}E_{\text{ads}} = -0.5$ eV). Thus, we observed a desorption with the TZV2P basis while employing the PBE functional (see Figure C.8).

Based on the above argument, it is clear that the similarities in the CO dynamics with different basis sets (while employing LDA) are only due to the insufficient collision energy considered (6 eV). Once higher collision energies are used, we observe a clear difference in the CO desorption behavior with different basis sets. This conclusion is based on the observation that even with LDA, we find an immense difference in the adsorption energy of the CO molecule (by ~ 0.4 eV) between the DZ and TZ basis (see Table C.15). Due to the substantial computational costs involved with the BOMD calculations, we have not verified this conclusion through additional simulations.

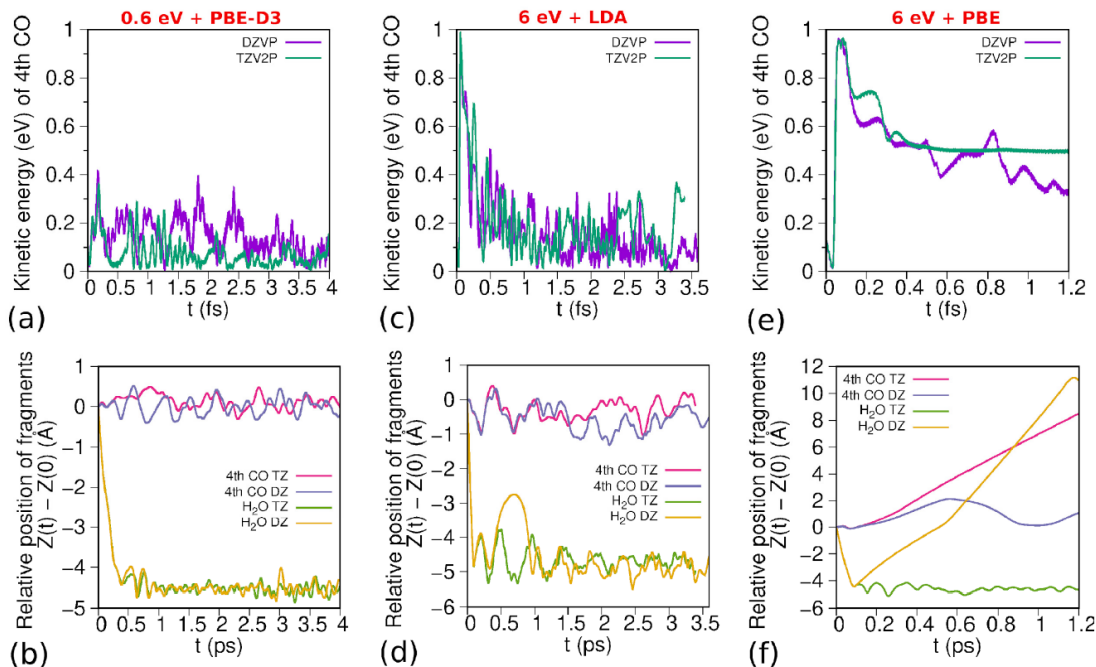


Figure C.8: Top panels (a), (c), and (e) show the changes in the KE of the 4th CO using both the DZ and TZ basis during the entire simulation time. Bottom panels (b), (d), and (f) show the relative change in the positions of the 4th CO and H₂O molecules along the c-direction of the lattice during the entire simulation time. From **left to right**, both the top and bottom panels correspond to simulations with the PBE-D3 (**0.6 eV** collision), LDA (**6 eV** collision), and PBE (**6 eV** collision) functionals, respectively.

Finally, we have also performed additional single-point adsorption energy calculations with the new strongly constrained and appropriately normed (SCAN) functional (with the optimized structure at the PBE level of theory) for the CO molecule with both the Tier-1 and Tier-2 basis sets using the FHI-aims package (see Table C.16). Once again, we found that there is a clear difference (~ 0.04 eV) in the adsorption energy with different basis sets. Similar results were also obtained with the hybrid B3LYP and M06-2X functionals (see Table C.17).

Table C.16: Adsorption energy of the CO molecule calculated using FHI-aims with different basis sets and exchange-correlation functionals.

System	Adsorption Energy with the Tier-1 basis, E_{Tier1} (eV)	Adsorption Energy with the Tier-2 basis, E_{Tier2} (eV)	Difference in Adsorption Energy, ($E_{\text{Tier2}} - E_{\text{Tier1}}$), (eV)
ZnO Slab + CO (SCAN)	-0.797	-0.760	0.037
ZnO Slab + CO (PBE)	-0.489	-0.398	0.091

Table C.17: Adsorption energy of the CO molecule on a Zn_9O_9 cluster calculated using the Gaussian16 software package with different basis sets and exchange-correlation functionals. All results were corrected for the basis set superposition error (BSSE) using the counterpoise method as implemented in Gaussian16.

Gaussian 16 Calculations	BSSE-Corrected Adsorption Energy, B3LYP E_{ads} (eV)	BSSE-Corrected Adsorption Energy, M06-2X E_{ads} (eV)	BSSE-Corrected Adsorption Energy, PBE E_{ads} (eV)
E_{DZ} basis (6-31g(d))	-0.368	-0.616	-0.536
E_{TZ} basis (6-311g(d,p))	-0.256	-0.416	-0.418
$E_{\text{DZ}} - E_{\text{TZ}}$ (eV)	-0.112	-0.200	-0.118

Based on all of these results, we believe that our main conclusion – the reaction mechanism predicted using BOMD simulations depends significantly on the employed basis set – is valid even with different exchange-correlation functionals. Also, we find that the simple diagnostic test of calculating the adsorption energies with various molecules involved in the reaction (to estimate how strongly the predicted reaction mechanism will be affected by a change in the employed basis set quality) still holds for various exchange-correlation functionals and basis sets.

C.7 BOMD Simulations with Changes in the Impact Position

We also performed a few additional calculations by impinging an H_2O molecule towards the ZnO surface such that (1) it hits the substrate in between the 4th and 2nd CO molecules, and (2) it interacts with the 2nd CO molecule during its flight towards the substrate. These calculations were performed using both the DZ and TZ basis sets, and, for each basis, both the presence and absence of the dispersion interactions were also examined. From the first set of calculations (see Figure C.9), we find that there is neither diffusion nor desorption of any of the CO molecules from the surface regardless of the inclusion/neglect of dispersion interactions and regardless of the basis set used. Thus, the mere collision of an H_2O molecule with the ZnO surface did not lead to the dissociation of any of the adsorbed CO molecules from the surface. For the second case (see Figure C.10), however, we find that the 2nd CO molecule desorbs from the surface when using a TZ basis, and it diffuses on the surface when a DZ basis is used – once again regardless of the inclusion/neglect of dispersion interactions. This second result reinforces our original result with the 4th CO molecule. Together, from these results, *it is quite evident that if an incoming H_2O molecule interacts with the adsorbed CO molecule during its flight, usage of a TZ basis will lead to the desorption of the CO molecule, whereas usage of a DZ basis will lead to the diffusion of the CO molecule on a ZnO surface instead.*

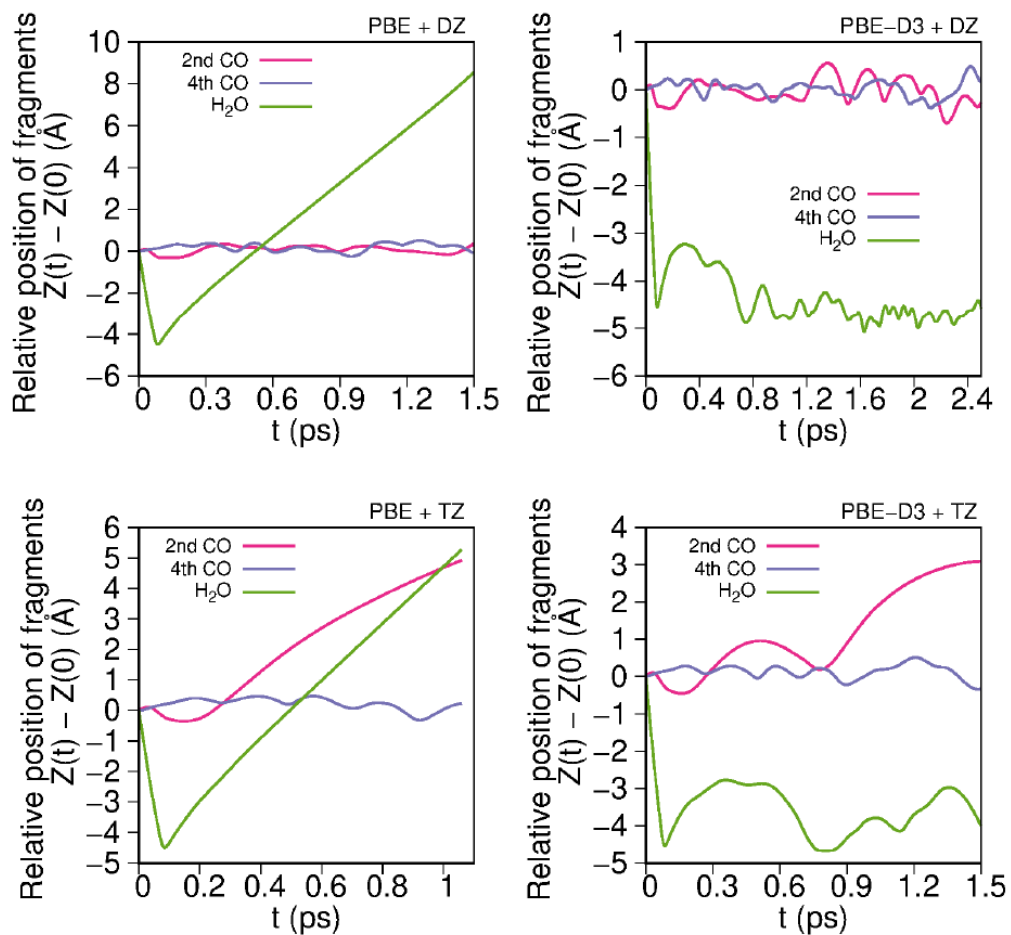


Figure C.9: Relative change in the positions of the 2nd CO, 4th CO, and H₂O molecules along the c-direction of the lattice. The results are presented for NVE simulations performed with the PBE (no dispersion correction) and PBE+D3 (dispersion-corrected) functionals and with the DZ and TZ basis sets. Here, the impinging H₂O molecule collides with the ZnO substrate in the vicinity of the 2nd CO molecule.

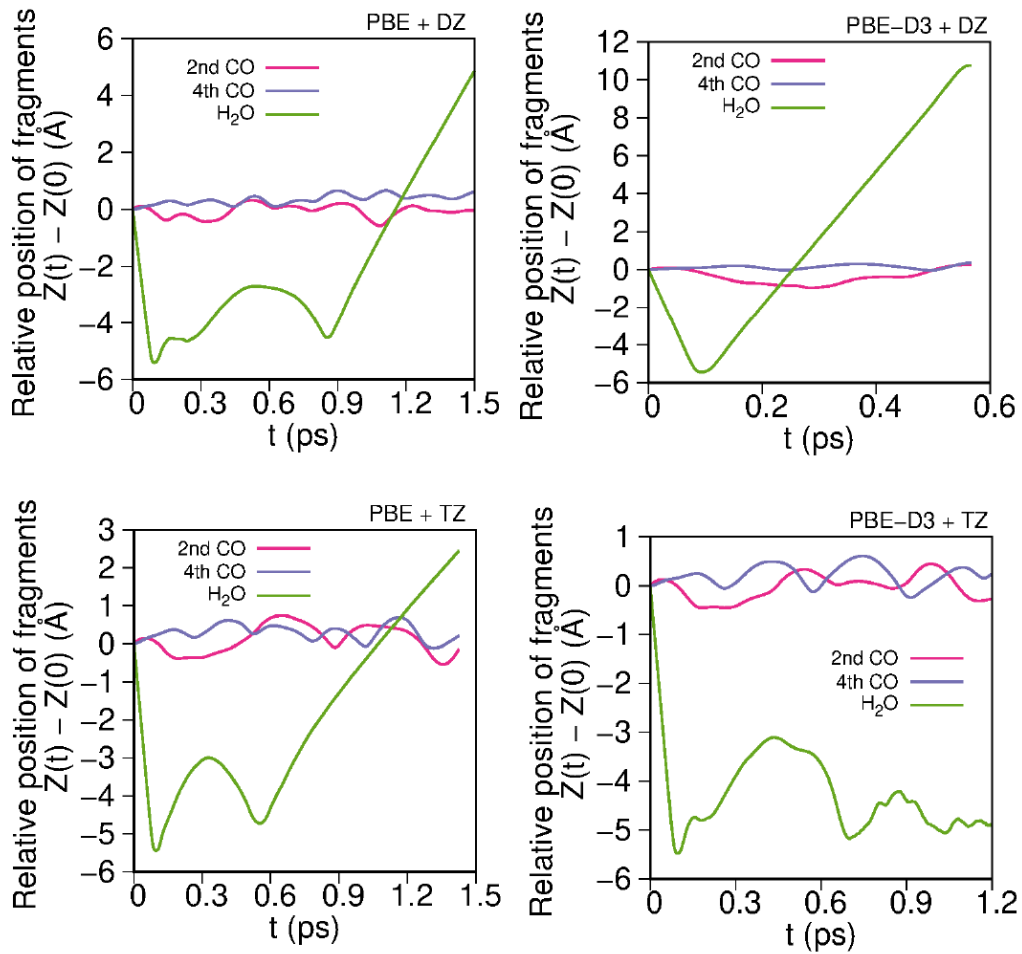


Figure C.10: Relative change in the positions of the 2nd CO, 4th CO, and H₂O molecules along the c-direction of the lattice. The results are presented for NVE simulations performed with the PBE (no dispersion correction) and PBE+D3 (dispersion-corrected) functionals and with the DZ and TZ basis sets. Here, the impinging H₂O molecule collides with the ZnO substrate between the 2nd and 4th CO molecules.

Appendix D

Charge Density Wave Hampers

Exciton Condensation in 1*T*-TiSe₂

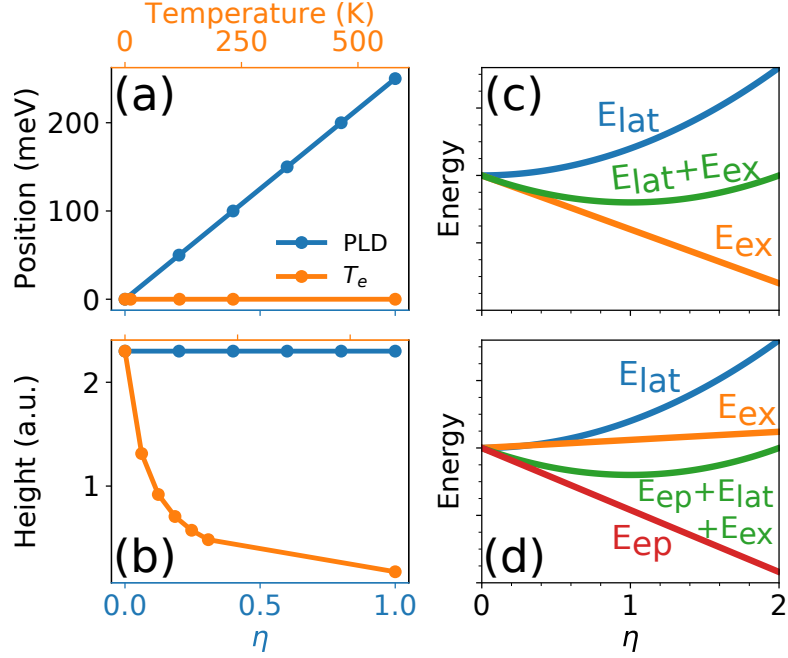


Figure D.1: The peak position $\epsilon(q = w)$ [Eq. 7.17] (a) and the peak height at $\epsilon(q = w)$ (b) of the soft mode as a combined effect of the PLD and T_e . (c)(d) The lattice energy E_{lat} , exciton energy E_{ex} , electron-phonon energy E_{ph} , and total energy E_{tot} as a function of the PLD, indicated by (c) the analytical model and (d) our DFT and experimental framework.

Combining our analysis of the effects of the PLD and T_e , we have developed a microscopic framework to explain the temperature dependence of the plasmon dispersion. As shown in Fig. D.1(a) and (b), the PLD raises the soft mode without smearing the plasmon peaks; in contrast, the electronic temperature smears the peaks without changing the dispersion of the soft mode. Based on this framework, the experimental observations can be explained by recognizing that at a high temperature of $T = 300$ K, TiSe_2 is semimetallic without a PLD. However, the thermal carriers occupy the E1 band and completely smear the soft mode signals; therefore, only the regular metallic plasmon can be observed in the experimental EELS. At a low temperature of $T = 17$ K, the thermal excitation is suppressed, while the energy required to create an exciton significantly increases due to the emerging

CDW gap. When the transition temperature is near $T = 185$ K, the PLD and carrier density are sufficiently small, resulting in the observed soft modes.

We directly compared calculated and experimental plasmon dispersions $\epsilon(T)$ at temperature T to verify our framework further. The effect of temperature is simulated with $\{T_e = T, \eta = \eta(T)\}$, where $\eta = \eta(T)$ is the PLD as a function of temperature, reproduced from [267] and shown in Fig. 7.1(j). We compare the experimental plasmon dispersion at $T = 17$ K, $T = 100$ K, and $T = 185$ K with our calculations with the parameters $\{T_e = 0$ K, $\eta = 1.0\}$, $\{T_e = 117$ K, $\eta = 0.8\}$, and $\{T_e = 232$ K, $\eta = 0.0\}$, respectively. As shown in Fig. 7.1(i), despite the difference in the absolute energy, the plasmon dispersions in our simulations are consistent with the experimental measurements for all three cases: at $T = 17$ K, the CDW gap raises the interband plasmon mode to the same energy range of the regular plasmon, resulting in a momentum-independent plasmon dispersion; at $T = 100$ K, the CDW gap decreases together with the PLD, presenting a soft mode with a finite energy gap at $q = w$; at $T_c = 185$ K, the soft mode is observed in both the experimental and simulated EELS as a result of the missing CDW gap and relatively small thermal excitation.

The complex effects of the thermal field obtained from our *ab initio* calculations provide a new mechanistic explanation that is quite distinct from the results of the analytical model commonly used in the scientific literature. In the analytical model, an additional BCS-like electron-hole coupling term [328] is added to the single particle tight-binding Hamiltonian [262–264, 268]. The exciton effect is overestimated because the BCS-like term is parameter-dependent, and the electron-phonon coupling is absent in the Hamiltonian.

The single-particle band structures are fixed to those at the normal state, and the CDW gap is omitted when the temperature decreases to T_c . In Ref. [268], the Hamiltonian without exciton interaction is

$$H_0 = \sum_{\mathbf{k}} \epsilon_v(\mathbf{k}) a^\dagger(\mathbf{k}) a(\mathbf{k}) + \sum_{i, \mathbf{k}} \epsilon_c^i(\mathbf{k}) b_i^\dagger(\mathbf{k}) b_i(\mathbf{k}) \quad (\text{D.1})$$

Here a^\dagger and b^\dagger are operators creating electrons with wave vector \mathbf{k} in the valence band and in the conduction band labeled i , respectively. To describe the exciton dissipation by the CDW, we propose a simple modification to the tight-binding model by adding the temperature-dependent CDW gap

$$\epsilon_{\text{CDW}}(T) = \begin{cases} 0, & \text{if } T \geq T_c \\ (1 - T/T_c)E_g, & \text{if } T < T_c \end{cases} \quad (\text{D.2})$$

to Eq. D.1 as

$$H_0 = \sum_{\mathbf{k}} \epsilon_v(\mathbf{k}) a^\dagger(\mathbf{k}) a(\mathbf{k}) + \sum_{i, \mathbf{k}} \left[\epsilon_c^i(\mathbf{k}) b_i^\dagger(\mathbf{k}) b_i(\mathbf{k}) + \epsilon_{\text{CDW}}(T) \right] \quad (\text{D.3})$$

where $E_g = 0.27$ eV is the CDW gap at $T = 0$ and $T_c = 180$ K is the transition temperature. As shown in Fig. D.1(c), the analytical model indicates that the CDW state is a compromise between the increasing lattice energy and the decreasing electron energy via the formation of excitons [328]. This suggests that the exciton condensation distorts the lattice and lowers the system's energy. Thus, the soft mode should not be destroyed by the PLD, provided that $T < T_c$. Instead, both the experimental spectra [259] and our EELS simulation show a different phenomenon: the PLD hinders the exciton condensation, which indicates another mechanism such as EPC is essential for stabilizing the CDW state [Fig. D.1(d)].

We note that the mechanism presented for the low-temperature suppression of condensation does not depend on the existence of a CDW phase. The isolated electron system may form an exciton condensation if the Coulomb binding is larger than the bandgap [262–264]. However, the interactions between the electron and phonon in this system introduce a channel for energy redistribution; specifically, the ionic energy can be increased by forming the PLD, and the electronic energy can be decreased by opening a CDW gap, which lowers the total energy. Additional mechanisms may also raise the energy of the soft plasmon mode in materials where no such CDW phase exists. We expect that interactions between electrons and different systems (other than phonons) may also introduce a channel for energy redistribution. For example, the interaction between the electron and spin may introduce a periodic spin order, i.e., a spin density wave (SDW) [515]. This will also open a bandgap and suppress the exciton condensation. An SDW state instead of the exciton condensation state emerges under the transition temperature.

Appendix E

Implementation of Real-Time TDDFT for Periodic Systems in the Open-Source PySCF Software Package

E.1 Nuclear Dynamics

To formulate mixed-quantum-classical nuclear dynamics in the Ehrenfest formulation, we follow the formulation in Ref. [516]. We first consider the time-derivative of the total energy and designate it as the work provided by the nuclear forces. The analytic time derivative of the total force becomes

$$\begin{aligned} \frac{dE_\ell}{dt} &= \langle \dot{\Psi} | \frac{-\nabla^2}{2} | \Psi \rangle + \langle \Psi | \frac{-\nabla^2}{2} | \dot{\Psi} \rangle + \int \dot{n}(\mathbf{r}) \frac{\delta(E_{xc} + E_{ext})}{\delta n(\mathbf{r})} + \dot{\mathbf{R}} \cdot \frac{\partial E}{\partial \mathbf{R}} \Big|_{D,n} \\ &= \sum_{\nu\mu} H_{\mu\nu}^{KS} \dot{D}_{\nu\mu} + \dot{\mathbf{R}} \cdot \frac{\partial E}{\partial \mathbf{R}} \Big|_{D,n}, \end{aligned} \quad (\text{E.1})$$

where $\frac{\partial E}{\partial \mathbf{R}} \Big|_{D,n}$ represents the nuclear coordinate partial derivative with the density matrix (D) and density (n) held fixed. We thus divided the time derivative into one part arising from the time evolution of the density matrix D and the other part that is independent from density matrix. The first part can be further decomposed as

$$\begin{aligned} \sum_{\mu\nu} H_{\mu\nu}^{KS} \dot{D}_{\nu\mu} &= \sum_{\nu\mu} F_{\mu\nu} \frac{1}{i\hbar} \left((S^{-1})_{\nu\sigma} \left(F - i\hbar \dot{\mathbf{R}} \cdot \mathbf{X} \right)_{\sigma\tau} D_{\tau\mu} - D_{\nu\sigma} \left(F + i\hbar \dot{\mathbf{R}} \cdot \mathbf{Y} \right)_{\sigma\tau} (S^{-1})_{\tau\mu} \right) \\ &= -\dot{\mathbf{R}} \cdot \sum_{\nu\mu} F_{\mu\nu} \left((S^{-1})_{\nu\sigma} \mathbf{X}_{\sigma\tau} D_{\tau\mu} + D_{\nu\sigma} \mathbf{Y}_{\sigma\tau} (S^{-1})_{\tau\mu} \right), \end{aligned} \quad (\text{E.2})$$

whereas the second part in Eq. (E.1) can be decomposed as

$$\begin{aligned} \frac{\partial E}{\partial \mathbf{R}} \Big|_{D,n} &= \sum_{\nu\mu} D_{\nu\mu} \dot{\mathbf{R}} \cdot \frac{\partial}{\partial \mathbf{R}} \langle \mu | \frac{-\nabla^2}{2} + v_{ext} | \nu \rangle \\ &+ \frac{1}{2} \sum_{\nu\mu\nu'\mu'} D_{\nu\mu} D_{\nu'\mu'} \dot{\mathbf{R}} \cdot \frac{\partial}{\partial \mathbf{R}} \langle \mu\mu' | | \nu\nu' \rangle + \sum_{\nu\mu} D_{\nu\mu} \dot{\mathbf{R}} \cdot \frac{\partial}{\partial \mathbf{R}} \langle \mu | \epsilon_{xc} | \nu \rangle. \end{aligned} \quad (\text{E.3})$$

The derivative matrices in Eq. (E.3) for Gaussian AOs are analytically calculable, and PySCF implements derivatives of these matrices in the FFT density-fitting scheme. We can also calculate them by numerical differentiation of the total energy, keeping D and n fixed.

The energy-conserving force associated with the a th nucleus is then given by

$$\mathbf{F}_a = F_{\mu\nu} \left(S^{-1} \mathbf{X}^{(a)} D + D \mathbf{Y}^{(a)} S^{-1} \right)_{\nu\mu} - \left. \frac{\partial E}{\partial \mathbf{R}_a} \right|_{D,n}. \quad (\text{E.4})$$

It should be noted that this expression reduces to the standard expression for the total ground state energy derivative for the ground state density matrix $D_{\mu\nu} = \sum_{\sigma} \sum_{j=1}^{N_{\sigma}} C_{\mathbf{k}\alpha}^{\mu} w_{\mathbf{k}\alpha} C_{\mathbf{k}\alpha}^{\nu*}$ as

$$F_{\mu\nu} \left(S^{-1} \mathbf{X}^{(a)} D + D \mathbf{Y}^{(a)} S^{-1} \right)_{\nu\mu} - \left. \frac{\partial E}{\partial \mathbf{R}_a} \right|_{D,n} = \sum_{\mu\nu} \left(\frac{\partial S_{\mu\nu}}{\partial \mathbf{R}} C_{\lambda}^{\nu} w_{\lambda\varepsilon\lambda} C_{\lambda}^{\mu} \right) - \left. \frac{\partial E}{\partial \mathbf{R}_a} \right|_{D,n} = - \left. \frac{\partial E}{\partial \mathbf{R}} \right|_{D,n}. \quad (\text{E.5})$$

In practical calculations, we calculate $\left. \frac{\partial E}{\partial \mathbf{R}_a} \right|_{D,n}$ by numerical differentiation with fixed D and n .