UNDERSTANDING NON-RADIATIVE RECOMBINATION PROCESSES OF THE OPTOELECTRONIC MATERIALS FROM FIRST PRINCIPLES

By

Yinan Shu

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Chemistry – Doctor of Philosophy

2017

ABSTRACT

UNDERSTANDING NON-RADIATIVE RECOMBINATION PROCESSES OF THE OPTOELECTRONIC MATIRALS FROM FIRST PRINCIPLES

By

Yinan Shu

The annual potential of the solar energy hit on the Earth is several times larger than the total energy consumption in the world. This huge amount of energy source makes it appealing as an alternative to conventional fuels. Due to the problems, for example, global warming, fossil fuel shortage, etc. arising from utilizing the conventional fuels, a tremendous amount of efforts have been applied toward the understanding and developing cost effective optoelectrical devices in the past decades. These efforts have pushed the efficiency of optoelectrical devices, say solar cells, increases from 0% to 46% as reported until 2015. All these facts indicate the significance of the optoelectrical devices not only regarding protecting our planet but also a large potential market. Empirical experience from experiment has played a key role in optimization of optoelectrical devices, however, a deeper understanding of the detailed electron-by-electron, atom-by-atom physical processes when material upon excitation is the key to gain a new sight into the field. It is also useful in developing the next generation of solar materials. Thanks to the advances in computer hardware, new algorithms, and methodologies developed in computational chemistry and physics in the past decades, we are now able to 1).model the real size materials, e.g. nanoparticles, to locate

important geometries on the potential energy surfaces(PESs); 2).investigate excited state dynamics of the cluster models to mimic the real systems; 3).screen large amount of possible candidates to be optimized toward certain properties, so to help in the experiment design. In this thesis, I will discuss the efforts we have been doing during the past several years, especially in terms of understanding the non-radiative decay process of silicon nanoparticles with oxygen defects using ab initio nonadiabatic molecular dynamics as well as the accurate, efficient multireference electronic structure theories we have developed to fulfill our purpose. The new paradigm we have proposed in understanding the nonradiative recombination mechanisms is also applied to other systems, like water splitting catalyst. Besides in gaining a deeper understanding of the mechanism, we applied an evolutionary algorithm to optimize promising candidates towards specific properties, for example, organic light emitting diodes (OLED).

Copyright by
YINAN SHU
2017

This dissertation is dedicated to my parents and grandparents, for their selfless supports, encouragements, and love.

ACKNOWLEDGMENTS

Everything has an end, so does the journey of getting a Ph.D. degree. I have never thought of what I am going to write at the very end of this dissertation, also the very end of my Ph.D. years. Now I find it is too hard to express the feeling into words. My life in these years is completely beyond a training process of being a researcher. It is an unforgettable experience of daily life in a different country, of learning to work and think like a scientist and of making friends from the world. As an international student from China, the very different cultures between the two countries make everything difficult at the beginning. However, I really want to express my most sincere gratitude and appreciation to my Ph.D. advisor Professor Benjamin G. Levine. It is his patience, helpfulness and encouragement throughout these years have made me survival on this hard road of being a researcher. He has taught me not only about those sciences but also how one should behave to be a young scientist.

In addition, I would like to thank Professor Piotr Piecuch, Professor Robert I. Cukier, Professor James E. Jackson, Professor Marcos Dantus for their valuable suggestions, and supports in these years of my graduate studies. Also to the group members of Levine research group, Garrett A. Meek, B. Scott Fales, Weitao Peng, Brandon Z. Child, Michael Esch, and David Bianchi, for all the help they provided, the stimulating discussions with them, and the wonderful research experiences we have shared together.

I would like to thank those of my friends who have accompanied me in these years

no matter I feel happy or sad. Yi Yi, Xiaopeng Yin, Shuai Nie, Chengpeng Chen and Xinliang Ding, as we are all Ph.D. students from China in the Department of Chemistry, who have helped me a lot during the years. Xinyue Zhang, Yunfeng Yang, Ruiyi Liu, Shaozu Wu and Lili Wu with whom I have shared a lot of memorable days in the United States. Also Zuo Yang, Mark Pu, Vincent Chen, Bernard Chen, Zhiling Chen and Jie Fang who have shared unforgettable fun moments of those days when I am in Ann Arbor.

Finally, to my parents, for their support and love in my whole life.

TABLE OF CONTENTS

LIST OF TABLES..... xiv

LIST OF FIGURES..... xxii

Chapter 1 Introduction of The Role of Conical Intersection in Optoelectronic Materials..... 1

1.1 The Role of Conical Intersection In Understanding Non-radiative Decay Process in Optoelectronic Materials..... 2

1.2 SiNC with Oxygen Defects..... 4

1.3 Dissertation Outline..... 7

REFERENCES..... 10

Chapter 2 Introduction of Electronic Structure Theories and Propagation of Dynamics..... 16

2.1 From Exact Theory to Born-Oppenheimer Approximation..... 18

2.2 Conical Intersection Facilitates Non-radiative Decay..... 20

2.3 Proper Description of The Electronic Structures At Conical Intersection Geometries..... 21

2.4 Propagate The Wavefunctions By Solving Time-Dependent Schrödinger Equation..... 25

REFERENCES..... 28

Chapter 3 Reducing The Propensity For Unphysical Wavefunction Symmetry Breaking In Multireference Calculations Of The Excited States..... 32

3.1 Method..... 38

3.1.1 Singly excited active space self-consistent field method..... 38

3.1.2 Accessing wavefunction symmetry breaking..... 42

3.2 Results and Discussion..... 50

3.2.1 Unphysical symmetry breaking in SA-CASSCF and SEAS-CASCI..... 50

3.2.2 Relative accuracy of SEAS-CASCI and SA-CASSCF..... 54

3.2.3 Unphysical wavefunction distortion in asymmetric molecules..... 56

3.2.4 SEAS-CASCI and biradicaloid systems..... 63

3.3 Conclusions..... 66

REFERENCES..... 68

Chapter 4 Configuration Interaction Singles Natural Orbitals: An Orbital Basis for an Efficient and Size Intensive Multireference Description of Electronic

Excited States..... 74
4.1 Method..... 77
4.1.1 CISNO-CASCI method..... 77
4.1.2 Computational details..... 80
4.2 Results and Discussion..... 82
4.2.1 Accuracy of vertical excitation energies..... 82
4.2.2 Comparison of variational efficiency of orbitals..... 91
4.2.3 Description of localized excited states..... 94
4.2.4 Size intensivity and size consistency..... 99
4.2.5 CISNO-CASCI and biradicaloid conical intersections..... 113
4.3 Conclusions..... 115
REFERENCES..... 117

Chapter 5 Recovering The Coupling Between Doubly Excited and Closed Shell Hartree-Fock Determinants On Single Excitation Level: The Class of Virtual-Occupied Coupled Orbital Configuration Interaction Singles (VOCO-CIS) Methods..... 123
5.1 Method..... 128
5.1.1 To go beyond the Brillouin's theorem..... 128
5.1.2 Practical recipe..... 130
5.2 Results and Discussion..... 133
5.2.1 Orbital optimization-CIS approach..... 133
5.2.2 The importance of virtual occupied orbital coupling..... 138
5.3 A New Conceptual Understanding of Iterative Orbital Optimization with Excited Determinants..... 146
5.4 Conclusions..... 147
APPENDIX..... 149
REFERENCES..... 158

Chapter 6 Efficient Electronic Structure Theories of Investigating Excited State Dynamics and Modeling Potential Energy Surfaces In The Vicinity of Conical Intersection..... 164
6.1 The Arising of Unphysical Symmetry Breaking of The Wavefunction in SA-CASSCF Wavefunction Ansatz..... 168
6.1.1 Orbital Pseudo-Jahn-Teller Effect (OPJTE)..... 168
6.1.2 Symmetry breaking in He_2^+ as an example of OPJTE..... 172
6.1.3 Singly-excited active space self-consistent field method..... 177
6.2 Effective Selection of The Determinants in Configuration Interaction Hamiltonian By Utilizing Natural Orbitals..... 178
6.2.1 SEAS and HF orbitals are not always most efficient orbitals in achieving small active spaces..... 178

6.2.2 Configuration interaction singles natural orbitals form an efficient orbital space for configuration interaction expansion..... 182
6.3 Benchmark Calculations for Two-step Methods..... 191
6.3.1 Size intensivity..... 191
6.3.2 Symmetry breaking..... 196
6.3.3 Two-step methods in describing n->π* excitation..... 199
6.3.4 Quantitative accuracy in description of the conical intersection..... 204
6.4 Conclusions and Outlook..... 206
APPENDIX..... 208
REFERENCES..... 231

Chapter 7 Non-Radiative Recombination via Conical Intersection at Silicon Epoxide Defect..... 236
REFERENCES..... 249

Chapter 8 Non-Radiative Recombination via Conical Intersections Arising at Defects on The Oxidized Silicon Surface..... 252
8.1 Computational Methods..... 253
8.2 Results..... 257
8.2.1 Dynamical simulation of the 3O cluster..... 257
8.2.2 Potential energy surface analysis of the 3O cluster..... 261
8.2.3 Dynamical simulation of the 5O cluster..... 270
8.2.4 Potential energy surface analysis of the 5O cluster..... 274
8.3 Discussion..... 281
8.4 Conclusions..... 284
REFERENCES..... 286

Chapter 9 Do Silicon-Oxygen Double Bonds Emit Light?..... 289
9.1 Methods..... 289
9.1.1 Nonadiabatic molecular dynamics simulations..... 289
9.1.2 Ab-Initio calculations..... 291
9.2 Results and Discussion..... 292
9.2.1 Dynamical simulations..... 292
9.2.2 Analysis of the potential energy surface..... 299
9.2.3 Discussion of excited state lifetime..... 308
9.3 Conclusions..... 310
REFERENCES..... 312

Chapter 10 Defect-Induced Conical Intersections Promote Non-Radiative Recombination: A Study at Nano-Sized Cluster..... 316
10.1 Cluster Models..... 318

10.2 Computational Methods..... 321
10.3 Size Insensitive Defects Conical Intersections..... 323
10.4 Recombination Rate Analysis..... 328
10.5 Conclusions..... 335
REFERENCES..... 336

Chapter 11 Silicon Cluster Terminated with H Atoms..... 342
11.1 Methods..... 344
11.1.1 Nonadiabatic molecular dynamics simulations..... 344
11.1.2 Quantum chemical calculations..... 345
11.2 Results and Discussion..... 346
11.2.1 Dynamical simulation of the silicon cluster..... 346
11.2.2 Potential energy surface analysis of the silicon cluster..... 353
11.3 Conclusions..... 364
APPENDIX..... 366
REFERENCES..... 372

Chapter 12 Silicon Nanocrystal with Hydroxide Group at Surface..... 376
12.1 Methods..... 377
12.1.1 Nonadiabatic molecular dynamics simulations..... 377
12.1.2 Quantum chemical study of silanol cluster..... 378
12.1.3 Quantum chemical study of SiNC..... 378
12.2 Results and Discussion..... 380
12.2.1 Dynamical simulation of the silicon cluster with hydroxide defect..... 380
12.2.2 PES analysis of silanol cluster..... 386
12.2.3 Conical intersection in SiNC..... 393
12.3 Conclusions..... 396
APPENDIX..... 398
REFERENCES..... 411

Chapter 13 Surface Structure and Silicon Nanocrystal Photoluminescence: The Role of Hypervalent Silyl Groups..... 415
13.1 Methods..... 418
13.1.1 Experimental methods..... 418
13.1.2 Computational details..... 419
13.2 Results and Discussion..... 421
13.2.1 Effect of heating on properties of SiNCs..... 421
13.2.2 Bonding and excited state dynamics of silyl defects..... 430
13.3 Conclusions..... 439
REFERENCES..... 441

Chapter 14 Recombination Mechanism of Tantalum Nitride Ta_3N_5: Why Is a Sacrificial Reagent Required? 446
14.1 Methods..... 451
14.1.1 Nonadiabatic molecular dynamics simulations..... 451
14.1.2 Ab-Initio calculations..... 452
14.2 Results and Discussion..... 453
14.2.1 Model cluster and ground state minimum..... 453
14.2.2 Dynamical simulation of the model cluster..... 456
14.2.3 Potential energy surface analysis..... 462
14.2.4 Spin-orbit coupling effect..... 467
14.3 Efficient Ta_3N_5 Based Water Splitting Photocatalyst Design..... 468
14.4 Conclusions..... 469
APPENDIX..... 470
REFERENCES..... 476

Chapter 15 Non-Radiative Recombination Mechanism of Lead Halide Perovskite Material..... 483
15.1 Methods..... 485
15.1.1 Nonadiabatic molecular dynamics simulations..... 485
15.1.2 Ab-Initio calculations..... 486
15.2 Results and Discussion..... 487
15.2.1 Model cluster and ground state minimum..... 487
15.2.2 Dynamical simulation of the CP model cluster..... 490
15.2.3 Potential energy surface analysis of CP cluster..... 498
15.3 Conclusions..... 509
APPENDIX..... 510
REFERENCES..... 516

Chapter 16 New Paradigm for Non-Radiative Recombination Mechanism in Optoelectronic Materials..... 524
16.1 Conical Intersection Facilitates the Non-Radiative Recombination in Optoelectronic Materials..... 529
16.2 Some Successful Applications of Localized Nuclear Deformation Induced MECI Prompted Non-Radiative Recombination Mechanism..... 533
16.2.1 SiNC with oxygen containing defects..... 533
16.2.2 Tantalum nitride..... 536
16.2.3 Lead Halide perovskites..... 536
16.3 Conclusions and Outlook..... 539
REFERENCES..... 542

Chapter 17 Simulated Evolution of Fluorophores for Light Emitting Diodes..... 551

17.1 Method..... 554
17.1.1 Tree-based genetic algorithm..... 554
17.1.2 Definition of molecular spaces..... 559
17.2 Fitness Evaluation..... 562
17.3 Results and Discussion..... 565
17.3.1 Overview of GA results..... 565
17.3.2 GA performance..... 567
17.3.3 Candidate fluorophores..... 571
17.4 Conclusions..... 577
REFERENCES..... 579

Chapter 18 Efficient Prediction of physical Properties Based On Machine Learning Techniques..... 586
18.1 Geometrical Many-Body Expansion Model and Molecular Representation..... 588
18.1.1 Geometrical many-body expansion model..... 588
18.1.2 Expansion truncation..... 591
18.1.3 Generating the many-body interaction vector..... 592
18.1.4 The Representation Model..... 594
18.2 Supervised Learning..... 595
18.2.1 Supervised linear learning problem..... 595
18.2.2 Cross-Validation for two-body interaction model..... 596
18.2.3 Three-body interaction model..... 598
18.2.4 Model improvement..... 599
18.3 Visualization of The High Dimensional Data With The Geometrical Many-Body Expansion Model..... 602
18.4 Conclusions..... 605
REFERENCES..... 608

LIST OF TABLES

Table 3-1. The standard deviations of symmetry equivalent geometric parameters of the three T_d molecules studied. The very small values demonstrate that the nuclear geometries are nearly but not exactly symmetric……..44

Table 3-2. For each of the three T_d molecules, this table includes the number of states averaged in SA-CASSCF calculations, the set of states which are expected to be degenerate by symmetry, and the sets of orbitals expected to be degenerate by symmetry. The average gap between the expected degenerate states is used as an indicator of whether the wavefunction preserves symmetry or not…… 45

Table 3-3. Dipole moments and the average energy gaps between states expected to be degenerate by symmetry calculated at single-reference levels of theory with symmetry-preserving references. All values are expected to be zero by symmetry. The very small deviations indicate that the slight asymmetries of the nuclear configurations do not result in large changes in these observables…… 46

Table 3-4. The number of symmetry-preserving active spaces for each combination of molecule and method. These counts are referenced to the total number of active spaces considered. Application of SEAS-CASCI and freezing the closed-shell orbital coefficients at their RHF values are successful strategies for enforcing the symmetry of the wavefunction…… 47

Table 4-1. Excitation energies of the low-lying excited states of the test set molecules computed at the state averaged CASSCF and CISNO-CASCI levels of theory. The active spaces and numbers of states averaged for these calculations are also presented. The number of states averaged includes the HF ground state in the case of CISNO-CASCI. MS-CASPT2 results are shown for comparison. The smaller active spaces and numbers of correlated orbitals used for the MS-CASPT2 calculations are shown in Table 4-2. Adiabatic state labels are assigned according to the state ordering at the CASSCF level of theory. The labels of the states computed by all other methods are assigned by comparison of chemical character (orbitals and CI vectors) to those of CASSCF. The error is defined as the difference between the excitation energy computed at the CASSCF or CISNO-CASCI level of theory, and that computed at the MS-CASPT2 level…… 85

Table 4-2. Excitation energies of the low-lying excited states of the test set molecules computed at the CISNO-CASCI, CISNO-CASCI-RS2, SA-CASSCF and MS-CASPT2 theoretical levels. Note that smaller active spaces are used than in Table 4-1 for computational convenience. All adiabatic state labels are assigned according to the characters of the CASSCF states in Table 4-1. Where possible, gas phase experimental data has been presented for comparison…….. 88

Table 4-3. Excitations energies computed at the SA-CASSCF/6-31G* and CISNO-CASCI/6-31G* levels of theory for ethylene, butadiene, hexatriene, and benzene microsolvated by 250 methanol molecules. The SA-CASSCF computations average over the ground state and excited states reported below. The CISNOs include only the ground state and lowest excited state in the averaging. The active spaces include all π and π* orbitals located on the chromophore…….. 98

Table 4-4. Absolute ground state energy factors (*y* as defined in Eq. (4-6)) computed at the SA-CASSCF, HF-CASCI, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, and FOMO-CASCI theoretical levels for a series of systems comprising different numbers of non-interacting hydrogen molecules…….. 105

Table 4-5. Absolute ground state energy factors (*y* as defined in Eq. (4-6)) computed at the SA-CASSCF, HF-CASCI, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, and FOMO-CASCI theoretical levels for a series of systems comprising different numbers of non-interacting ethylene molecules…….. 106

Table 5-1. The S_1, S_0 and S_1-S_0 energy gap (all in the unit of eV) computed with SA-CASSCF, CISNO-CASC, CISNOr-CASCI, SEAS-CASCI, HF-CASCI and MRCI(Q) methods are shown at FC geometry and MECI geometries optimized at SA-CASSCF, CISNO-CASCI, CISNOr-CASCI, SEAS-CASCI theoretical levels……..141

Table 5-2. The S_1, S_0 and S_1-S_0 energy gap (all in the unit of eV) computed with SA-CASSCF, CISNO-CASC, CISNOr-CASCI, SEAS-CASCI and HF-CASCI methods without doubly excited determinants in the CASCI Hamiltonian are shown at the MECI geometries optimized at SA-CASSCF, CISNO-CASCI, CISNOr-CASCI, SEAS-CASCI theoretical levels…….. 143

Table 5-S1. The absolute energies of the FC and various optimized geometries of the ethylene pyramidalized MECI at the CASSCF level of theory (with active space and basis sets as described in the text)…….. 150

Table 5-S2. The absolute energies of the FC and various optimized geometries of the ethylene pyramidalized MECI at the CISNO-CASCI level of theory (with active space and basis sets as described in the text)..... 150

Table 5-S3. The absolute energies of the FC and various optimized geometries of the ethylene pyramidalized MECI at the CISNOr-CASCI level of theory (with active space and basis sets as described in the text)..... 151

Table 5-S4. The absolute energies of the FC and various optimized geometries of the ethylene pyramidalized MECI at the SEAS-CASCI level of theory (with active space and basis sets as described in the text)..... 151

Table 5-S5. The absolute energies of the FC and various optimized geometries of the ethylene pyramidalized MECI at the HF-CASCI level of theory (with active space and basis sets as described in the text)..... 152

Table 5-S6. The absolute energies of the FC and various optimized geometries of the ethylene pyramidalized MECI at the MRCI(Q) level of theory (with active space and basis sets as described in the text)..... 152

Table 5-S7. The absolute energies of the various optimized geometries of the ethylene pyramidalized MECI at the SA-CASSCF without doubly excited determinant level of theory (with active space and basis sets as described in the text)..... 153

Table 5-S8. The absolute energies of the various optimized geometries of the ethylene pyramidalized MECI at the CISNO-CASCI without doubly excited determinant level of theory (with active space and basis sets as described in the text)..... 153

Table 5-S9. The absolute energies of the various optimized geometries of the ethylene pyramidalized MECI at the CISNOr-CASCI without doubly excited determinant level of theory (with active space and basis sets as described in the text)..... 154

Table 5-S10. The absolute energies of the various optimized geometries of the ethylene pyramidalized MECI at the SEAS-CASCI without doubly excited determinant level of theory (with active space and basis sets as described in the text)..... 154

Table 5-S11. The absolute energies of the various optimized geometries of the ethylene

pyramidalized MECI at the HF-CASCI without doubly excited determinant level of theory (with active space and basis sets as described in the text)..... 155

Table 6-1. The mean absolute error and mean signed error of S_1 excitation energy between two-step approaches and EOM-CCSD method. The S_1 state has n->π* excitation character..... 205

Table 6-S1. The absolute energies of S_0 and S_1 states of $Ti(OH)_4$ FC geometry at SEAS-RASCI, CISNOr-RASCI and CISNO-RASCI theoretical levels with various EFOs, corresponding to Figure 6-3(a), Figure 6-4(a) and Figure 6-6(a)..... 209

Table 6-S2. The absolute energies of S_0 and S_1 states of $Ni(CO)_4$ FC geometry at SEAS-RASCI, CISNOr-RASCI and CISNO-RASCI theoretical levels with various EFOs, corresponding to Figure 6-3(b), Figure 6-4(b) and Figure 6-6(b)..... 210

Table 6-S3. The absolute energies of S_0 and S_1 states of $Ti(OH)_4$ FC geometry at CISNOr-CASCI, SEAS-CASCI, HF-CASCI theoretical levels with various active spaces, corresponding to Figure 6-5(a)..... 211

Table 6-S4. The absolute energies of S_0 and S_1 states of $Ni(CO)_4$ FC geometry at CISNOr-CASCI, SEAS-CASCI, HF-CASCI theoretical levels with various active spaces, corresponding to Figure 6-5(b) 212

Table 6-S5. The absolute energies of S_0 and S_1 states of $Ti(OH)_4$ and $Ni(CO)_4$ FC geometries at EOM-CCSD theoretical level, corresponding to Figure 6-5..... 213

Table 6-S6. The absolute energies of S_0 and S_1 states of pyrazine FC geometries at SA-CASSCF, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, FOMO-CASCI, CISNOr-CASCI, SA-ASCISSCF with various active spaces and EOM-CCSD theoretical levels, corresponding to Figure 6-11..... 214

Table 6-S7. The absolute energies of S_0 and S_1 states of thymine FC geometries at SA-CASSCF, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, FOMO-CASCI, CISNOr-CASCI, SA-ASCISSCF with various active spaces and EOM-CCSD theoretical levels, corresponding to Figure 6-11..... 216

Table 6-S8. The absolute energies of S_0 and S_1 states of formaldehyde FC geometries at SA-CASSCF, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, FOMO-CASCI, CISNOr-CASCI, SA-ASCISSCF with various active spaces and EOM-CCSD theoretical

levels, corresponding to Figure 6-11…… 219

Table 6-S9. The absolute energies of S_0 and S_1 states of pyradizine FC geometries at SA-CASSCF, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, FOMO-CASCI, CISNOr-CASCI, SA-ASCISSCF with various active spaces and EOM-CCSD theoretical levels, corresponding to Figure 6-11…….. 222

Table 6-S10. The absolute energies of S_0 and S_1 states of cytosine FC geometries at SA-CASSCF, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, FOMO-CASCI, CISNOr-CASCI, SA-ASCISSCF with various active spaces and EOM-CCSD theoretical levels, corresponding to Figure 6-11…….. 224

Table 6-S11. The absolute energies of S_0 and S_1 states of uracil FC geometries at SA-CASSCF, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, FOMO-CASCI, CISNOr-CASCI, SA-ASCISSCF with various active spaces and EOM-CCSD theoretical levels, corresponding to Figure 6-11…….. 226

Table 8-1. The size of the active space (abbreviated: number of active electrons/number of active orbitals) and number of states to be averaged in our SA-CASSCF calculations for each cluster. Also given are the numbers of occupied orbitals to be correlated in the MS-CASPT2 and MRCI(Q) calculations. Note that more orbitals are correlated in the MS-CASPT2 calculations at the spawning points from our AIMS simulations (reported in Figures 8-7 and 13) than in the remainder of the MS-CASPT2 calculations reported in this work…… 254

Table 8-2. The S_0 and S_1 energies (in eV) at five important points on the PES of the 3O cluster calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory. S_1 minima and S_1-S_0 MECI geometries have been optimized at both the CASSCF and MS-CASPT2 levels of theory…… 264

Table 8-3. The S_0 and S_1 energies (in eV) at four important points on the PES of the 5O cluster calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory. S_1 minimum and S_1-S_0 MECI geometries have been optimized at both the CASSCF and MS-CASPT2 levels of theory…… 276

Table 9-1. The S_0 and S_1 energies (in eV) at five important points on the PES of the Si=O bond containing cluster calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory. Excited state geometries are optimized at both the CASSCF and MS-CASPT2 levels of theory. The pyramidalized geometries are referred to generically as S_1-S_0 CI and S_1 min, while additional features in the

oxygen insertion region of the PES are prefixed by O-i…….…… 301

Table 10-1. Energies of the first singlet excited states of the three SiNCs pictured in Figure 10-1 at their respective Franck-Condon points (FC) and optimized conical intersections (CI). Energies are reported relative to the ground state minimum energy.…….…… 330

Table 11-1. The S_0 and S_1 energies (in eV) at three important points on the PES of the silicon cluster calculated at the SA- CASSCF, MS-CASPT2, and MRCI(Q) levels of theory. S_1-S_0 MECI geometries are optimized at both the SA-CASSCF and MS-CASPT2 levels of theory…….…… 356

Table 11-S1. The absolute energies of the FC and various optimized geometries of the S_0/S_1 MECIs at the SA-CASSCF level of theory (with active space and basis sets as described in the text)…….. 367

Table 11-S2. The absolute energies of the FC and various optimized geometries of the S_0/S_1 MECIs at the MS-CASPT2 level of theory (with active space and basis sets as described in the text)…….. 367

Table 11-S3. The absolute energies of the FC and various optimized geometries of the S_0/S_1 MECIs at the MRCI(Q) level of theory (with active space and basis sets as described in the text)…….. 368

Table 11-S4. The absolute energies of the FC geometry at the EOM-CCSD level of theory…….. 368

Table 12-1. The S_0 and S_1 energies (in eV) at three important points on the PES of the silicon hydroxide cluster calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory. S_1-S_0 MECI geometries are optimized at both the CASSCF and MS-CASPT2 levels of theory…….…… 388

Table 12-S1. The absolute energies of the various optimized geometries of the cluster model at the CASSCF level of theory (with active space and basis sets as described in the text)…….. 399

Table 12-S2. The absolute energies of the various optimized geometries of the cluster model at the MS-CASPT2 level of theory (with active space and basis sets as described in the text)…….. 399

Table 12-S3. The absolute energies of the various optimized geometries of the cluster

model at the MRCISD(Q) level of theory (with active space and basis sets as described in the text)..... 400

Table 12-S4. The absolute energies of the Franck-Condon point of the cluster model at the EOM-CCSD level of theory..... 400

Table 12-S5. The absolute energies of the various optimized geometries of the cluster model at the CISNOr-CASCI level of theory (with active space and basis sets as described in the text)..... 400

Table 12-S6. The absolute energies of the various optimized geometries of the nanoparticle at the CISNOr-CASCI level of theory (with active space and basis sets as described in the text)..... 401

Table 14-1. The S_0 and S_1 energies (in eV) at two important points on the PES of the model tantalum nitride cluster calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory. S_1-S_0 MECI geometries are optimized at both the CASSCF and MS-CASPT2 levels of theory..... 464

Table 14-S1. The absolute energies of the various optimized geometries of the Ta_3N_5 cluster model at the SA-CASSCF level of theory (with active space and basis sets as described in the text)..... 471

Table 14-S2. The absolute energies of the various optimized geometries of the Ta_3N_5 cluster model at the MS-CASPT2 level of theory (with active space and basis sets as described in the text)..... 471

Table 14-S3. The absolute energies of the various optimized geometries of the Ta_3N_5 cluster model at the MRCI(Q) level of theory (with active space and basis sets as described in the text)..... 472

Table 14-S4. The absolute energies of the FC geometry of the Ta_3N_5 cluster model at the EOM-CCSD level of theory..... 472

Table 15-1. Size of the Active Space (Abbreviated: Number of Active Electrons/Number of Active Orbitals) and Number of States to Be Averaged in Our SA-CASSCF Calculations for CP Cluster. Also given are the numbers of occupied orbitals to be correlated in the MS-CASPT2 calculations. Note that more orbitals are correlated in the MS-CASPT2..... 491

Table 15-2. The S_0 and S_1 energies (in eV) at three important points on the PES of the CP

cluster calculated at the CASSCF and MS-CASPT2 levels of theory. S_1-S_0 MECI and S_1min geometries are optimized at both the CASSCF and MS-CASPT2 levels of theory. Note that all SA-CASSCF calculations are performed with SBKJC_VDZ basis set while all MS-CASPT2 calculations are performed with SBKJC_VDZ(p,2d) basis set…… 503

Table 15-S1. The absolute energies of the various optimized geometries of the CP cluster model at the SA-CASSCF level of theory (with active space and basis sets as described in the text)…… 511

Table 15-S2. The absolute energies of the various optimized geometries of the CP cluster model at the MS-CASPT2 level of theory (with active space and basis sets as described in the text)…… 512

Table 15-S3. The absolute energies of the FC geometry of the CP cluster model at the EOM-CCSD level of theory…… 512

Table 17-1. Information pertaining to the eight individual optimizations performed in this work is presented, including the user defined parameters (N_{pool} and P_{mut}), the total number of fitness evaluations performed, and the structures, fitness values, singlet-triplet energy gaps, and S_1 transition dipole moments of the three fittest genes computed in each optimization…… 568

LIST OF FIGURES

Figure 3-1. a) and c) The average energy difference between degenerate states (as defined in Table 3-2) as a function of active space for sila-adamantane (a) and adamantane (c) calculated using various multireference methods. b) and d) The S_0 dipole moment as a function of active space for sila-adamantane (b) and adamantane (d) calculated using the same set of multireference methods. Deviation of either value from zero indicates wavefunction symmetry breaking. SEAS-CASCI predicts a symmetry-preserving wavefunction in more cases than CASSCF. Freezing of closed orbitals in their RHF configurations also decreases the tendency toward symmetry breaking. The sila-adamantane and adamantane molecules are pictured in the insets of panels a) and c), with silicon, carbon, and hydrogen atoms represented in blue, gray, and white, respectively..... 51

Figure 3-2. a) The average energy difference between degenerate states (as defined in Table 3-2) as a function of active space for the linear titanium hydroxide cluster calculated using various multireference methods. b) The S_0 dipole moment as a function of active space for the linear titanium hydroxide cluster using the same set of multireference methods. Deviation of either value from zero indicates wavefunction symmetry breaking. SEAS-CASCI predicts a symmetry-preserving wavefunction in more cases than CASSCF. Freezing of closed orbitals in their Hartree-Fock configurations also decreases the tendency toward symmetry breaking. The linear $Ti(OH)_4$ cluster is pictured in the inset of a), with titanium, oxygen, and hydrogen atoms represented in yellow, red, and white, respectively..... 53

Figure 3-3. The first vertical excitation energies (averaged over nominally degenerate states) of the three T_d clusters as a function of active space calculated using several different multireference methods. A thin black line indicates the EOM-CCSD reference energies. In all cases errors relative to EOM-CCSD are significant, but SEAS-CASCI performs similarly well to CASSCF..... 55

Figure 3-4. a) The PES of the first excited state of beryllium hydroxide as a function of a rigid asymmetric stretching (AS) coordinate calculates at three levels of theory. CASSCF predicts a large cusp in the PES at the C_{2v} symmetric (zero AS) geometry, while the SEAS-CASCI surface is very flat with respect to AS. The shape of the SEAS-CASCI curve is in much better agreement with the

symmetry-preserving EOM-CCSD reference. b) The z component of the S_1 dipole moment of beryllium hydroxide as a function of the AS coordinate calculated at the same levels of theory. This value should be zero by symmetry at the C_{2v} structure, but CASSCF predicts a large non-zero value, indicating a symmetry-broken wavefunction. A sudden change in the dipole at zero indicates discontinuous behavior of the CASSCF wavefunction. The relatively smooth curve predicted by SEAS-CASCI is in better agreement with the EOM-CCSD reference..... 58

Figure 3-5. The a) PES and b) norm of the trace of the dipole moment operator over states S_1, S_2, and S_3 (M; defined in Eq.(3-4)) for the linear $Ti(OH)_4$ molecule as a function of an asymmetric stretching coordinate (illustrated in the inset). The SEAS-CASCI PES (orange) is smoothly varying and in qualitative agreement with EOM-CCSD (black), while the CASSCF surface (red) has unphysical cusps at the T_d symmetric geometry and is in poor agreement with EOM-CCSD. SEAS-CASCI predicts a dipole moment curve in good agreement with EOM-CCSD, while the CASSCF dipole moment curve changes discontinuously at the T_d geometry..... 60

Figure 3-6. The a) PES and b) norm of the trace of the dipole moment operator over states S_1, S_2, and S_3 (M; defined in Eq. (3-4)) for the bent $Ti(OH)_4$ molecule as a function of an asymmetric stretching coordinate (illustrated in the inset). The SEAS-CASCI PES (orange) is smoothly varying and in near-quantitative agreement with EOM-CCSD (black), while the CASSCF surface (red) is both in poor agreement with EOM-CCSD and relatively unsmooth. SEAS-CASCI predicts a dipole moment in near-quantitative agreement with EOM-CCSD..... 63

Figure 3-7. The PES of ethylene as a function of the rigid twisting and pyramidalization of fully optimized conical intersection structures (pictured at the top of each panel) calculated at the CASSCF (a) and SEAS-CASCI (b) levels of theory. The SEAS-CASCI optimized structure and PES are in excellent agreement with CASSCF, indicating the ability of SEAS-CASCI to describe biradicaloid regions of the PES, despite the lack of invariance to rotations between HOMO and LUMO in the initial SEAS SCF step..... 66

Figure 4-1. The state-averaged absolute energies of (a) pyrazine, (b) thymine, (c) pentadiene iminium, (d) benzene, (e) trihydroxy(oxo) vanadium, (f) ferrocyanide, (g) tetracyanonickelate, (h) iron pentacarbonyl, and (i) sila-adamantane molecules computed at the SA-CASSCF, CISNO-CASCI, SEAS-CASCI, and IVO-CASCI theoretical levels as a function of active space..... 92

Figure 4-2. The π* orbital of ethylene microsolvated by 250 methanol molecules plotted for different orbital choices: (a) a SA-2-CASSCF natural orbital, (b) a CISNO averaged over the ground state and one excited state, (c) a HF virtual orbital LUMO+10 of the cluster, (d) a HF virtual orbital LUMO+11 of the cluster, (e) a HF virtual orbital LUMO+12 of the cluster. The CISNO and SA-CASSCF orbitals are plotted with an isovalue of 0.01. The HF virtual orbitals are plotted with an isovalue of 0.025. All computations described in this Figure were performed in a 6-31G basis set..... 97

Figure 4-3. (a) The excitation energies of the monomer-localized singly excited states of systems of non-interacting H_2 molecules computed at the SA-CASSCF, HF-CASCI, SEAS-CASCI, CISNO-CASCI, IVO-CASCI, FOMO-CASCI, and EOM-CCSD theoretical levels as a function of the number of non-interacting subsystems are shown in red, green, orange, blue, purple, cyan, and gray, respectively. (b) Same as (a), but with a narrower energy scale for clarity. (c) Same as (a), but for systems of non-interacting ethylene molecules..... 101

Figure 4-4. Excitation energies of the singly excited state of monomers of hydrogen and ethylene, and of the doubly excited states of non-interacting dimers of these systems in which each monomer is singly excited. Notably, the excitation energies of the dimers are exactly twice those of the monomers, indicating excited state size consistency..... 108

Figure 4-5. Illustration of the density matrices defined in equations (10) and (11) for case $M = 2$..... 110

Figure 4-6. PESs of ethylene as a function of rigid twisting and pyramidalization coordinates computed at the (a) SA-CASSCF and (b) CISNO-CASCI theoretical levels. The intersection points in (a) and (b) reflect the MECIs optimized at their respective levels of theory..... 115

Figure 5-1. S_1 and S_0 PESs PESs near conical intersection geometries of (a) ethylene molecule, calculated on SA-CASSCF theoretical level, (b) ammonia molecule, calculated on SA-CASSCF theoretical level, (c) ethylene molecule, calculated on SA-CISSCF theoretical level, (d) ammonia molecule, calculated on SA-CISSCF theoretical level, (e) ethylene molecule, calculated on HF-CIS theoretical level, (f) ammonia molecule, calculated on HF-CIS theoretical level..... 134

Figure 5-2. (a) S_1 and S_0 PESs of ammonia molecule near conical intersection geometry calculated on CISNOr-CIS theoretical level. (b), a corresponding contour plot

of the S_1-S_0 energy gap…….. 137

Figure 6-1. Computational cost comparison between SA-CASSCF and CISNO-CASCI for sila-adamantane and titanium hydroxide system for various active spaces…….. 167

Figure 6-2. Upper panel: Black shows the eigenvalues (E_+ and E_-) of the CAS(3/2)/6-31G Hamiltonian describing He2+ as a function of a distortion of the orbitals from variationally optimal symmetry constrained orbitals ($C=0$) to fully optimized orbitals (no symmetry constraint, $C=\delta C$). The diagonal matrix elements of the Hamiltonian (E_A and E_B) are presented in blue. Despite the fact that the diagonal elements of the Hamiltonian are increased by symmetry breaking, the variationally optimal solution ($E_-(\delta C)$) is symmetry-broken because symmetry breaking introduces strong coupling between configurations A and B. The square markers indicate data that was calculated, while the lines connecting them are interpolated assuming quadratic behavior of the Hamiltonian matrix elements. Lower panel: The orbital basis from which the CAS(3/2) Hamiltonian is constructed. The dashed lines guide the eye to see that the orbitals are symmetric when $C=0$ and symmetry-broken when $C=\delta C$. Configuration A corresponds to double occupation of the σ orbital and single occupation of the σ* orbital; the occupations are reversed in configuration B…….. 176

Figure 6-3. (a) The SEAS orbital of $Ti(OH)_4$ and associated closed shell determinant constituted by SEAS orbitals, the number indexes the orbitals with the vertical alignment of the orbitals indicate the degeneracy. The inset shows the $Ti(OH)_4$ molecule. (b) The absolute difference between SEAS-RASCI and EOM-CCSD of $Ti(OH)_4$ in the unit of eV. RASCI is defined in the text. (c) The SEAS orbital of $Ni(CO)_4$ and associated closed shell determinant constituted by SEAS orbitals, the number indexes the orbitals with the vertical alignment of the orbitals indicate the degeneracy. The inset shows the $Ni(CO)_4$ molecule. (b) The absolute difference between SEAS-RASCI and EOM-CCSD of $Ni(CO)_4$ in the unit of eV. RASCI is defined in the text…….. 181

Figure 6-4. (a) The CISNOr of $Ti(OH)_4$ and associated closed shell determinant constituted by CISNOr, the number indexes the orbitals with the vertical alignment of the orbitals indicate the degeneracy of the occupation numbers of CISNOr. The inset shows the $Ti(OH)_4$ molecule. (b) The absolute difference between CISNOr-RASCI and EOM-CCSD of $Ti(OH)_4$ in the unit of eV. RASCI is defined in the text. (c) The CISNOr of $Ni(CO)_4$ and associated

closed shell determinant constituted by CISNOr, the number indexes the orbitals with the vertical alignment of the orbitals indicate the degeneracy of the occupation numbers of CISNOr. The inset shows the $Ni(CO)_4$ molecule. (b) The absolute difference between CISNOr-RASCI and EOM-CCSD of $Ni(CO)_4$ in the unit of eV. RASCI is defined in the text…… 186

Figure 6-5. The S_1 excitation energy as a function of active spaces for (a) $Ti(OH)_4$ and (c) $Ni(CO)_4$ molecules calculated at CISNOr-CASCI, SEAS-CASCI, HF-CASCI and EOM-CCSD theoretical levels. (b) and (d) lists the active spaces employed in CISNOr-CASCI, SEAS-CASCI and HF-CASCI…… 188

Figure 6-6. (a) The CISNO of $Ti(OH)_4$ and associated closed shell determinant constituted by CISNO, the number indexes the orbitals with the vertical alignment of the orbitals indicate the degeneracy of the occupation numbers of CISNO. The inset shows the $Ti(OH)_4$ molecule. (b) The absolute difference between CISNO-RASCI and EOM-CCSD of $Ti(OH)_4$ in the unit of eV. RASCI is defined in the text. (c) The CISNO of $Ni(CO)_4$ and associated closed shell determinant constituted by CISNO, the number indexes the orbitals with the vertical alignment of the orbitals indicate the degeneracy of the occupation numbers of CISNO. The inset shows the $Ni(CO)_4$ molecule. (b) The absolute difference between CISNO-RASCI and EOM-CCSD of $Ni(CO)_4$ in the unit of eV. RASCI is defined in the text…… 189

Figure 6-7. (a) The PES of the first excited state of beryllium hydroxide as a function of a rigid asymmetric stretching (AS) coordinate (as illustrated in the inset) calculates at three levels of theory. CASSCF predicts a large cusp in the PES at the C_{2v} symmetric (zero AS) geometry, while the CISNO-CASCI surface is much smoother with respect to AS. The shape of the CISNO-CASCI curve is in much better agreement with the symmetry-preserving EOM-CCSD reference. (b) and (c) The PESs for the linear $Ti(OH)_4$ and bent $TI(OH)_4$ molecules as a function of an asymmetric stretching coordinate (illustrated in the insets). The CISNO-CASCI PESs (blue) are smoothly varying and in qualitative agreement with EOM-CCSD (black), while the CASSCF surfaces (red) have unphysical cusps at the T_d symmetric geometry for linear $Ti(OH)_4$ and is in poor agreement with EOM-CCSD…… 190

Figure 6-8. (a) The excitation energies of the monomer-localized singly excited states of systems of non-interacting H_2 molecules computed at the CIS, EOM-CCSD, Full CI, SA-CASSCF, SEAS-CASCI, CISNO-CASCI, CISNOr-CASCI, IVO-CASCI, FOMO-CASCI, HF-CASCI, CISNOr-CIS and SA-CISSCF theoretical levels as a

function of the number of non-interacting subsystems are shown in grey, dark grey, black, red, orange, blue, green, purple, cyan, light red, magenta and yellow respectively. (b) Same as (a), but with a narrower energy scale for clarity…… 192

Figure 6-9. The HF ground state and CIS excited states relaxed first order reduced density matrices for one, two and three non-interacting subsystem…… 195

Figure 6-10. (a) The averaged energy difference between degenerate states as a function of active spaces for the linear titanium hydroxide cluster calculated at SA-CASSCF, SEAS-CASCI, CISNO-CASCI, CISNOr-CASCI, IVO-CASCI, FOMO-CASCI and EOM-CCSD theoretical levels. (b) The active spaces employed at each level of multi-reference theories…… 198

Figure 6-11. The excitation energies of n->π* state as a function of active spaces employed in SA-CASSCF and various two-step methods. The SA-CASSCF, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, FOMO-CASCI, CISNO-CASCI, SA-ASCISSCF and EOM-CCSD are shown as red, blue, orange, green, purple, cyan, yellow and grey respectively for (a) cytosine, (b) formaldehyde, (c) pyradizine, (d) pyrazine, (e) thymine and (f) uracil molecules…… 202

Figure 6-12. The n->π* excited state characters for (a) cytosine, (b) formaldehyde, (c) pyradizine, (d) pyrazine, (e) thymine and (f) uracil molecules. The n and π* orbitals shown here are HF orbitals…… 203

Figure 7-1. The population of adiabatic states S_0 and S_1 as a function of time after excitation. The average over all 20 initial nuclear basis functions is shown in bold red and blue for S_0 and S_1, respectively, while the values associated with individual initial basis functions are represented by thin lines. The $Si_8H_{12}O$ cluster is shown in the inset, with silicon, oxygen, and hydrogen atoms represented by blue, red, and white, respectively…… 240

Figure 7-2. The a) S_1-S_0 energy gap and b) the Si-O-Si angle of the epoxide ring (defined in the inset) as a function of time after excitation. The average over the 20 initial basis functions on S_1 is shown in bold red lines, while the values of the individual basis functions are presented in thin gray lines. (Population on S_0 is excluded from this average.)…… 242

Figure 7-3. The S_0 and S_1 energies of three important points on the PES of the silicon epoxide cluster: the Franck-Condon (FC) point, the S_0/S_1 MECI, and the S_1 min. Black and red lines correspond to the energies of the CASSCF

optimized geometries computed at the CASSCF and CASPT2 levels of theory, respectively. Energies are shifted so that the S_0 energy of the FC point is zero for both levels of theory. Only one black line is visible for the S_0/S_1 MECI because S_0 and S_1 are degenerate. Optimized geometries are shown as insets..... 243

Figure 7-4. Partially occupied natural orbitals of the CASSCF wavefunctions at the FC point (left) and the optimized S_0/S_1 MECI structure (right). Two different views of each orbital are presented for the FC point. Occupation numbers, *n*, are presented for each orbital. Orbitals were determined as described in the text..... 245

Figure 8-1. The 3O (left) and 5O (right) clusters..... 257

Figure 8-2. The population of S_1 (bold blue) and S_0 (bold red) of the 3O cluster as a function of time after excitation averaged over all simulations. Thin blue and pink lines report the populations corresponding to the twenty individual simulations..... 258

Figure 8-3. Average S_1-S_0 energy gap of the 3O cluster as a function of time (bold red), with the energy gaps of the twenty individual initial nuclear basis functions shown in gray. Snapshots from representative trajectories are shown in the inset, with silicon, oxygen, and hydrogen atoms colored blue, red, and white, respectively..... 259

Figure 8-4. (a). Average O-SiOO pyramidalization angle of the 3O cluster as a function of time (bold red) with pyramidalization angles of the twenty individual initial nuclear basis functions (gray). The inset indicates the atoms that define the pyramidalization angle. (b). The distance between epoxide silicon atoms as a function of time averaged over twenty simulations (bold red) and the values corresponding to the twenty individual simulations themselves (gray). The inset indicates the two silicon atoms between which the distance was calculated..... 262

Figure 8-5. (a). S_0 and S_1 energies of the Franck-Condon (FC) point, the CASSCF-optimized S_1 minima, and the CASSCF-optimized S_1-S_0 minimal energy conical intersections of the 3O cluster as calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory (black, red, and blue bars, respectively). Geometries are shown as insets. (b). Same as (a) except that the excited state minima and MECIs have been optimized at the MS-CASPT2 level of theory..... 263

Figure 8-6. Partially occupied CASSCF natural orbitals of the 3O cluster at the FC point (left) and the closed MECI (right) are presented with occupation numbers. The FC natural orbitals are computed by diagonalization of the S_1 density matrix, while the S_1−S_0 CI natural orbitals are computed from the state-averaged S_1−S_0 density matrix. Two views of each orbital are presented for clarity..... 267

Figure 8-7. Scatter plot of the S_1-S_0 energy gap (a) and S_1 energy (b) of the 3O cluster at all spawning geometries computed at the CASSCF and MS-CASPT2 theoretical levels..... 269

Figure 8-8. Populations of S_1 (bold blue) and S_0 (bold red) of the 5O cluster as a function of time after excitation averaged over all twenty simulations. Thin blue and pink lines report the population corresponding to the twenty individual simulations before averaging..... 271

Figure 8-9. The average S_1-S_0 energy gap of the 5O cluster as a function of time after excitation (bold red) with the energy gaps of the twenty individual initial nuclear basis functions shown in gray. Snapshots from representative trajectories are shown in the inset..... 272

Figure 8-10. (a) Average O-SiOO pyramidalization angle as a function of time (bold lines) with the angles of the twenty individual initial nuclear basis functions (thin lines). The greater angle for each trajectory is shown in red, while the lesser is shown in blue. The inset indicates the atoms that define the pyramidalization angle. (b) The average distance between the epoxide silicon atoms as a function of time (bold red) with the values for the twenty individual initial nuclear basis functions in thin gray lines..... 274

Figure 8-11. (a) The S_0 and S_1 energies of the CASSCF-optimized excited state minimum and S_1-S_0 conical intersections of the 5O cluster as calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory (black, red, and blue bars, respectively). Geometries are shown as insets. Energies at the Franck−Condon (FC) point are shown for comparison. (b) The same as (a) but excited state geometries are optimized at the MS-CASPT2 level of theory..... 275

Figure 8-12. Partially occupied CASSCF natural orbitals of the 5O cluster at the FC point (left), Str S_1-S_0 MECI (middle), and Unpyr S_1-S_0 MECI (right) are presented with occupation numbers. The FC natural orbitals are computed by

diagonalization of the S_1 density matrix, while those of the S_1-S_0 MECIs are computed using the state-averaged S_1-S_0 density matrix..... 279

Figure 8-13. Scatter plot of the S_1-S_0 energy gap (a) and S_1 energy (b) of the 5O cluster at all spawning geometries computed at the CASSCF and MS-CASPT2 theoretical levels..... 281

Figure 9-1. The average S_1-S_0 energy gap as a function of time (bold red) with the energy gaps of the twenty initial nuclear basis functions (gray). Snapshots from representative trajectories are shown in the inset, with silicon, oxygen, and hydrogen atoms colored blue, red, and white, respectively..... 293

Figure 9-2. The population of S_1 (bold blue) and S_0 (bold red) as a function of time after excitation averaged over all twenty initial nuclear basis functions. Thinner blue and pink lines reflect the population corresponding to the twenty individual basis functions before averaging..... 294

Figure 9-3. A histogram of the S_1-S_0 energy gap of the excited population collected over all twenty initial nuclear basis functions integrated over the duration of the simulations..... 295

Figure 9-4. The pyramidalization angle (defined in the text) about the double-bonded Si atom as a function of time, with the average over all excited population shown in red and the values corresponding to the twenty individual initial nuclear basis functions shown in gray..... 296

Figure 9-5. The Si=O bond length as a function of time, with the average in blue and the twenty individual initial nuclear basis functions in gray. Coherent oscillations in the bond length are observed for the entire simulation window..... 298

Figure 9-6. The S_0 and S_1 energies of the CASSCF-optimized excited state minima and minimal energy conical intersections as calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory (black, red, and blue bars, respectively). Geometries are shown as insets. Energies at the Franck-Condon (FC) point are shown for comparison. The S_1-S_0 CI and S_1 min correspond to the pyramidalized region of the PES explored by the AIMS simulations. The oxygen insertion (O-i) intersection and minimum were not observed in the dynamical simulations, but are at lower energies than the pyramidalized minimum..... 300

Figure 9-7. The CASSCF S_0 (black) and S_1 (red) energies as a function of a linearly interpolated/extrapolated coordinate between the S_1-S_0 CI and S_1 min geometries…… 302

Figure 9-8. The S_0 and S_1 energies of MS-CASPT2 optimized minima and intersections, as calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory (black, red, and blue, respectively). Geometries are shown as insets…… 304

Figure 9-9. Partially occupied CASSCF natural orbitals at the FC point (left) and the S_1-S_0 CI (right) are presented with occupation numbers. The FC natural orbitals are computed by diagonalization of the S_1 density matrix, while the S_1-S_0 CI natural orbitals are computed from the state-averaged S_1-S_0 density matrix..... 307

Figure 10-1. Panels A, B, and C show key features of the three CIs discussed in this work. The defects are defined in the first column, the optimized CI structures are shown in the second column, and two different views of the HOMOs and LUMOs involved in the non-radiative transitions at each CI are shown in the third column..... 320

Figure 10-2. Selected geometric parameters at the Franck-Condon (left) and optimized conical intersection (right) structures for each defect. For the Si=O defect we report the pyramidalization angle, defined as the angle between the Si=O bond and the plane defined by the double-bonded silicon atom and the two silicon atoms to which it is bound..... 325

Figure 10-3. Particle size dependence of the CI energy (relative to the ground state minimum energy) of the type 1 epoxide (red), type 2 epoxide (orange), and Si=O (blue) defects..... 327

Figure 10-4. Upper panel: Idealized representation of the PESs of the ground (gray), quantum-confined (green), and defect-localized (red) states of SiNCs of different sizes illustrating how the barrier to localization (crossing between the red and green curves), and thus the rate of recombination, can be strongly size-dependent, though the energy of the defect-localized conical intersection (crossing between the red and gray curves) is not. PESs are represented as a function of a coordinate which locally distorts the nuclear structure of the defect so as to reduce the energy of the defect-localized excited state (e.g. epoxide ring opening or Si=O pyramidalization). Lower panel: Estimated activation barriers for localization to the three defects

studied here, $\Delta G^{\ddagger}$, are plotted as a function of the excitation energy to the quantum-confined state..... 331

Figure 11-1. Population of S_1 (bold blue) and S_0 (bold red) as a function of time after excitation averaged over all 20 initial nuclear basis functions. Thinner blue and pink lines reflect the population corresponding to the 20 individual basis functions before averaging..... 349

Figure 11-2. Average S_1-S_0 energy gap of silicon cluster as a function of time (bold red) with the energy gaps of 20 initial nuclear basis functions (gray). Snapshots from representative trajectories are shown in the inset, with silicon and hydrogen atoms colored blue and white, respectively..... 350

Figure 11-3. (a) Averaged Si-Si bond distance as a function of time (bold red) with bond distance of 19 initial nuclear basis functions (gray), (b) One trajectory that undergoes cluster opening process and no population transfer was observed during the simulation with the snap shots from the dynamics, (c) One trajectory stays near Frank-Condon region in the whole simulation with the snap shots from the dynamics, (d) One trajectory undergoes hydrogen transfer process with the snap shots from the dynamics, (e) indicates the hydrogen transfer process in (d) with the transferred hydrogen shown in red circle. (b),(c) and (d) graphs show the Si-Si bond distance as a function of time (red) as well as S_1 state population as a function of time (blue) for this specific trajectory..... 351

Figure 11-4. Two H-SiSiSi pyramidalization angle as a function of time with snapshots from dynamics and indicating corresponding atoms associated with the pyramidalization angle, H(1) and H(2) are the labels for the two different hydrogen atoms. Si(X), Si(Y) and Si(Z) are three silicon atoms associated with the definition of pyramidalization angle. Noticed that the two hydrogen atoms rotated during the simulation. The graph also shows the S_1 state population as a function of time (pink color)..... 353

Figure 11-5. S_0 and S_1 energies of SA-CASSCF and MS-CASPT2 optimized bond breaking (BD) and square splanar (SqP) minimal energy conical intersections of silicon cluster as calculated at the SA-CASSCF, MS-CASPT2, and MRCI(Q) levels of theory (black, red, and blue bars, respectively). Geometries are shown as insets. Energies at the Franck−Condon (FC) geometry are shown for comparison..... 357

Figure 11-6. (a) The S_1-S_0 energy gap correlation between SA-CASSCF and MRCI(Q)

theoretical levels with 6 orbitals correlated in the MRCI(Q) calculations. (b) The adiabatic S_1 excitation energy correlation between SA-CASSCF and MRCI(Q) theoretical levels..... 358

Figure 11-7. Hole and particle orbitals at IG and FSG are shown in (a) and (b) respectively, with the fist and second columns show the perspective 1 and 2 at IG and FSG, the yellow shade area indicates the local area where re-hybridization happens. The hole and particle orbitals correspond to the HOMO and LUMO SA-CASSCF natural orbitals..... 360

Figure 11-8. (a) The hole and particle orbitals at IG and FSG from with the red circle indicates the anti-bonding interaction between the Si p and H s orbitals. The hole and particle orbitals are the same as those shown in Figure 11-7(c). (b) A schematic graph shows the coordination of the Si atom, these atoms correspond to the yellow shade area as shown in Figure 11-7. The dashed cubic is used as a reference to clarify the positions of each atoms. Red arrow on the left panel shows the direction of rotation of the Si-H bonds. (c) A simplified cartoon to illustrate the anti-bonding character between Si p and H s orbitals at tetrahedral coordination as shown in the left panel. While this anti-bonding interaction does not exist in the right panel which shows the square planer coordination..... 363

Figure 11-9. Partially occupied SA-CASSCF natural orbitals of silicon cluster on two different views at FC geometry (left), BD MECI (middle) and SqP MECI (right) are presented with occupation numbers. All natural orbitals are computed from the state-averaged S_1−S_0 density matrix..... 364

Figure 12-1. Average S_1-S_0 energy gap of silicon cluster as a function of time (bold red) with the energy gaps of 20 initial nuclear basis functions (gray). Snapshots from representative trajectories are shown in the inset, with silicon, hydrogen and oxygen atoms colored blue, white and red, respectively..... 381

Figure 12-2. Population of S_1 (bold blue) and S_0 (bold red) as a function of time after excitation averaged over all 20 initial nuclear basis functions. Thinner blue and pink lines reflect the population corresponding to the 20 individual basis functions before averaging..... 382

Figure 12-3. Average Si-Si bond distance as a function of time(bold red) with the Si-Si bon distance of 20 initial nuclear basis sets(grey), the inset shows the atoms associated with the atom distance concerned in graph, with red, blue and

grey represents oxygen, silicon and hydrogen atoms respectively..... 383

Figure 12-4. Average H(2)OSiH(1) dihedral angle as a function of time(bold red) with the H(2)OSiH(1) dihedral angle of 20 initial nuclear basis sets(grey), the inset shows the atoms associated with the dihedral angle concerned in graph, with red, blue and grey represents oxygen, silicon and hydrogen atoms respectively..... 385

Figure 12-5. Histogram of the S_1-S_0 energy gap of the excited population collected over all 20 initial nuclear basis functions integrated over the duration of the simulations..... 386

Figure 12-6. The S_0 and S_1 energies of three important points on the PES of the silicon hydroxide cluster: the Franck-Condon (FC) point, the S_0/S_1 MECI, and the S_1 min. Black, red and blue lines correspond to the energies of the SA-CASSCF optimized geometries computed at the SA-CASSCF, MS-CASPT2 and MRCISD(Q) levels of theory, respectively. Energies are shifted so that the S_0 energy of the FC point is zero for all levels of theory. Only one black line is visible for the S_0/S_1 MECI because S_0 and S_1 are degenerate. Optimized geometries are shown as insets..... 389

Figure 12-7. The S_0 and S_1 energies of three important points on the PES of the silicon hydroxide cluster: the Franck-Condon (FC) point, the S_0/S_1 MECI, and the S_1 min. Black, red and blue lines correspond to the energies of the MS-CASPT2 optimized geometries computed at the SA-CASSCF, MS-CASPT2 and MRCISD(Q) levels of theory, respectively. Energies are shifted so that the S_0 energy of the FC point is zero for all levels of theory. Only one black line is visible for the S_0/S_1 MECI because S_0 and S_1 are degenerate. Optimized geometries are shown as insets..... 390

Figure 12-8. (a) The S_1-S_0 energy gap correlation between SA-CASSCF and MS-CASPT2 theoretical levels with 7 orbitals correlated in the MS-CASPT2 calculations. (b) The adiabatic S_1 excitation energy correlation between SA-CASSCF and MS-CASPT2 theoretical levels..... 392

Figure 12-9. Partially occupied SA-CASSCF natural orbitals of silicon hydroxide cluster at FC (left) and MECI (right) geometries are presented with occupation numbers. The FC natural orbitals are computed by diagonalization of the S_1 density matrix, while S_1−S_0 MECI natural orbitals are computed from the state-averaged S_1−S_0 density matrix..... 395

Figure 12-10. The S_0/S_1 MECI structure of the SiNC (left) and cluster model (right) as optimized at the CISNOr-CASCI/LANL2DZ level. The HOMOs (top) and LUMOs (bottom) shown correspond to the CISNOr orbitals determined by diagonalization of the state-averaged relaxed CIS density matrix…….. 396

Figure 13-1. PL (a) and FTIR (b) from as-produced SiNCs (blue) and SiNCs refluxed in mesitylene to 160°C (red). The PL spectra are normalized to their respective absorption peaks (not shown), allowing direct comparison of the PL intensities of the two samples…….. 423

Figure 13-2. *In-situ* FTIR spectra (SiH_x region) of heated SiNCs. Purple shows the as-produced sample, with colors moving towards red as the temperature is increased to 550°C, with ΔT = 50°C. Spectra are normalized to the highest point in the displayed region…….. 424

Figure 13-3. EPR signal amplitude (a) and spectral representation (b) from SiNCs as a function of *in-situ* heating time at 125°C. The PL QYs from an SiNC sample heated under the same conditions are shown in (c), and FTIR spectra of an as-produced sample and a sample heated to 125°C is pictured in (d). Error bars in the QY measurement reflect standard deviation…….. 425

Figure 13-4. RGA mass spectra from SiNCs (top) shows removal of SiH_x groups beginning between 100-200°C. These peaks are not evident in the control mass spectra (bottom). Silyl peaks are circled in black…….. 429

Figure 13-5. RGA mass spectra for sample held at 400°C. While silyl peaks (circled in black) are visible at 1 minute, they are reduced after 5 minutes and not visible above the noise after 10 minutes, demonstrating the removal of these groups during heating…….. 430

Figure 13-6. Potential energy surface as a function of silyl bond stretching in the $[Si(SiH_3)_4H]^-$ cluster computed at the QCISD(T) level of theory. The energy is shifted such that it is zero at infinite separation. The insets show the atom labels and optimized Si-Si bond lengths (left) and the orbital responsible for the bonding (right) …….. 432

Figure 13-7. The simulated population of S_0 and S_1 of $[Si(SiH_3)_4H]^-$ as a function of time after excitation. The average over all 20 trajectories is shown in bold, with the populations of the individual trajectories shown by thin lines…….. 434

Figure 13-8. The Si2-Si5 distance (illustrated in the inset) of $[Si(SiH_3)_4H]^-$ as a function of

time after excitation. The average over all 20 trajectories is shown in bold red, with the values corresponding to the individual trajectories shown by thin gray lines..... 435

Figure 13-9. The energy gap between the two lowest many-electron states (S_0 and S_1) of $[Si(SiH_3)_4H]^-$ as a function of time after excitation. The average over all 20 trajectories is shown in bold red, with the values corresponding to the individual trajectories shown by thin gray lines. Insets show a representative geometry at the beginning of the simulation and another at a point of near-zero energy gap..... 436

Figure 13-10. Energies and geometries of three important points on the PES of $[Si(SiH_3)_4H]^-$. S_0 and S_1 energies are shown for all geometries. All structures are optimized at the highly accurate CASPT2 level of theory. CASPT2 energies are presented in red. Single point calculations at the CASSCF level of theory used in our AIMS simulations are presented in black for comparison..... 438

Figure 14-1. Processes involved in photocatalytic overall water splitting on a heterogeneous photocatalyst in a (a) quasi-particle picture, (b) many electron wavefunction picture..... 448

Figure 14-2. (a) Ground state minimum geometry of the model cluster $Ta_2N_{10}H_{20}$, with Ta, N and H represented by yellow, blue and grey colors respectively. The yellow dotted arrow indicates the center of symmetry, the yellow solid arrow shows a sample pair of N atoms with symmetric positions. The red numbers show the Ta-N bond length. (b) The S_1 state has an excitation energy of 4.28eV computed on EOM-CCSD level, the HF orbitals indicate the S_1 character is mostly described by excitations from N p orbitals to Ta d orbitals..... 455

Figure 14-3. Population of S_1 (bold blue) and S_0 (bold red) as a function of time after excitation averaged over all 20 initial nuclear basis functions. Thinner blue and pink lines reflect the population corresponding to the 20 individual basis functions before averaging..... 457

Figure 14-4. Average S_1-S_0 energy gap of the model Ta_3N_5 cluster as a function of time (bold red) with the energy gaps of 20 initial nuclear basis functions (gray). Snapshots from representative trajectories are shown in the inset, with tantalum, nitrogen and hydrogen atoms colored yellow, blue and grey, respectively..... 458

Figure 14-5. Averaged Ta-N bond length as a function of time (bold red) with bond length of 20 initial nuclear basis functions (gray)…… 460

Figure 14-6. (a) Averaged two different N-TaNN pyramidalization angles as a function of time (bold red and blue) with bond length of 20 initial nuclear basis functions (thin red and blue). The corresponding pyramidalization angles are shown in the insets. (b) The atoms defined in the definition of pyramidalization angle…… 461

Figure 14-7. The S_0 and S_1 energies of three important points on the PES of the tantalum nitride model cluster: the Franck-Condon (FC) point, S_0/S_1 MECI and S_1 min optimized on SA-CASSCF theoretical level. Black, red and blue lines correspond to the energies of the single point calculation computed at the SA-CASSCF, MS-CASPT2 and MRCISD(Q) levels of theory, respectively. Energies are shifted so that the S_0 energy of the FC point is zero for all levels of theory. Optimized geometries are shown as insets…… 463

Figure 14-8. Partially occupied SA-CASSCF natural orbitals of silicon hydroxide cluster at FC (left) and MECI (right) geometries are presented with occupation numbers. The FC natural orbitals are computed by diagonalization of the S_1 density matrix, while S_1-S_0 MECI natural orbitals are computed from the state-averaged S_1-S_0 density matrix…… 467

Figure 15-1. Ground state minimum geometry of the CP model cluster $PbCl_6Cs_4$, with Pb, Cl and Cs represented by black, green and brown colors respectively. The blue thin line linked together to show the octahedral structure of $[PbCl_6]^{4-}$. (b) The first four excited states are nearly degenerate with the associated excitation energies and characters calculated at EOM-CCSD theoretical level shown in the graph. Single arrow with the same color indicates the same excited state. The S_1, S_2, S_3, and S_4 are represented with blue, green, orange and gray colors respectively…… 488

Figure 15-2. Average S_1-S_0 energy gap of the model CP cluster as a function of time (bold red) with the energy gaps of 20 initial nuclear basis functions (gray)…… 492

Figure 15-3. (a) Top six panels: six averaged Pb-Cl bond lengths as a function of time(bold red) with twenty individual initial nuclear basis functions shown in gray color. The small blue box in each panel spans the same period of time. (b) Bottom panel: the averaged RMSD of the six Pb-Cl bond lengths as a function of time (bold red) with twenty individual initial nuclear basis functions

shown in gray color…… 493

Figure 15-4. (a) Top twelve panels: twelve averaged Cl-Pb-Cl rectangle angles as a function of time(bold red) with twenty individual initial nuclear basis functions shown in gray color. (b) Bottom left panel indicates the representative atoms associated with a rectangle angle. (c) Bottom right panel: the averaged RMSD of the twelve Cl-Pb-Cl rectangle angles as a function of time (bold red) with twenty individual initial nuclear basis functions shown in gray color…… 495

Figure 15-5. (a) Top three panels: three averaged Cl-Pb-Cl straight angles as a function of time(bold green) with twenty individual initial nuclear basis functions shown in gray color. (b) Bottom left panel indicates the representative atoms associated with a straight angle. (c) Bottom right panel: the averaged RMSD of the three Cl-Pb-Cl straight angles as a function of time (bold red) with twenty individual initial nuclear basis functions shown in gray color…… 497

Figure 15-6. Scanned PES along a symmetric stretch of Pb-Cl bond lengths. The S_1 and S_0 state absolute SA-CASSCF energy PESs are shown as red curves; the S_1-S_0 energy gap is shown as the black curve. The blue dotted line indicates the two ends of the PES well where the right-hand side end corresponds to the FC geometry; the left-hand side end corresponds to the other geometry that has been reached during dynamical simulations…… 499

Figure 15-7. The S_0 and S_1 energies of FC geometry and FC-S_1 min geometries optimized at both SA-CASSCF and MS-CASPT2 theoretical levels. Black and red lines correspond to the energies of the single point calculation computed at the SA-CASSCF and MS-CASPT2 theoretical levels with SBKJC_VDZ and SBKJC_VDZ(p,2d) basis sets with ECPs. Energies are shifted so that the S_0 energy of the FC point is zero for all levels of theory. Optimized geometries are shown as insets…… 501

Figure 15-8. The S_0 and S_1 energies of three important points on the PES of the CP model cluster: the Franck-Condon (FC) point, S_0-S_1 MECI, S_1 min optimized on SA-CASSCF theoretical level. Black and red lines correspond to the energies of the single point calculation computed at the SA-CASSCF and MS-CASPT2 theoretical levels. Energies are shifted so that the S_0 energy of the FC point is zero for all levels of theory. Optimized geometries are shown as insets…… 504

Figure 15-9. The S_0 and S_1 energies of three important points on the PES of the CP model cluster: the Franck-Condon (FC) point, S_0-S_1 MECI, S_1 min optimized on SA-CASSCF theoretical level. Black and red lines correspond to the energies of the single point calculation computed at the SA-CASSCF and MS-CASPT2 theoretical levels. Energies are shifted so that the S_0 energy of the FC point is zero for all levels of theory. Optimized geometries are shown as insets……. 505

Figure 15-10. The S_0 and S_1 energies of FC geometry and Pb-Cl MECI geometries optimized at both SA-CASSCF and MS-CASPT2 theoretical levels. Black and red lines correspond to the energies of the single point calculation computed at the SA-CASSCF and MS-CASPT2 theoretical levels with SBKJC_VDZ and SBKJC_VDZ(p,2d) basis sets with ECPs. Energies are shifted so that the S_0 energy of the FC point is zero for all levels of theory…….. 507

Figure 15-11. Partially occupied SA-CASSCF natural orbitals of silicon hydroxide cluster at FC (left), MECI (middle) and PbCl MECI(right) geometries are presented with occupation numbers. The FC natural orbitals are computed by diagonalization of the S_1 density matrix, while S_1−S_0 MECI natural orbitals are computed from the state-averaged S_1−S_0 density matrix……. 508

Figure 16-1. Band structure at a fixed time t=0, with valence bands, conduction bands and the trapped state shown in black, grey and red respectively. (b) Potential energy surface as a function of nuclear distortion with ground state, quantum confined excited state and defect state shown in black, grey and red respectively. The nuclear distortion is also a function of time. From time t=0 to t=t', it corresponds to a series processes of generated exciton localize to a defect state and go through non-radiative decay to ground state through the conical intersection between the ground state and defect state……. 527

Figure 16-2. An energy level diagram shows the MPE mechanism. The valence potential well, conduction potential well and the trap potential well are shown in black, grey and red respectively. The potential well is defined as a sum of electronic and lattice vibration energies. The potential wells are plotted as a function of lattice coordinate……. 528

Figure 16-3. (a) First excited state of SiNC and SiNC with epoxide defect as a function of different cluster sizes represented by black and red respectively. These two states are representative state for quantum confined delocalized state and

localized defect state. (b) The corresponding excitation characters of SiNC with epoxide defect with 10, 146 and 269 silicon atoms cluster size. It is clear that the excitation characters are localized at the defect. However, as can been seen from the hole of 269 silicon atom cluster, the electronic structure of the hole has a mixing character involving epoxide defect and quantum confined state, this is due to the fact that as the cluster size increases, the excitation energy of quantum confined state reduces, and they are close enough to mix at the 269 silicon atom cluster. This is also the reason why the energy of defect localized state of 269 silicon atom cluster risen a little bit compared to 146 silicon atom cluster…… 531

Figure 16-4. (a) Various MECIs energy levels of model silicon cluster with oxygen defects. With the green arrow points to the specific MECI structure and the grey shadow indicates the localized part where strong nuclear deformation is observed. The red rectangle shade under the energy level diagram corresponds the MECIs that are responsible for experimental observed PL peaks change after the SiNC is oxidized. (b) The PL peaks of pristine SiNC and oxidized SiNC as a function of SiNC size. The red shade close to the energy level of MECIs locate in (a). This explains why the excitons have energy level that is higher that ~2eV (as the red shade indicates) are quenched. (b) is reproduced from the experimental data of paper: Electronic states and luminescence in porous silicon quantum dots: the role of oxygen, MV Wolkin, J Jorne, PM Fauchet, G Allan, C Delerue, Phys. Rev. Lett. 82, 197…….. 535

Figure 16-5. (a). A matrix showing MECI energy levels of $APbX_3$ perovskite material. With each row represents the same type of halide anion but different types of A position cation, each column represents the same type of A position cation but different types of halide anions. The color of the dot represents the energy levels of MECI geometries, and the insets show the MECI geometries. The right hand side color coordinate indicates the energy level in the unit of eV. (b). The ground state minimum geometries of symmetric and non-symmetric conformations of $(NH_3CH_3)_4PbCl_6$ model cluster with two perspectives. The circled $NH_3CH_3^+$ cation indicates the one that has been rotated from symmetric conformation……. 538

Figure 17-1. Illustration of the tree-based genetic algorithm employed in this work…… 556

Figure 17-2. The pictured fragments define the molecular space explored in this work. The X sites on the acceptor groups are allowed to form bonds with the Y sites

on the withdrawing and donor groups..... 561

Figure 17-3. HOMO and LUMO orbitals of the 4CzIPN molecule computed at the CAM-B3LYP/6-31G** level of theory..... 564

Figure 17-4. A scatter plot of the S_1-T_1 energy gap and S_1-S_0 transition dipole moment of the 7518 molecules computed in this study. The color of the points indicates the fitness values of the genes according to the scale on the right. The red dashed line is the Pareto optimal frontier of the data set. The two inset graphs show the same data on different scales for clarity..... 566

Figure 17-5. The average fitness of the pool as a function of generation for series of GA optimizations with (a) identical N_{pool}, but varying P_{mut} and (b) identical P_{mut}, but varying N_{pool}..... 569

Figure 17-6. This scatter plot relates the initial content of the pool to the final content. Three points corresponding to the three most common acceptor groups (benzene, black; pyridine, red; furan, blue) are plotted for each optimization, with the fraction of genes based on that acceptor in the initial pool on the x-axis and the same fraction in the final pool on the y-axis..... 571

Figure 17-7. Matrices showing the numbers of fit molecules (those with $F > 0.8$) containing specific pairs of fragments. (a) shows the number of fit molecules containing the acceptor group on the y-axis and the substituent (donor or withdrawing group) on the x-axis. (b) shows the number of fit molecules containing one copy of each of the substituents on the x- and y-axes. Colored dots guide the eye by reflecting the value of each element, with smaller elements indicated by colder colors and larger elements by warmer colors..... 573

Figure 17-8. This histogram presents the 3792 molecules with $F > 0.8$ binned according to their respective S_1 vertical excitation energies. Insets show representative molecules from left to right having calculated excitations energies of 1.65 eV, 2.79eV, 3.10eV, 3.53eV, 3.77eV, and 4.55 eV..... 576

Figure 18-1. Graphical representation of a molecule with root and six fragments (A, B, A, C, E, B) linked to the root. The $1\sigma_B$, $2\sigma_{BA}$ and $3\sigma_{EBA}$ represent the one-body, two-body and three-body interactions..... 590

Figure 18-2. The root mean square error as a function of the size of the training set for (a) S_1-T_1 energy gap and (b) S_1-S_0 transition dipole moment with a model up to two-body interactions…… 598

Figure 18-3. The graphical represented molecules with different symmetry (a) and (b). However, in the two-body interaction model, they are represented by the same many-body interaction vector…… 600

Figure 18-4. The scatter plot of fitted fitness value and calculated fitness value as shown in black dots for individual molecules. The red line is a fitted line with r^2=0.941…… 601

Figure 18-5. The distribution of the residual between fitted and calculated fitness values…… 602

Figure 18-6. A projected distance matrix of the molecules on the first two principle components with the color represents the fitness value of the molecule…… 604

Figure 18-7. A projected distance matrix of the molecules on the first two principle components with the color represents the fitness value of the molecule. The two zoom in part of the graph is shown with randomly selected molecules and their fitness values…… 605

Chapter 1

Introduction of The Role of Conical Intersection in Optoelectronic Materials

"You can't say A is made of B or vice versa. All mass is interaction."

-Richard P. Feynman(1918-1988)

Portions of this chapter are drawn from Shu, Y. et al. "Nonradiative recombination via conical intersections arising at defects on the oxidized silicon surface," submitted to J. Phys. Chem. C, 119, 1737 (2015)

Optoelectronic devices provide the pathways to convert photon energy to direct current electricity or the corresponding reverse process has been one of the most active research fields in the past decades.[1-7] Due to its wide applications and potentially huge market,[8-10] the research progress have been contributed by scientists from a variety fields in natural science and engineering. Just to name a few, the successfully commercialized products include light emitting diodes (LED)[11-13], solar cells[14-18] etc. For both LED and solar cells, the most important performance metrics are efficiency, stability, and cost.[8-11] Optoelectronic devices can be made of both inorganic materials and organic solids. Inorganic materials usually have better stability, higher efficiency but also harder to

manipulate the properties, and higher cost compared to organic solids.[8,19,20] As efficiency being one of the most important concerns of the optoelectronic materials, a comprehensive understanding of the efficiency limiting processes is required. Non-radiative recombination[21,22] which causes carriers recombine with heat releasing is one main process that limits the efficiency of the optoelectronic materials.[23-30] During my Ph.D. study, we have investigated the mechanism of non-radiative recombination process in various materials with special emphasis on silicon nanocrystal (SiNC) with oxygen defects. To introduce the following chapters with various materials we have studied, I will first briefly present the new paradigm we proposed in understanding the non-radiative recombination process in materials in section 1.1. This section is followed by a detailed introduction of SiNC with oxygen defects as an example in section 1.2. When coming to the investigation of other materials in the later chapters, specific introduction of the material combined with experimental observation and the purpose of our theoretical study will be given at the beginning of the corresponding chapter. Section 1.3 introduces the general topics of each chapter in this dissertation and how they are connected.

1.1 The Role of Conical Intersection In Understanding Non-radiative Decay Process in Optoelectronic Materials

In molecular photochemistry, theory of predicting geometries of molecules that will undergo efficient non-radiative decay has been well established. Over the last several

decades it has been demonstrated that conical intersections (CIs), points of degeneracy between the potential energy surfaces (PESs) of different electronic states, provide pathways for such decay.[31-34] A simplified mathematical explanation will be provided in Chapter 2 section 2.2. In general, a molecule will decay quickly and efficiently to the electronic ground state after excitation if a) a CI is present between the ground and lowest excited state, b) the energy of that CI is below the vertical excitation energy of the lowest electronic excited state, and c) there is not a large energetic barrier in the excited state PES preventing access from the Franck-Condon point.

Defects in nanocrystals (NC) are similar to molecules in that their electronic excitations involve similar amounts of energy (a few eV) and are localized to approximately the same volume (that of a few atoms). Thus, it is reasonable to hypothesize that defects will undergo non-radiative decay via CIs just as molecules do, and that the presence of a CI meeting the above criteria may indicate that a particular defect facilitates non-radiative recombination. For pure NCs, exciton will relax along the excited state PES and go through the nuclear distortion usually involves a few atoms. This is due to the finite temperature ensemble of the NCs sample which prevents the NC being highly symmetric. The non-symmetric distortion of the NC can be treated as defects as well.

Our hypothesis is that the NC will undergo exciton localization process after vertical excitation assume the barrier is not energetically large. If energy levels of CIs introduced by the local nuclear geometrical distortion are accessible by the exciton, the exciton will be quenched efficiently.

1.2 SiNC With Oxygen Defects

In this section, I will introduce one of the main systems I have studied during my Ph.D. years, SiNC with oxygen defects, as an example to show how our hypothesis can be applied to this system to provide a reasonable explanation of the experiment phenomena.

Since the early 1990s when Canham and co-workers discovered the efficient PL of SiNC in visible light range[35], the research field related to Si-based nano-material is intensively active. Due to the low cost, low-toxicity, biological compatibility, applications of SiNC range a variety of regions including light emitting devices[36-38], bioimaging[39-41], photosensitizers[42], and photovoltaics.[43,44] Their tunable PL behavior arises from both quantum confinement effects and the large surface-to-volume ratio of nanocrystals, which allows surface modifications to have dramatic effects on properties.[45] The PL spectrum of nanostructured silicon[35] is one property that shows a strong dependence on surface chemistry. Excited hydrogen-terminated particles emit light with an energy which can be tuned from the bulk silicon band gap of 1.1 eV to approximately 5 eV by varying the nanocrystal size.[46,47] The particles have very high PL quantum yields as synthesized, but significantly lower yields upon oxidation.[48] That defects in semiconductors introduce pathways for non-radiative recombination is well known,[49,50] and thus it is not surprising that defects created during oxidation would decrease the PL yield. However, the exact mechanism by which oxidation introduces non-radiative recombination pathways is not yet clearly established.

Surface oxidation also has a surprising effect on the color of the PL of SiNCs. Upon oxidation, the PL energies of SiNCs with gaps larger than 2.1 eV shift to the red and appear to become insensitive to the average particle size.[46,51] The resulting orange PL is known as the slow (or S-) band PL of nanostructured silicon due to its microsecond lifetime.[52] A higher energy (2.1-2.7 eV) band with a shorter nanosecond lifetime, known as the fast (or F-) band, is also often observed upon oxidation, but will not be discussed further in this thesis.[53-55]

The nature of the excitation responsible for the S-band PL of SiNCs has been the topic of many computational studies, but it remains controversial. It has been suggested that the size-independent S-band is the result of emission from a defect-localized excited state. Perhaps the most frequently cited hypothesis along these lines is that silicon-oxygen double bond (Si=O) defects may be responsible for the emission.[46,56-61] However, as will be discussed in detail in Chapter 9, that simulations by us suggest that excited Si=O defects undergo non-radiative recombination on a timescale considerably faster than the experimentally observed lifetime of the S-band.[62] The stability of Si=O defects is also questionable. Though synthesis of molecules containing such a bond has been reported[63,64], these species are known to be very reactive. Other theoretical works suggest that electronically excited states localized not at Si=O bonds but instead at bridging (Si-O-Si) bonds[65] may be responsible for the unusual PL.

Not all evidence supports the existence of an emissive defect, however. It has been established that the gaps of SiNCs depend on the precise details of the surface

structure even in the complete absence of an emissive defect.[66-72] States delocalized over the oxidized surface of the material have been implicated as the source of the S-band, as well.[73] Several experiments suggest that quantum-confined rather than defect-localized states may be responsible for the S-band emission. The PL lifetimes of SiNCs embedded in SiO_2 vary smoothly from 20 to 200 μs as the diameter is increased from 2.5 nm to 7 nm, suggesting that quantum-confined states are responsible for emission over this entire range.[52] In addition, narrow PL linewidths—consistent with emission from discrete, atom-like quantum-confined states—have been observed in SiNCs of similar size.[74]

The evidence in favor of attributing the S-band to a defect-localized state is hard to reconcile with that in favor of a quantum-confined state, but the proposed hypothesis as has been discussed in the Section1.1 account for both sets of observations. We suggest that the emission occurs from a quantum-confined state and that the unexpected size insensitivity of the PL energy arises due to non-radiative decay processes introduced by CIs whose energy level is size insensitive as will be proved in Chapter 10. A series of works related to SiNC with oxygen defect will be discussed in detail in Chapter 7-12. These works have demonstrated the existence of conical intersections which can be accessed via distortion of oxygen-containing defects on the SiNC surface, viz. Si=O double bonds, silicon epoxide rings (three-membered rings containing two silicon and one oxygen atom) and a combination of silicon epoxide rings and silicon-oxygen bridging bonds (Si-O-Si).[75] To understand how size-insensitive PL may arise from the presence of such

intersections, it is important to recognize that 1) the emission observed in ref[46] is that of ensembles comprising nanocrystals of a range of sizes, and 2) the energies of conical intersections between defect-localized excited states and the ground state are size-insensitive. Thus, when the smaller SiNCs in the ensemble (those with larger band gaps) are excited, they have enough energy to access the conical intersection, and the exciton can, therefore, undergo non-radiatively decay quickly and efficiently. In contrast, larger excited SiNCs—with smaller gaps—do not have enough energy to access the defect-localized excited state or the associated conical intersection and therefore remain in their quantum-confined state long enough to emit. As such, regardless of the average diameter of the SiNCs comprising the ensemble, the emission arises only from the larger SiNCs. Thus the conical intersections introduced by oxidation effectively cap the ensemble PL energy.

1.3 Dissertation Outline

In this section, I will give a brief outline of this dissertation and how these chapters are connected. Generally speaking, this dissertation contains three parts, they are electronic structure theory part, material excited state dynamics part and material optimization part. As has been discussed above, the CI plays an important role in the non-radiative decay processes of the materials. Hence a comprehensive understanding of the non-radiative decay process requires a proper description of the PES in the vicinity of CI. This motivates the electronic structure theory part which introduces the ideas of

how to correctly describe the CI region of the PES and new methodologies developed in our group. The electronic structure theory part includes Chapter 2-6. Chapter 2 gives a brief introduction of the electronic structure theory, special emphasis will be put on describing the CI region of the PES. A brief introduction of how to solve the Time-Dependent Schrödinger Equation to propagate the nuclear dynamics will be discussed as well. Chapter 3-5 introduces the new electronic structure theories that we have developed with various advantages toward the current state of the art method of calculating the PES near the CI region. Chapter 6 closes the first part with some benchmark studies and possible future directions to go.

With the help of existing method and those methodologies developed in our group that are introduced in the electronic structure theory part, we are able to model possible pathways of the non-radiative decay processes of the materials from small model clusters with a couple of heavy atoms to the nano-sized cluster. These application works constitute the second part of this thesis, material excited state dynamics part, including Chapter 7-16. In this part, I will present a couple of cases we have studied with the goal of understanding the non-radiative recombination mechanisms. Chapter 7-11 show investigations of SiNC with different oxygen containing defects. Chapter 12 studies the pristine SiNC capped with H atoms for comparison. Chapter 13 shows a combined theoretical and experimental study of SiNC with hypervalently bonded Silyl group to illustrate the success of our hypothesis further. In Chapter 14 and Chapter 15, I will present two other systems we have investigated, tantalum nitride (Ta3N5) which is a

promising photocatalyst for water splitting and lead halide perovskite (AMX_3) material with organic and inorganic A position cations. The second part is closed by Chapter 16, in which I will summarize the application works we have done and how our hypothesis is successfully applied to understand the non-radiative recombination mechanisms of optoelectronic materials.

Understanding the non-radiative decay pathways of the materials helps us design the next generation efficient optoelectronic materials. This motivates the development of automated optimization strategy with the help of which, large molecular configuration space can be efficiently searched and hence materials with desired properties can be designed. This is described in the material optimization part of the thesis including Chapter 17-18. This part shows how we have applied large scale computational screening and an efficient optimization algorithm to optimize materials toward certain properties. Chapter 17 will introduce the algorithm and an example to show how this algorithm works. Chapter 18 will discuss some machine learning based studies to show how we utilize our database to build an efficient model to calculate physical and chemical properties of the molecules.

REFERENCES

REFERENCES

(1) Meredith, P.; Bettinger, C. J.; Irimia-Vladu, M.; Mostert, A. B.; Schwenn, P. E. *Rep Prog Phys* **2013**, *76*.

(2) Scaccabarozzi, A. D.; Stingelin, N. *J Mater Chem A* **2014**, *2*, 10818.

(3) Li, Y.; Qian, F.; Xiang, J.; Lieber, C. M. *Mater Today* **2006**, *9*, 18.

(4) Rakic, A. D.; Djurisic, A. B.; Elazar, J. M.; Majewski, M. L. *Appl Optics* **1998**, *37*, 5271.

(5) Auston, D. H. *Appl Phys Lett* **1975**, *26*, 101.

(6) Forrest, S. R.; Thompson, M. E. *Chem Rev* **2007**, *107*, 923.

(7) Duan, X. F.; Huang, Y.; Cui, Y.; Wang, J. F.; Lieber, C. M. *Nature* **2001**, *409*, 66.

(8) Brabec, C. J.; Hauch, J. A.; Schilinsky, P.; Waldauf, C. *Mrs Bull* **2005**, *30*, 50.

(9) Coe-Sullivan, S. *Nat Photonics* **2009**, *3*, 315.

(10) Barr, J. Evaluation of the commercial potential of novel organic photovoltaic technologies. M Eng, Massachusetts Institute of Technology, 2005.

(11) Pimputkar, S.; Speck, J. S.; DenBaars, S. P.; Nakamura, S. *Nat Photonics* **2009**, *3*, 179.

(12) Ye, S.; Xiao, F.; Pan, Y. X.; Ma, Y. Y.; Zhang, Q. Y. *Mat Sci Eng R* **2010**, *71*, 1.

(13) Zhao, H. P.; Liu, G. Y.; Zhang, J.; Poplawsky, J. D.; Dierolf, V.; Tansu, N. *Opt Express* **2011**, *19*, A991.

(14) Gunes, S.; Neugebauer, H.; Sariciftci, N. S. *Chem Rev* **2007**, *107*, 1324.

(15) Li, G.; Zhu, R.; Yang, Y. *Nat Photonics* **2012**, *6*, 153.

(16) Narayan, M. R. *Renew Sust Energ Rev* **2012**, *16*, 208.

(17) Lunt, R. R.; Bulovic, V. *Appl Phys Lett* **2011**, *98*.

(18) Zhao, Y. M.; Meek, G. A.; Levine, B. G.; Lunt, R. R. *Adv Opt Mater* **2014**, *2*, 606.

(19) Brabec, C. J. *Sol Energ Mat Sol C* **2004**, *83*, 273.

(20) Nguyen, T. P. *Surf Coat Tech* **2011**, *206*, 742.

(21) Teller, E. *J Phys Chem-Us* **1937**, *41*, 109.

(22) Kilina, S.; Kilin, D.; Tretiak, S. *Chem Rev* **2015**, *115*, 5929.

(23) Galland, C.; Ghosh, Y.; Steinbruck, A.; Sykora, M.; Hollingsworth, J. A.; Klimov, V. I.; Htoon, H. *Nature* **2011**, *479*, 203.

(24) Kamat, P. V. *Accounts Chem Res* **2012**, *45*, 1906.

(25) Rao, A.; Chow, P. C. Y.; Gelinas, S.; Schlenker, C. W.; Li, C. Z.; Yip, H. L.; Jen, A. K. Y.; Ginger, D. S.; Friend, R. H. *Nature* **2013**, *500*, 435.

(26) Gabriel, M. M.; Grumstrup, E. M.; Kirschbrown, J. R.; Pinion, C. W.; Christesen, J. D.; Zigler, D. F.; Cating, E. E. M.; Cahoon, J. F.; Papanikolas, J. M. *Nano Lett* **2014**, *14*, 3079.

(27) Brawand, N. P.; Voros, M.; Galli, G. *Nanoscale* **2015**, *7*, 3737.

(28) Najafi, E.; Scarborough, T. D.; Tang, J.; Zewail, A. *Science* **2015**, *347*, 164.

(29) Wang, L. J.; Long, R.; Prezhdo, O. V. *Annu Rev Phys Chem* **2015**, *66*, 549.

(30) Wang, H. N.; Zhang, C. J.; Rana, F. *Nano Lett* **2015**, *15*, 339.

(31) Bernardi, F.; Olivucci, M.; Robb, M. A. *Chemical Society Reviews* **1996**, *25*, 321.

(32) Yarkony, D. R. *Reviews of Modern Physics* **1996**, *68*, 985.

(33) Levine, B. G.; Martinez, T. J. *Annual Review of Physical Chemistry* **2007**, *58*, 613.

(34) Domcke, W.; Yarkony, D. R. *Annu. Rev. Phys. Chem.* **2012**, *63*, 325.

(35) Canham, L. T. *Applied Physics Letters* **1990**, *57*, 1046.

(36) Erdem, T.; Demir, H. V. *Nature Photonics* **2011**, *5*, 126.

(37) Canham, L. *Nature* **2000**, *408*, 411.

(38) Walters, R. J.; Bourianoff, G. I.; Atwater, H. A. *Nature Materials* **2005**, *4*, 143.

(39) Cui, Y.; Wei, Q. Q.; Park, H. K.; Lieber, C. M. *Science* **2001**, *293*, 1289.

(40) He, Y.; Zhong, Y. L.; Peng, F.; Wei, X. P.; Su, Y. Y.; Lu, Y. M.; Su, S.; Gu, W.; Liao, L. S.; Lee, S. T. *Journal of the American Chemical Society* **2011**, *133*, 14192.

(41) Erogbogbo, F.; Yong, K. T.; Roy, I.; Hu, R.; Law, W. C.; Zhao, W. W.; Ding, H.; Wu, F.; Kumar, R.; Swihart, M. T.; Prasad, P. N. *Acs Nano* **2011**, *5*, 413.

(42) Pikulev, V. B.; Kuznetsov, S. N.; Saren, A. A.; Gardin, Y. E.; Gurtov, V. A. *Physica Status Solidi a-Applications and Materials Science* **2007**, *204*, 1266.

(43) Perez-Wurfl, I.; Hao, X. J.; Gentle, A.; Kim, D. H.; Conibeer, G.; Green, M. A. *Appl Phys Lett* **2009**, *95*.

(44) Kim, S.; Jeon, K.; Lee, J. C.; Swihart, M. T.; Yang, M. *Appl Phys Express* **2012**, *5*.

(45) Dohnalova, K.; Gregorkiewicz, T.; Kusova, K. *J Phys-Condens Mat* **2014**, *26*.

(46) Wolkin, M. V.; Jorne, J.; Fauchet, P. M.; Allan, G.; Delerue, C. *Physical Review Letters* **1999**, *82*, 197.

(47) Reboredo, F. A.; Franceschetti, A.; Zunger, A. *Physical Review B* **2000**, *61*, 13073.

(48) Jurbergs, D.; Rogojina, E.; Mangolini, L.; Kortshagen, U. *Appl Phys Lett* **2006**, *88*.

(49) Hall, R. N. *Physical Review* **1952**, *87*, 387.

(50) Shockley, W.; Read, W. T. *Physical Review* **1952**, *87*, 835.

(51) von Behren, J.; van Buuren, T.; Zacharias, M.; Chimowitz, E. H.; Fauchet, P. M. *Solid State Communications* **1998**, *105*, 317.

(52) Garcia, C.; Garrido, B.; Pellegrino, P.; Ferre, R.; Moreno, J. A.; Morante, J. R.; Pavesi, L.; Cazzanelli, M. Applied Physics Letters 2003, 82, 1595.

(53) Andrianov, A. V.; Kovalev, D. I.; Shuman, V. B.; Yaroshetskii, I. D. JETP Lett. 1992, 56, 236.

(54) Ito, T.; Ohta, T.; Hiraki, A. *Jpn. J. Appl. Phys. 2* **1992**, *31*, L1.

(55) Calcott, P. D. J.; Nash, K. J.; Canham, L. T.; Kane, M. J.; Brumhead, D. J. Phys.: Condens. Matter 1993, 5, L91.

(56)Puzder, A.; Williamson, A. J.; Grossman, J. C.; Galli, G. Journal of Chemical Physics 2002, 117, 6721.

(57) Vasiliev, I.; Chelikowsky, J. R.; Martin, R. M. *Physical Review B* **2002**, *65*.

(58) Puzder, A.; Williamson, A. J.; Grossman, J. C.; Galli, G. Journal of the American Chemical Society 2003, 125, 2786.

(59) Luppi, M.; Ossicini, S. *Physical Review B* **2005**, *71*.

(60) Chen, X. B.; Pi, X. D.; Yang, D. R. *Journal of Physical Chemistry C* **2010**, *114*, 8774.

(61) Pennycook, T. J.; Hadjisavvas, G.; Idrobo, J. C.; Kelires, P. C.; Pantelides, S. T. Physical Review B 2010, 82.

(62) Shu, Y.; Levine, B. G. *J. Phys. Chem. C* **2014**, *118*, 7669.

(63) Xiong, Y.; Yao, S. L.; Driess, M. *J. Am. Chem. Soc.* **2009**, *131*, 7562.

(64) Filippou, A. C.; Baars, B.; Chernov, O.; Lebedev, Y. N.; Schnakenburg, G. Angew. Chem.-Int. Edit. 2014, 53, 565.

(65) Degoli, E.; Guerra, R.; Iori, F.; Magri, R.; Marri, I.; Pulci, O.; Bisi, O.; Ossicini, S. C. R. Phys. 2009, 10, 575.

(66) Daldosso, N.; Luppi, M.; Ossicini, S.; Degoli, E.; Magri, R.; Dalba, G.; Fornasini, P.; Grisenti, R.; Rocca, F.; Pavesi, L.; Boninelli, S.; Priolo, F.; Spinella, C.; Iacona, F. Physical Review B 2003, 68.

(67) Ramos, L. E.; Furthmuller, J.; Bechstedt, F. *Physical Review B* **2004**, *70*.

(68) Ramos, L. E.; Furthmuller, J.; Bechstedt, F. *Physical Review B* **2005**, *71*.

(69) Guerra, R.; Degoli, E.; Ossicini, S. *Physical Review B* **2009**, *80*.

(70) Konig, D.; Rudd, J.; Green, M. A.; Conibeer, G. Solar Energy Materials and Solar Cells 2009, 93, 753.

(71) Seino, K.; Bechstedt, F.; Kroll, P. *Nanotechnology* **2009**, *20*.

(72) Li, T. S.; Gygi, F.; Galli, G. Physical *Review Letters* **2011**, *107*.

(73) Zhou, Z. Y.; Brus, L.; Friesner, R. *Nano Lett.* **2003**, *3*, 163.

(74) Sychugov, I.; Juhasz, R.; Valenta, J.; Linnros, J. *Phys. Rev. Lett.* **2005**, *94*.

(75) Shu, Y. N.; Levine, B. G. *J Phys Chem C* **2015**, *119*, 1737.

Chapter 2

Introduction of Electronic Structure Theories and Propagation of Dynamics

"There are no secrets about the world of nature. There are secrets about the thoughts and intentions of men."

-J. Robert Oppenheimer(1904-1967)

Theoretical quantum chemistry as a field of applying quantum mechanics to chemical systems, its main goal is to solve Schrödinger equation to obtain the electronic structures of many electron systems which determine the chemical and physical properties of the matter. The birth of quantum chemistry is dated back to 1927 when Heitler and London[1] investigated covalent bond in the hydrogen molecule. As most of the chemical systems have more than one electron, obtaining the exact solution of the Schrödinger equation is generally impossible.

The first step towards the numerical results, quantum chemistry has largely relied on the Born-Oppenheimer approximation[2] which states that the motion of nuclei and electrons can be separated. As a result, multilevel approximations were made in order to solve the electronic Schrödinger equation numerically, among which, the most recognized ansatz including Hartree-Fock, truncated configuration interaction, many-body

perturbation, coupled-cluster and complete active space methods.[3-5] While it is the goal of the existing theory to explain and match with the experimentally observed phenomena such as matter electromagnetic radiation interaction, which is one of the most widely applied experimental methodologies. Resolving these spectroscopic properties of matters requires the knowledge of energy differences as well as transition probabilities between ground and excited states.[6] Thus investigating the matter electromagnetic radiation interaction as well as how matter behaves after absorbing the electromagnetic radiation is of particular importance.

As a main project during my Ph.D. study, understanding the mechanism of the non-radiative decay process which explained by E. Teller as "the molecule can get from a higher potential surface to a lower one, transforming in this way electronic excitation energy into kinetic energy of atoms and finally into heat"[7], requires the knowledge of multistate potential energy surfaces (PESs) as well as going beyond the Born-Oppenheimer approximation. This chapter will be organized as the following, first I will introduce the adiabatic and Born-Oppenheimer approximations, followed by the rising of the necessity to go beyond the Born-Oppenheimer approximation for the purpose of studying non-radiative decay process. Second, I will introduce the importance of employing multi-configurational electronic structure theory to obtain a qualitatively correct description of the PESs at regions where non-radiative decay happens. Last, algorithm, full multiple spawning algorithm in specific, to solve the time-dependent Schrödinger equation will be introduced briefly.

2.1 From Exact Theory to Born-Oppenheimer Approximation

The total Hamiltonian of the molecule has the form,

$$\hat{H} = \hat{T}_N + \hat{H}_e \tag{2-1}$$

where $\hat{T}_N = -\frac{\hbar^2}{2\mu}\Delta_{\mathbf{R}}$ is the nuclear kinetic energy operator, $\Delta_{\mathbf{R}} = \frac{\partial^2}{\partial \mathbf{R}^2}$ is the Laplacian. $\hat{H}_e = -\frac{\hbar^2}{2m}\Delta_{\mathbf{r}} + \hat{V}$ is the electronic Hamiltonian operator contains electronic kinetic energy as well as all the columbic interactions between nuclei and electrons $\hat{V}$.

In exact non-adiabatic theory, the total wavefunction is expanded as linear combination of electronic and nuclear wavefunctions in the space of infinite numbers of electronic states ($N = \infty$).

$$\Psi(\mathbf{r},\mathbf{R}) = \sum_i^N \psi_i(\mathbf{r};\mathbf{R})\chi_i(\mathbf{R}) \tag{2-2}$$

where $\psi_i(\mathbf{r};\mathbf{R})$ is the electronic wavefunction as eigenfunctions of $\hat{H}_e$,

$$\hat{H}_e\psi_i(\mathbf{r};\mathbf{R}) = \varepsilon_i\psi_i(\mathbf{r};\mathbf{R}) \tag{2-3}$$

and $\chi_i(\mathbf{R})$ is the nuclear wavefunction. Operating the Hamiltonian on the exact wavefunction gives,

$$\left(\hat{T}_N + \hat{H}_e\right)\sum_i^N \psi_i(\mathbf{r};\mathbf{R})\chi_i(\mathbf{R}) = E\sum_i^N \psi_i(\mathbf{r};\mathbf{R})\chi_i(\mathbf{R}) \tag{2-4}$$

multiply both sides by $\psi_j^*(\mathbf{r};\mathbf{R})$ and integrate over electronic coordinates $\mathbf{r}$, and ignore the coordinates in the parentheses for simplicity,

$$\sum_i^N \left\langle \psi_j \middle| \hat{T}_N + \hat{H}_e \middle| \psi_i\chi_i \right\rangle_{\mathbf{r}} = E\sum_i^N \left\langle \psi_j \middle| \psi_i \right\rangle_{\mathbf{r}} \chi_i \tag{2-5}$$

due to the orthonormalization condition $\left\langle \psi_j \middle| \psi_i \right\rangle_{\mathbf{r}} = \delta_{ij}$,

$$\varepsilon_j \chi_j + \sum_i^N \left\langle \psi_j \middle| \hat{T}_N \middle| \psi_i \chi_i \right\rangle_{\mathbf{r}} = E \chi_j \tag{2-6}$$

with

$$\begin{aligned} \left\langle \psi_j \middle| \hat{T}_N \middle| \psi_i \chi_i \right\rangle_{\mathbf{r}} = \left(1-\delta_{ij}\right)\left(-\frac{\hbar^2}{\mu}\right)\left\langle \psi_j \middle| \frac{\partial}{\partial \mathbf{R}} \middle| \psi_i \right\rangle_{\mathbf{r}} \frac{\partial}{\partial \mathbf{R}} \chi_i - \delta_{ij} \frac{\hbar^2}{2\mu} \frac{\partial^2}{\partial \mathbf{R}^2} \chi_i \\ + \left\langle \psi_j \middle| \hat{T}_N \middle| \psi_i \right\rangle_{\mathbf{r}} \chi_i \end{aligned} \tag{2-7}$$

reorganization of the equations gives,

$$\left[-\frac{\hbar^2}{2\mu} \Delta_{\mathbf{R}} + \varepsilon_j + \left\langle \psi_j \middle| \hat{T}_N \middle| \psi_j \right\rangle_{\mathbf{r}} - E \right] \chi_j = -\sum_{i(i \neq j)}^N \Theta_{ij} \chi_i \tag{2-8}$$

with

$$\Theta_{ij} = -\frac{\hbar^2}{\mu} \left\langle \psi_j \middle| \frac{\partial}{\partial \mathbf{R}} \middle| \psi_i \right\rangle_{\mathbf{r}} \frac{\partial}{\partial \mathbf{R}} + \left\langle \psi_j \middle| \hat{T}_N \middle| \psi_i \right\rangle_{\mathbf{r}} \tag{2-9}$$

Adiabatic approximation ignores the Θ_{ij}, and the Schrödinger equation reduces to,

$$\left[-\frac{\hbar^2}{2\mu} \Delta_{\mathbf{R}} + \varepsilon_j + \left\langle \psi_j \middle| \hat{T}_N \middle| \psi_j \right\rangle_{\mathbf{r}} \right] \chi_j = E \chi_j \tag{2-10}$$

Born-Oppenheimer approximation further treats $\left\langle \psi_j \middle| \hat{T}_N \middle| \psi_j \right\rangle_{\mathbf{r}}$ as 0, and ends up with an even simpler form of the Schrödinger equation,

$$\left[-\frac{\hbar^2}{2\mu} \Delta_{\mathbf{R}} + \varepsilon_j \right] \chi_j = E \chi_j \tag{2-11}$$

It is evident from the previous derivation that under adiabatic and Born-Oppenheimer approximations, the nuclear wavefunctions propagate on one potential energy surface, there is no term accounts for the electronic transition among multiple electronic states. Thus to investigate the non-radiative decay processes which involve

electronic population transfer among multiple electronic states, we have to go beyond the Born-Oppenheimer and adiabatic approximation[①].

2.2 Conical Intersection Facilitates Non-radiative Decay

Now let us consider the time-dependent Schrödinger equation for electrons only and at the moment, treat nuclei as classical particles. The time-dependent Schrödinger equation,

$$i\hbar\frac{\partial\Psi\left(\mathbf{r},\mathbf{R},t\right)}{\partial t}=\hat{H}\Psi\left(\mathbf{r},\mathbf{R},t\right) \tag{2-12}$$

with the total wavefunction expanded as,

$$\Psi\left(\mathbf{r},\mathbf{R},t\right)=\sum_{i}^{N}c_{i}\left(t\right)\psi_{i}\left(\mathbf{r};\mathbf{R}\right) \tag{2-13}$$

Insert the expansion of total wavefunction into the time-dependent Schrödinger equation and multiply on the left with $\psi_{j}\left(\mathbf{r};\mathbf{R}\right)$, integrate over electronic coordinates,

$$i\hbar\dot{c}_{j}\left(t\right)=\sum_{i}^{N}c_{i}\left(t\right)\left(H_{ji}-i\hbar\dot{\mathbf{R}}\bullet\mathbf{d}_{ji}\right) \tag{2-14}$$

with,

$$H_{ji}=\left\langle\psi_{j}\left(\mathbf{r};\mathbf{R}\right)\middle|\hat{H}_{e}\middle|\psi_{i}\left(\mathbf{r};\mathbf{R}\right)\right\rangle_{\mathbf{r}} \tag{2-15}$$

$$d_{ji}=\left\langle\psi_{j}\left(\mathbf{r};\mathbf{R}\right)\middle|\frac{\partial}{\partial\mathbf{R}}\middle|\psi_{i}\left(\mathbf{r};\mathbf{R}\right)\right\rangle_{\mathbf{r}} \tag{2-16}$$

[①] It is important to point out that in literature, people usually do not distinguish between the Born-Oppenheimer approximation and adiabatic approximation. When people say Born-Oppenheimer approximation, usually mean both Born-Oppenheimer approximation and adiabatic approximation. In the following text, we will only use the terminology Born-Oppenheimer approximation unless specifically mentioned.

The derivative coupling d_{ji} can be expressed further in terms of energies of adiabatic states as,

$$d_{ji} = \left\langle \psi_j(\mathbf{r};\mathbf{R}) \left| \frac{\partial}{\partial \mathbf{R}} \right| \psi_i(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} = \frac{\left\langle \psi_j(\mathbf{r};\mathbf{R}) \left| \frac{\partial \hat{H}_e}{\partial \mathbf{R}} \right| \psi_i(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}}}{E_i - E_j} \quad (2\text{-}17)$$

The derivative coupling is thus inversely proportional to the energy difference between the two electronic states. As can be seen from equation (2-14), the change of the population for the specific state is propagated by the derivative coupling and nuclear velocity when adiabatic representation is employed. Hence, it is justified from equation (2-14) and (2-17) that smaller energy gaps between two different adiabatic states facilitates efficient population transfer between the two states. The crosses among the adiabatic PESs are called conical intersections (without rigorous definition), these points form an N-2 dimension seam[8-12], with N represents the nuclear degrees of freedom.

2.3 Proper Description of The Electronic Structures At Conical Intersection Geometries

As seen from the previous section, conical intersection facilitates the non-radiative decay processes. It is thus important to properly locate the conical intersection geometry, at which, associated energy gap of the two adiabatic states is zero. Let us start with two state model Hamiltonian[13],

$$H(\mathbf{R}) = \begin{bmatrix} H_{11}(\mathbf{R}) & H_{12}(\mathbf{R}) \\ H_{21}(\mathbf{R}) & H_{22}(\mathbf{R}) \end{bmatrix} \quad (2\text{-}18)$$

with $H_{ji}(\mathbf{R}) = \left\langle \varphi_j(\mathbf{r};\mathbf{R}) \left| \hat{H}_e \right| \varphi_i(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}}$. $\varphi_i(\mathbf{r};\mathbf{R})$ and $\varphi_j(\mathbf{r};\mathbf{R})$ are not adiabatic

states. The eigenvalues are,

$$E_{\pm} = \frac{H_{11}(\mathbf{R}) + H_{22}(\mathbf{R}) \pm \sqrt{\Delta H(\mathbf{R})^2 + H_{12}(\mathbf{R})^2}}{2} \tag{2-19}$$

with $\Delta H(\mathbf{R}) = H_{11}(\mathbf{R}) - H_{22}(\mathbf{R})$. Moreover, zero energy gap means $E_+ = E_-$, which ends up with conditions for locating conical intersections,

$$H_{11}(\mathbf{R}) = H_{22}(\mathbf{R}) \tag{2-20}$$

$$H_{12}(\mathbf{R}) = 0 \tag{2-21}$$

This also explains why the conical intersection dimensionality is N-2.

If we expand the Hamiltonian matrix elements near the conical intersection in a Taylor series and truncate it as first order,

$$H(\mathbf{R}) = H(\mathbf{R_{CI}}) + \nabla H(\mathbf{R_{CI}}) \bullet \delta\mathbf{R} \tag{2-22}$$

where $\mathbf{R_{CI}}$ represents the conical intersection geometry. Then the conditions for locating conical intersection as shown in eq. (2-20) and eq. (2-21) becomes,

$$\Delta H(\mathbf{R_{CI}}) = 0 + \nabla \Delta H(\mathbf{R_{CI}}) \bullet \delta\mathbf{R} = 0 \tag{2-23}$$

$$H_{12}(\mathbf{R_{CI}}) = 0 + \nabla H_{12}(\mathbf{R_{CI}}) \bullet \delta\mathbf{R} = 0 \tag{2-24}$$

And we define,

$$\mathbf{g} = \nabla \Delta H(\mathbf{R_{CI}}) \tag{2-25}$$

$$\mathbf{h} = \nabla H_{12}(\mathbf{R_{CI}}) \tag{2-26}$$

Hence the associated two vectors that linearly lift the gap between the two eigenvalues (under the first order Taylor expansion approximation) are termed as g and h vectors with the following definition,

$$g^{ij} = \left\langle \psi_i(\mathbf{r};\mathbf{R}) \left| \frac{\partial \hat{H}_e}{\partial \mathbf{R}} \right| \psi_i(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} - \left\langle \psi_j(\mathbf{r};\mathbf{R}) \left| \frac{\partial \hat{H}_e}{\partial \mathbf{R}} \right| \psi_j(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} \quad (2\text{-}27)$$

$$h^{ij} = \left\langle \psi_i(\mathbf{r};\mathbf{R}) \left| \frac{\partial \hat{H}_e}{\partial \mathbf{R}} \right| \psi_j(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} \quad (2\text{-}28)$$

A Proper description of the conical intersection, requires the electronic structure theories applied to calculate the PES has the capability to capture all these features.

The widely used and conceptually simple ways to get excited state information are configuration interaction singles (CIS) and its associated density functional theory (DFT) version, named, time-dependent DFT (TDDFT)[14,15] with Tamm-Dancoff approximation (TDA)[16]. However, it is pointed out by Levine et al.[17] that neither CIS based on Hartree-Fock (HF) orbitals nor TDDFT with TDA are not suitable in describing the conical intersection region of the PES in a qualitative sense. The reason can be explained in the fact that the h vector is zero at the conical intersection geometry at both theoretical levels. The h vector at conical intersection geometry is further derived and re-written as shown in eq. (2-29),

$$\begin{aligned} h^{ij} &= \left\langle \psi_i(\mathbf{r};\mathbf{R}) \left| \frac{\partial \hat{H}_e}{\partial \mathbf{R}} \right| \psi_j(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} \\ &= \left\langle \frac{\partial \psi_i(\mathbf{r};\mathbf{R})}{\partial \mathbf{R}} \middle| \hat{H}_e \middle| \psi_j(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} + \left\langle \psi_i(\mathbf{r};\mathbf{R}) \middle| \frac{\partial \hat{H}_e}{\partial \mathbf{R}} \middle| \psi_j(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} \\ &\quad + \left\langle \psi_i(\mathbf{r};\mathbf{R}) \middle| \hat{H}_e \middle| \frac{\partial \psi_j(\mathbf{r};\mathbf{R})}{\partial \mathbf{R}} \right\rangle_{\mathbf{r}} \\ &= \frac{\partial \left\langle \psi_i(\mathbf{r};\mathbf{R}) \middle| \hat{H}_e \middle| \psi_j(\mathbf{r};\mathbf{R}) \right\rangle}{\partial \mathbf{R}}_{\mathbf{r}} \end{aligned} \quad (2\text{-}29)$$

The second equality holds because of eq. (2-30),

$$
\begin{aligned}
&\left\langle \frac{\partial \psi_i(\mathbf{r};\mathbf{R})}{\partial \mathbf{R}} \middle| \hat{H}_e \middle| \psi_j(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} + \left\langle \psi_i(\mathbf{r};\mathbf{R}) \middle| \hat{H}_e \middle| \frac{\partial \psi_j(\mathbf{r};\mathbf{R})}{\partial \mathbf{R}} \right\rangle_{\mathbf{r}} \\
&= \left\langle \frac{\partial \psi_i(\mathbf{r};\mathbf{R})}{\partial \mathbf{R}} \middle| \varepsilon_j \psi_j(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} + \left\langle \varepsilon_i \psi_i(\mathbf{r};\mathbf{R}) \middle| \frac{\partial \psi_j(\mathbf{r};\mathbf{R})}{\partial \mathbf{R}} \right\rangle_{\mathbf{r}} \\
&\Downarrow \text{at CI, states are degenerate and hence } \varepsilon_j = \varepsilon_i = \varepsilon \\
&= \varepsilon \left[\left\langle \frac{\partial \psi_i(\mathbf{r};\mathbf{R})}{\partial \mathbf{R}} \middle| \psi_j(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} + \left\langle \psi_i(\mathbf{r};\mathbf{R}) \middle| \frac{\partial \psi_j(\mathbf{r};\mathbf{R})}{\partial \mathbf{R}} \right\rangle_{\mathbf{r}} \right] \\
&= \varepsilon \frac{\partial \left\langle \psi_i(\mathbf{r};\mathbf{R}) \middle| \psi_j(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}}}{\partial \mathbf{R}} \\
&\Downarrow \text{states are orthorgonal} \\
&= 0
\end{aligned}
\tag{2-30}
$$

and due to the Brillouin's theorem[4], $\left\langle \psi_{HF}(\mathbf{r};\mathbf{R}) \middle| \hat{H}_e \middle| \psi_i^a(\mathbf{r};\mathbf{R}) \right\rangle = 0$. It ends up with $h^{0j} = 0$ at CIS level. Hence it is concluded that doubly excited determinants are necessary in order to capture a qualitatively correct description of the PES near conical intersection geometry between ground and excited states. However, it has to be emphasized here that CIS or TDDFT with TDA are suitable methods in describing the conical intersections between excited states.

It is the state-averaged complete active space self-consistent field (SA-CASSCF)[18] method has long been the state of the art electronic structure theory in describing the conical intersection regions of the PES as the zeroth order approximation of the wavefunction. SA-CASSCF has a good balance between computational cost and accuracy with the possibility of computing more than 20 heavy atoms system in a reasonable computational time. SA-CASSCF and its analytical gradient, analytical implementation of non-adiabatic coupling vector are also widely available in current electronic structure theory codes.[19-23] However, SA-CASSCF also exists undesirable behaviors, for example,

the accuracy depends on the user chosen active space, convergence problems, PES discontinuities,[24-26] unphysical symmetry breaking of the wavefunctions,[27] and not being size intensive.[28] Some of these problems are arising from the fact that orbital coefficients and configuration interaction vector are simultaneously optimized in SA-CASSCF. In the following chapters, I will introduce the multi-reference electronic structure theories developed in our group for the purpose of solving or alleviating the problems in SA-CASSCF.

2.4 Propagate The Wavefunctions By Solving Time-Dependent Schrödinger Equation

Now let us get back to the time-dependent Schrödinger equation,

$$i\hbar\frac{\partial\Psi\left(\mathbf{r},\mathbf{R},t\right)}{\partial t}=\hat{H}\Psi\left(\mathbf{r},\mathbf{R},t\right) \tag{2-31}$$

the full expansion of the wavefunction can be written as,

$$\Psi\left(\mathbf{r},\mathbf{R},t\right)=\sum_{J}^{N}\psi_{J}\left(\mathbf{r};\mathbf{R}\right)\chi_{J}\left(\mathbf{R},t\right) \tag{2-32}$$

with similar notation as equation (2-12) and (2-2). The approximations of the total wavefunction expansion can generally be dived into three categories, one treats the nuclei as classical particles, and thus the nuclear wavefunction $\chi_{J}\left(\mathbf{R},t\right)$ reduces to time-dependent coefficients as the way we described in section 1.2. Methods in this type of mixed quantum/classical treatment including widely used trajectory surface hopping[29,30] and Ehrenfest[31-34] dynamics. The other treats the nuclei quantum mechanically, for example, time-dependent Hartree[35], multi-configuration time-dependent Hartree[36-38],

multiple spawning[39,40], multiple cloning[41] and path integral[42,43] methods. And semi-classical methods[44-50] partially captures the quantum effects of the nuclei.

We have applied *ab-initio* multiple spawning (AIMS) method to investigate the excited state dynamics of materials. AIMS employs the procedure of "*on-the-fly*" computing of the PES where nuclei propagate on. As compared with the mixed quantum/classical dynamics, the quantum mechanical effects of the nuclei are captured by the expansion of the nuclear wavefunction into frozen Gaussian basis[51] functions with time-dependent weights through a spawning procedure,

$$\chi_J(\mathbf{R},t) = \sum_j^{N_J(t)} C_j^J(t)\,\chi_j^J\left(\mathbf{R};\bar{\mathbf{R}}_j^J(t),\bar{\mathbf{P}}_j^J(t),\bar{\gamma}_j^J(t),\alpha_j^J\right) \tag{2-33}$$

$$\begin{aligned}&\chi_j^J\left(\mathbf{R};\bar{\mathbf{R}}_j^J(t),\bar{\mathbf{P}}_j^J(t),\bar{\gamma}_j^J(t),\alpha_j^J\right) \\ &= e^{i\gamma_j^J(t)}\prod_\eta^{3N_{atom}}\chi_{j\eta}^J\left(R_\eta;\bar{R}_{j\eta}^J(t),\bar{P}_{j\eta}^J(t),\alpha_{j\eta}^J\right)\end{aligned} \tag{2-34}$$

$$\begin{aligned}&\chi_{j\eta}^J\left(R_\eta;\bar{R}_{j\eta}^J(t),\bar{P}_{j\eta}^J(t),\alpha_{j\eta}^J\right) \\ &= \left(\frac{2\alpha_{j\eta}^J}{\pi}\right)^{1/4}\exp\left[-\alpha_{j\eta}^J\left(R_\eta-\bar{R}_{j\eta}^J(t)\right)^2+i\bar{P}_{j\eta}^J(t)\left(R_\eta-\bar{R}_{j\eta}^J(t)\right)\right]\end{aligned} \tag{2-35}$$

where $C_j^J(t)$ Is the time dependent complex coefficients, χ_j^J is the time dependent basis functions, $N_J(t)$ is the number of basis functions on J th electronic state at time t, $\bar{\mathbf{R}}_j^J(t)$ is the averaged position, $\bar{\mathbf{P}}_j^J(t)$ is the averaged momentum, $\bar{\gamma}_j^J(t)$ is the phase factor, α_j^J is the width of the frozen Gaussian basis function and η indexes the nuclear degrees of freedom. The $\bar{\mathbf{R}}_j^J(t)$ and $\bar{\mathbf{P}}_j^J(t)$ are propagated according to the classical equation of motion,

$$\frac{\partial \bar{R}_{j\eta}^{J}(t)}{\partial t} = \frac{\bar{P}_{j\eta}^{J}(t)}{m_{\eta}} \tag{2-36}$$

$$\frac{\partial \bar{P}_{j\eta}^{J}(t)}{\partial t} = -\frac{\partial V_{JJ}(\mathbf{R})}{\partial R_{\eta}}\bigg|_{\bar{\mathbf{R}}_{j}^{J}(t)} \tag{2-37}$$

where m_{η} is the mass of ηth nuclear degree of freedom and $V_{JJ}(\mathbf{R})$ is the potential energy experienced by the nuclei at $\mathbf{R}$ on the J th electronic state. The phase factor $\bar{\gamma}_{j}^{J}(t)$ is propagated semiclassically,

$$\frac{\partial \bar{\gamma}_{j}^{J}(t)}{\partial t} = -V_{JJ}\left(\bar{\mathbf{R}}(t)\right) + \sum_{\eta}^{3N} \frac{\left(\bar{P}_{j\eta}^{J}(t)\right)^{2}}{2m_{\eta}} \tag{2-38}$$

The spawning procedure ensures the expansion of the nuclear basis functions only allowed when a certain threshold is reached (for example, when non-adiabatic effect becomes important). Hence controls the growth of computational cost while captures the quantum mechanical effects of the nuclei. For detailed discussions of the AIMS method, I refer the readers to the references 39 and 40.

I will present cases in Chapter 7-15 to show the applications of AIMS to understand the non-radiative decay mechanisms of various optoelectronic materials.

REFERENCES

REFERENCES

(1) Heitler, W. L., F. *Z Phys* **1927**, *44*, 18.

(2) Born, M.; Oppenheimer, R. *Ann Phys-Berlin* **1927**, *84*, 0457.

(3) Pyykko, P.; Stanton, J. F. *Chem Rev* **2012**, *112*, 1.

(4) Szabó, A.; Ostlund, N. S. *Modern quantum chemistry : introduction to advanced electronic structure theory*; Free Press ;Collier Macmillan: New York,London, 1982.

(5) Piela, L. *Ideas of quantum chemistry*, 1st ed.; Elsevier: Amsterdam ; Boston, 2007.

(6) Serrano-Andres, L.; Merchan, M. *J Mol Struc-Theochem* **2005**, *729*, 99.

(7) Teller, E. *J Phys Chem-Us* **1937**, *41*, 109.

(8) Yarkony, D. R. *Reviews of Modern Physics* **1996**, *68*, 985.

(9) Yarkony, D. R. *Chem Rev* **2012**, *112*, 481.

(10) Bernardi, F.; Olivucci, M.; Robb, M. A. *Chemical Society Reviews* **1996**, *25*, 321.

(11) Worth, G. A.; Cederbaum, L. S. *Annu Rev Phys Chem* **2004**, *55*, 127.

(12) Levine, B. G.; Martinez, T. J. *Annual Review of Physical Chemistry* **2007**, *58*, 613.

(13) von Neumann, J.; Wigner, E. *Phys Z* **1929**, *30*, 467.

(14) Runge, E.; Gross, E. K. U. *Physical Review Letters* **1984**, *52*, 997.

(15) Chong, D. P. Recent advances in density functional methods; World Scientific: Singapore ; River Edge, N.J., 1995.

(16) Hirata, S.; Head-Gordon, M. *Chemical Physics Letters* **1999**, *314*, 291.

(17) Levine, B. G.; Ko, C.; Quenneville, J.; Martinez, T. J. *Molecular Physics* **2006**, *104*, 1039.

(18) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. *Chemical Physics* **1980**, *48*, 157.

(19) Werner, H. J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schutz, M. Wiley Interdisciplinary Reviews-Computational Molecular Science 2012, 2, 242.

(20) Gordon, M. S.; Schmidt, M. W. Theory and Applications of Computational Chemistry: The First Forty Years 2005, 1167.

(21) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. Journal of Computational Chemistry 1993, 14, 1347.

(22) Aquilante, F.; Autschbach, J.; Carlson, R. K.; Chibotaru, L. F.; Delcey, M. G.; De Vico, L.; Fdez Galvan, I.; Ferre, N.; Frutos, L. M.; Gagliardi, L.; Garavelli, M.; Giussani, A.; Hoyer, C. E.; Li Manni, G.; Lischka, H.; Ma, D.; Malmqvist, P. A.; Muller, T.; Nenov, A.; Olivucci, M.; Pedersen, T. B.; Peng, D.; Plasser, F.; Pritchard, B.; Reiher, M.; Rivalta, I.; Schapiro, I.; Segarra-Marti, J.; Stenrup, M.; Truhlar, D. G.; Ungur, L.; Valentini, A.; Vancoillie, S.; Veryazov, V.; Vysotskiy, V. P.; Weingart, O.; Zapata, F.; Lindh, R. Journal of Computational Chemistry 2015.

(23) Lischka, H.; Shepard, R.; Pitzer, R. M.; Shavitt, I.; Dallos, M.; Muller, T.; Szalay, P. G.; Seth, M.; Kedziora, G. S.; Yabushita, S.; Zhang, Z. Y. Physical Chemistry Chemical Physics 2001, 3, 664.

(24) Demeras, A. S.; Lepetit, M. B.; Malrieu, J. P. *Chemical Physics Letters* **1990**, *172*, 163.

(25) Zaitsevskii, A.; Malrieu, J. P. *Chemical Physics Letters* **1994**, *228*, 458.

(26) Bauschlicher, C. W.; Langhoff, S. R. *Journal of Chemical Physics* **1988**, *89*, 4246.

(27) Shu, Y. N.; Levine, B. G. *Journal of Chemical Physics* **2013**, *139*.

(28) Shu, Y. A.; Hohenstein, E. G.; Levine, B. G. *Journal of Chemical Physics* **2015**, *142*.

(29) Tully, J. C.; Preston, R. K. *Journal of Chemical Physics* **1971**, *55*, 562.

(30) Tully, J. C. *Journal of Chemical Physics* **1990**, *93*, 1061.

(31) Mclachlan, A. D. *Molecular Physics* **1964**, *8*, 39.

(32) Micha, D. A. *Journal of Chemical Physics* **1983**, *78*, 7138.

(33) Kirson, Z.; Gerber, R. B.; Nitzan, A.; Ratner, M. A. *Surface Science* **1984**, *137*, 527.

(34) Sawada, S. I.; Nitzan, A.; Metiu, H. *Physical Review B* **1985**, *32*, 851.

(35) Gerber, R. B.; Buch, V.; Ratner, M. A. *Journal of Chemical Physics* **1982**, *77*, 3022.

(36) Beck, M. H.; Jackle, A.; Worth, G. A.; Meyer, H. D. Physics Reports-Review Section of Physics Letters 2000, 324, 1.

(37) Wang, H. B.; Thoss, M. *Journal of Chemical Physics* **2003**, *119*, 1289.

(38) Meyer, H. D.; Worth, G. A. *Theoretical Chemistry Accounts* **2003**, *109*, 251.

(39) Ben-Nun, M.; Martinez, T. J. *Journal of Chemical Physics* **1998**, *108*, 7244.

(40) Levine, B. G.; Coe, J. D.; Virshup, A. M.; Martinez, T. J. *Chemical Physics* **2008**, *347*, 3.

(41) Makhov, D. V.; Glover, W. J.; Martinez, T. J.; Shalashilin, D. V. Journal of Chemical Physics 2014, 141.

(42) Topaler, M.; Makri, N. *Journal of Physical Chemistry* **1996**, *100*, 4430.

(43) Shim, S.; Aspuru-Guzik, A. *Journal of Chemical Physics* **2012**, *137*.

(44) Miller, W. H. *Chemical Physics Letters* **1970**, *7*, 431.

(45) Herman, M. F. *Journal of Chemical Physics* **1984**, *81*, 754.

(46) Sun, X.; Wang, H. B.; Miller, W. H. *Journal of Chemical Physics* **1998**, *109*, 7064.

(47) Thoss, M.; Stock, G. *Physical Review A* **1999**, *59*, 64.

(48) Bonella, S.; Coker, D. F. *Journal of Chemical Physics* **2001**, *114*, 7778.

(49) Virshup, A. M.; Punwong, C.; Pogorelov, T. V.; Lindquist, B. A.; Ko, C.; Martinez, T. J. Journal of Physical Chemistry B 2009, 113, 3280.

(50) Zhu, C. Y.; Teranishi, Y.; Nakamura, H. Advances in Chemical Physics, Vol. 117 2001, 117, 127.

(51) Heller, E. J. *Journal of Chemical Physics* **1981**, *75*, 2923.

Chapter 3

Reducing The Propensity For Unphysical Wavefunction Symmetry Breaking In Multireference Calculations Of The Excited States

"Happy the man who has been able to know the causes of things."

-Publius Vergilius Maro (70 BC-19 BC)

This chapter is reproduced from paper: Reducing the propensity for unphysical wavefunction symmetry breaking in multireference calculations of the excited states of semiconductor clusters. Y. Shu, B. G. Levine, J. Chem. Phys., 139, 074102 (2013).

In recent years, nonadiabatic molecular dynamics simulation has served as a powerful tool to elucidate the microscopic dynamics of excited semiconducting systems.[1-7] In such simulations the potential energy surface (PES) and electronic structure are often computed on-the-fly using one of several single-reference electronic structure methods, such as configuration interaction singles[8] or time-dependent density functional theory.[9,10] Though these methods are well-suited for the study of transitions between excited electronic states and have a desirable computational cost, they fail to properly describe conical intersections[11] between the closed-shell ground state and open-shell

excited states,[12] the proper description of which is vital to accurately modeling non-radiative decay. Thus, it would be desirable to employ multireference electronic structure methods, which have long been the methods of choice for modeling non-radiative decay in molecules. Specifically, the complete active space self-consistent field (CASSCF) method[13] is a very popular choice for modeling such decay owing to its low computational cost, its ability to provide a qualitatively accurate description of complex PESs, and the availability of analytic calculations for gradients and nonadiabatic coupling vectors.[14]

Though widely used and effective, CASSCF sometimes exhibits undesirable behavior. It is well known that the results of CASSCF calculations depend strongly and unpredictably on the choice of the orbital active space, with errors in relative energies of several electronvolts observed in some cases. This sensitivity can be alleviated by including dynamic electron correlation via perturbation theory[15] or configuration interaction (CI), but such approaches increase the computational cost of calculations dramatically, rendering on-the-fly dynamic simulations possible only for relatively small molecules.[16] In addition to the strong dependence of results on the choice of active space, CASSCF also exhibits convergence difficulties, multiple solutions, and PES discontinuities,[17-19] all of which are highly undesirable in the context of AIMD simulations. These issues arise from the nonlinearity of the CASSCF problem, in which the orbitals and CI coefficients are self-consistently optimized.

Artificial wavefunction symmetry breaking—which in many cases persists for all

reasonably sized active spaces—is another source of error in CASSCF calculations of the low-lying excited states of molecules. As first pointed out by Löwdin, there is no guarantee that variational optimization of a particular form of approximate wavefunction will result in a solution that transforms as an irreducible representation of the point group of the Hamiltonian.[20] Spatial wavefunction symmetry breaking in self-consistent field (SCF) calculations has been studied in a range of systems, particularly radicals and excited states, where the insufficiency of a single-determinant trial wavefunction results in a variationally optimal solution that does not reflect the molecular symmetry.[21-49] Such symmetry breaking often occurs when there exist multiple nearly degenerate low-lying electronic configurations of different irreducible representations that differ from each other by a single excitation. In such cases, discontinuities in the gradient of the energy or the energy itself with changes in nuclear geometry are often observed, and such artifacts can be exacerbated when post-Hartree-Fock treatments of electron correlation are performed with a symmetry-broken reference.[30] In many systems this unphysical symmetry breaking can be eliminated by application of more complex reference wavefunctions or ensembles, such as those employed in non-orthogonal CI,[21,32] multireference and valence bond methods (based on a carefully constructed and often large configuration spaces),[23,24,26,27,36,42,44] state-averaging/ensemble approaches,[37,43] and Brueckner coupled cluster approaches.[25,28,29,33,40,50,51] In some cases, it can be argued that a symmetry-broken solution actually provides a more accurate physical description of the system than a symmetry-preserving one,[31,46,47,52,53] but such a solution is not

desirable when continuous behavior of the electronic structure with respect to nuclear motion is required, as is the case in AIMD simulations. In these cases it would be desirable to determine a qualitatively accurate, symmetry-preserving wavefunction at a reasonable computational cost.

The root of unphysical symmetry breaking of multireference wavefunctions can be understood in similar terms to that of single-reference wavefunctions. It is instructive to consider such symmetry breaking in analogy to the pseudo-Jahn-Teller effect (PJTE).[54] In systems exhibiting the PJTE, the lowest energy molecular geometry is of lower symmetry because nearly degenerate electronic configurations of different irreducible representations can couple and split when the symmetry of the Hamiltonian is reduced by distortion of the nuclear structure, thus lowering the energy of the lower state. This splitting arises because non-zero off-diagonal Hamiltonian matrix elements between these electronic configurations are introduced by the reduction of symmetry. At the CASSCF level of theory, however, no reduction of the symmetry of the geometry of the molecule is required to introduce such off-diagonal elements. Instead, when CASSCF orbitals are variationally optimized without enforcement of symmetry, the reduction of the symmetry of the orbitals themselves can introduce analogous non-zero off-diagonal Hamiltonian matrix elements between configurations of nominally different point groups. Like in the PJTE, these off-diagonal elements result in mixing and splitting of the configurations involved and potentially a symmetry-broken variational solution. This can be thought of as a PJTE in orbital space, as opposed to in nuclear position space.

An often discussed example of unphysical spatial symmetry breaking is the He2+ ion.[24,32] The wavefunction of the He2+ ion can be described in terms of two low-lying, nearly degenerate, symmetry-adapted molecular orbitals: σg (bonding) and σu (antibonding). Though usually discussed as an example of the symmetry breaking of a single determinantal wavefunction, symmetry breaking is also observed in the variational optimization of a two-determinant multiconfigurational SCF (MCSCF) wavefunction constructed from determinants σg2σu and σgσu2. (In fact, such an MCSCF wavefunction is just a different representation of the single determinant Hartree-Fock wavefunction.) The σg2σu and σgσu2 configurations become nearly degenerate at large but finite He-He distance. Though uncoupled when σg and σu are perfectly symmetric, the distortion of the orbitals from perfect even or odd symmetry allows the coupling and splitting of these configurations which in turn reduces the variational energy of the ground state. Thus, the variational solution is symmetry broken (skewed to have a greater electron density on one He atom or the other).[55] In this specific case, the driving force for symmetry breaking is often termed the orbital size effect; the wavefunction is stabilized by expansion away from the nucleus on the He atom which has the less positive charge, and contracting towards the nucleus on the side with the more positive charge. However, this is only one example of the more general PJTE-like splitting.

It is reasonable to assert that this PJTE-like symmetry breaking arises when the choice of active space is insufficient to properly describe the electronic structure of the molecule. In the case of He2+, a symmetry preserving wavefunction can be created by inclusion of

more electronic configurations in the wavefunctions. However, as will be demonstrated by example below, at times no reasonably sized active space eliminates artifactual symmetry breaking. When a multireference description of the electronic structure is required, but CASSCF fails to provide a suitable description of the electronic structure regardless of the choice of active space, an alternative approach is necessary. With this in mind, we turn our attention to two-step complete active space configuration interaction (CASCI) methods, which decouple the optimization of the orbitals from that of the CI coefficients. CASCI approaches based on improved Hartree-Fock virtual orbitals (IVO),[56] unrestricted Hartree-Fock natural orbitals,[57] natural orbitals of correlated wavefunction calculations,[58] natural orbitals of high-multiplicity states,[59] and floating occupation molecular orbitals (FOMO)[60,61] have shown great promise as algorithmically simpler and computationally less expensive alternatives to CASSCF. In addition to their relative simplicity and low computational cost, a careful choice of the orbital optimization approach can lead to a two-step approach with more desirable behavior than CASSCF, despite the fact that the resulting wavefunctions are variationally inferior.

With the goal of developing multireference electronic structure methods as tools for the study of semiconductor photochemistry, we have applied SA-CASSCF to compute the excited states of several small semiconductor clusters and related molecules. Herein we present results demonstrating that such calculations often result in unphysical wavefunction symmetry breaking for all reasonable choices of active space. Based on the premise that eliminating off-diagonal Hamiltonian matrix elements from the energy

expression used to variationally optimize the orbitals would eliminate the driving force for such symmetry breaking, we introduce a two-step multireference method—the singly excited active space self-consistent field (SEAS-) CASCI method—which is demonstrated to reduce the propensity for unphysical wavefunction symmetry breaking. It is further demonstrated that, relative to SA-CASSCF, qualitative errors in the PES and molecular properties attributable to artifactual symmetry breaking are eliminated at the SEAS-CASCI level of theory in several cases, and that unphysical wavefunction distortion at completely asymmetric geometries is also corrected.

3.1 Method

3.1.1 Singly excited active space self-consistent field method

In the first step of the SEAS-CASCI procedure, a SEAS SCF calculation is done to determine a set of orbitals. The initial orbital determination step is followed by the construction and diagonalization of a CASCI Hamiltonian.

To define the energy expression optimized in the SEAS procedure we first partition the molecular orbitals (MOs) into several sets: closed orbitals (which are doubly-occupied in all configurations), active orbitals, and virtual orbitals. The active orbitals are further partitioned into a lower active set, which would be occupied in a restricted-Hartree-Fock-like (RHF-like) closed-shell single determinant wavefunction, and the upper active orbitals, which are unoccupied in this RHF-like determinant. The SEAS energy to be optimized is the average energy of this RHF-like determinant and all possible singlet configuration state

functions created by exciting a single electron from a lower active to an upper active orbital,

$$E_{SEAS} = \frac{1}{1+\frac{M}{2}\left(N-\frac{M}{2}\right)}\left[\left\langle \Psi_0 \middle| \hat{H} \middle| \Psi_0 \right\rangle + \sum_i \sum_a \left\langle {}^1\Psi_i^a \middle| \hat{H} \middle| {}^1\Psi_i^a \right\rangle\right] \tag{3-1}$$

with *M* and *N* representing the number of active electrons and orbitals respectively, $\left|\Psi_0\right\rangle$ representing a closed-shell RHF-like determinant, and $\left|{}^1\Psi_i^a\right\rangle$ representing a singlet configuration state function with an electron excited from lower active orbital i to upper active orbital a. We term E_{SEAS} the SEAS energy.

The SEAS energy is independent of the off-diagonal Hamiltonian matrix elements between the included configurations, and therefore one would not expect to observe PJTE-like orbital symmetry breaking driven by the off-diagonal elements between included configurations. The independence of the SEAS energy from the off-diagonal elements of the Hamiltonian allows it to be rewritten in the form of a generalized Hartree-Fock expression,

$$\begin{aligned}
E_{SEAS} = & \sum_k 2h_{kk} + \sum_i \left(2-\frac{N}{n}+\frac{M}{2n}\right)h_{ii} + \sum_a \frac{M}{2n}h_{aa} + \sum_k \sum_l \left[2(kk\,|\,ll) - (kl\,|\,kl)\right] \\
& + \sum_k \sum_i \left(2-\frac{N}{n}+\frac{M}{2n}\right)\left[(kk\,|\,ii) - \frac{1}{2}(ki\,|\,ki)\right] + \sum_a \sum_k \frac{M}{2n}\left[(kk\,|\,aa) - \frac{1}{2}(ka\,|\,ka)\right] \\
& + \sum_i \sum_j \left[\left(2-\frac{N}{n}+\frac{M}{2n}\right)(jj\,|\,ii) - \left(1-\left(\frac{N}{2n}-\frac{M}{4n}\right)(1-\delta_{ij})\right)(ji\,|\,ji)\right] \\
& + \sum_i \sum_a \left[\frac{M}{4n}(ii\,|\,aa) + \frac{6-M}{8n}(ia\,|\,ia)\right]
\end{aligned} \tag{3-2}$$

Where h_{ii} represents a diagonal element of the one-electron Hamiltonian, $(ij\,|\,kl)$ represents a two-electron integral in chemists' notation, a indexes upper active orbitals, i

and j index lower active orbitals, and k and l index closed orbitals. The total number of spin-adapted configurations averaged in the SEAS energy is given by

$$n = 1 + \frac{M}{2}\left(N - \frac{M}{2} \right) \tag{3-3}$$

The above procedure defines the SEAS orbital determination step. The second step in the SEAS-CASCI procedure is to construct and diagonalize a CASCI Hamiltonian using the SEAS orbitals. Despite the single-reference-like nature of the SEAS equations, the final CASCI step results in a fully multireference wavefunction which is invariant to rotations among active orbitals (though the SEAS SCF energy, Eq. (3-1), of the initial orbital determination step is not invariant to rotations between the upper active and lower active spaces).

The SEAS-CASCI method offers specific advantages over the existing methods for eliminating wavefunction symmetry breaking. The SEAS-CASCI method was conceived to produce a symmetry-preserving wavefunction even when the standard state-averaged multireference approaches fail. Relative to non-orthogonal CI approaches, the inclusion of dynamic electron correlation via MRCI or perturbative expansions will be relatively straight forward. Brueckner orbitals generally relieve problems with symmetry breaking on a state-at-a-time basis, but are best thought of as reflecting relaxation due to electron correlation. In contrast, the present approach was conceived to determine a set of orbitals in which to efficiently expand electronically excited wavefunctions to zeroth order. Additionally, the state-at-a-time approach does not facilitate the proper treatment of

surface crossings, which are important features of the PES when modeling non-radiative decay.

Though discussed in the context of symmetry breaking in the present work, SEAS-CASCI offers a conceptual advantage over existing two-step CASCI methods as a tool for describing electronic excited states. In a similar spirit to SA-CASSCF, the SEAS energy expression was designed to variationally optimize the occupied and virtual orbitals to provide a nearly optimal zeroth-order description of low-lying excited states. Two-step CASCI approaches based on the natural orbitals of unrestricted Hartree-Fock[57] or correlated[58] calculations of the ground state provide an admirable description of bond breaking processes, but would not be expected to provide an efficient basis in which to expand excited states. The IVO-CASCI procedure of Freed and coworkers[56] optimizes the virtual orbitals to provide an efficient description of low-lying excited states, but leaves the occupied orbitals in their Hartree-Fock form, which are optimal for electron removal rather than excitation. In FOMO-CASCI,[60,61] both occupied and virtual orbitals are optimized to minimize the average energy of an ensemble of states including the neutral ground state, neutral excited states, and various ground and excited ionic states. The advantages of optimization of the occupied orbitals for excitation in the context of CASCI methods will be the focus of a future study.

We have performed all SEAS orbital determinations presented below by use of the state-averaged MCSCF function in the MolPro electronic structure package[62] to minimize

E_{SEAS} as represented in Eq. (3-1). The SEAS orbitals are then used to build a full CASCI Hamiltonian, which is diagonalized to determine the final SEAS-CASCI energies with a second call to the same function. We have verified the validity of Eq. (3-2) using the generalized Hartree-Fock code in the GAMESS software package[63] to reproduce orbitals identical to those determined by our MolPro implementation of Eq. (3-1) in several cases.

3.1.2 Accessing wavefunction symmetry breaking

The majority of the molecules which are notorious for unphysical symmetry breaking have open-shell ground states, but the present method was specifically developed with a different class of symmetry-breaking systems in mind: the open-shell, singlet excited states of closed shell systems. Though many radical molecules which exhibit unphysical wavefunction symmetry breaking are quite small, it is our experience that symmetry breaking in the excited states of closed-shell molecules occurs predominantly in larger systems. Thus, our study was focused on three moderately sized molecules relevant to semiconductor photochemistry for which unphysical symmetry breaking is observed at the SA-CASSCF level of theory: sila-adamantane ($Si_{10}H_{16}$), adamantane ($C_{10}H_{16}$), and a titanium (IV) hydroxide cluster with all Ti-O-H angles constrained to be linear, which we hereafter referred to as linear $Ti(OH)_4$ (pictured in the insets of Figures 3-1a, 3-1c, and 3-2a, respectively). The two adamantane systems represent clusters cut from the bulk diamond and silicon crystal lattices, respectively, and terminated with hydrogen atoms, while tetrahedral titanium oxide centers are reported to have interesting photocatalytic properties.[64,65] All three clusters are of T_d symmetry.

The geometries of our chosen molecules were optimized at the Moller-Plesset second order perturbation theory (MP2) level of theory[66] with the 6-31G** basis set.[67] All optimizations were initialized from asymmetric geometries and allowed to converge to nearly symmetric final structures. These slight deviations from symmetry were allowed to facilitate the symmetry breaking of the wavefunction in cases where it is variationally favored. The deviation from symmetry of the final structures is measured by calculation of the standard deviation of the symmetry equivalent geometric parameters, which are reported in Table 3-1 for the three T_d clusters. For each cluster, equation-of-motion coupled cluster singles and doubles (EOM-CCSD)[68] calculations based on a symmetry-preserving RHF/6-31G** reference were performed to identify the characters and degeneracies of the lowest electronic excited states, and the sets of states observed to be degenerate are reported in Table 3-2. Table 3-3 presents the average energy gaps between the degenerate states calculated at this level of theory. The deviations of the geometries (Tables 3-1) and of the excited EOM-CCSD wavefunctions (as measured by the energy gap between nominally degenerate states; Table 3-3) from symmetry are very small, indicating that large dipoles and energy gaps predicted in other calculations are artifacts of the choice of method, not physically correct descriptions of the slightly asymmetric molecular geometries.[69]

Table 3-1. The standard deviations of symmetry equivalent geometric parameters of the three T_d molecules studied. The very small values demonstrate that the nuclear geometries are nearly but not exactly symmetric.

Molecule	Titanium Hydroxide - Linear		Adamantane			Sila-adamantane		
Geometric Parameter	Ti-O bonds / Angstrom	O-Ti-O angles / Degree	C-C bonds / Angstrom	C-CH-C angles / Degree	C-CH_2-C angles / Degree	Si-Si bonds / Angstrom	Si-SiH-Si angles / Degree	Si-SiH_2-Si angles / Degree
Standard Deviation	$1.1*10^{-5}$	$9.4*10^{-3}$	$3.1*10^{-5}$	$7.7*10^{-3}$	$2.5*10^{-3}$	$3.2*10^{-5}$	$5.3*10^{-5}$	$2.5*10^{-5}$

Table 3-2. For each of the three T_d molecules, this table includes the number of states averaged in SA-CASSCF calculations, the set of states which are expected to be degenerate by symmetry, and the sets of orbitals expected to be degenerate by symmetry. The average gap between the expected degenerate states is used as an indicator of whether the wavefunction preserves symmetry or not.

Molecule	States Averaged	Expected Degenerate States	Expected Sets of Degenerate Orbitals
Sila-adamantane	5	S_2,S_3,S_4	71,72 (Si-Si σ) 73,74,75 (Si-Si σ) 76,77,78 (HOMO, Si-Si σ) 79,80,81 (LUMO, Si-Si σ*) 82 (Si-Si σ*, non-degenerate) 83,84 (Si-Si σ*)
Adamantane	4	S_1,S_2,S_3	31,32 (C-C σ) 33,34,35 (C-C σ) 36,37,38 (HOMO, C-C σ) 39 (LUMO, C-C σ*, non-degenerate) 40,41,42 (C-C σ*) 43,44,45 (C-C σ*)
Linear $Ti(OH)_4$	4	S_1,S_2,S_3	22,23 (O lone pair) 24,25,26 (O lone pair) 27,28,29 (HOMO, O lone pair) 30,31 (LUMO, Ti 3d orbitals) 32,33,34 (Ti 3d orbitals) 35 (Ti 4s orbital, non-degenerate)

Table 3-3. Dipole moments and the average energy gaps between states expected to be degenerate by symmetry calculated at single-reference levels of theory with symmetry-preserving references. All values are expected to be zero by symmetry. The very small deviations indicate that the slight asymmetries of the nuclear configurations do not result in large changes in these observables.

	CCSD Dipole Moment / Debye	RHF Dipole Moment / Debye	EOM-CCSD Average Energy Gap between Degenerate States / eV
Sila-adamantane	$2.15*10^{-3}$	$2.12*10^{-3}$	$2.88*10^{-5}$
Adamantane	$1.06*10^{-5}$	$2.39*10^{-5}$	$1.07*10^{-4}$
Linear $Ti(OH)_4$	$4.20*10^{-5}$	$1.13*10^{-4}$	$1.07*10^{-4}$

Table 3-4. The number of symmetry-preserving active spaces for each combination of molecule and method. These counts are referenced to the total number of active spaces considered. Application of SEAS-CASCI and freezing the closed-shell orbital coefficients at their RHF values are successful strategies for enforcing the symmetry of the wavefunction.

	CASSCF	CASSCF-froz	SEAS-CASCI	SEAS-CASCI-froz
Sila-adamantane	0 out of 6	2 out of 6	4 out of 6	6 out of 6
Adamantane	3 out of 8	6 out of 8	5 out of 8	8 out of 8
Linear $Ti(OH)_4$	0 out of 9	1 out of 9	3 out of 9	9 out of 9
Total	**3 out of 23**	**9 out of 23**	**12 out of 23**	**23 out of 23**

In all cases, the excitations in the adamantane clusters are of σ→σ* character, while in linear $Ti(OH)_4$ the lowest excitations are characterized by the transfer of a single electron from the oxygen lone pair orbitals of the hydroxide ligands to the *d*-shell of the titanium atom. The smallest active spaces considered (6/7 in the case of sila-adamantane, 6/4 for adamantane, and 6/5 for linear $Ti(OH)_4$) contain all of the orbitals necessary to describe the dominant character of the excitations predicted by EOM-CCSD. All active spaces (state averaged ensemble) are chosen such that sets of orbitals (states) that are expected to be degenerate by symmetry are not split. The size of the state averages employed and the expected state and orbital degeneracies used to determine the set of reasonable active spaces for each cluster are presented in Table 3-2. State averages are chosen such that the sets of states expected to be degenerates and all lower energy states are included in the average. In all three cases we have considered all active spaces that meet these criteria and can be performed at a reasonable computational cost. The definition of "reasonable" here is rather inclusive, and we do not mean it to indicate that these active spaces are reasonable to use in practice, only that they are reasonable to try. Here we consider all of these active spaces in our analysis to demonstrate that the symmetry-breaking behavior observed below is not the result of a single poor choice of active space.

For each active space determined to be reasonable, a set of four multireference methods were applied: state-averaged CASSCF, state-averaged CASSCF with all closed-shell orbital coefficients frozen at their symmetry-preserving RHF values (CASSCF-froz),

standard SEAS-CASCI (as described above), and SEAS-CASCI with all closed-shell orbital coefficients frozen at their RHF values (SEAS-CASCI-froz). All presented multireference calculations were performed using converged RHF orbitals for the neutral ground state of the system as the starting guess. In all cases, the shape and degeneracy of the starting orbitals reflected the full symmetry of the molecule. (Though not presented here, other initial guesses including Kohn-Sham orbitals and symmetry-preserving SEAS orbitals were also employed, with no significant change in the tendency towards symmetry breaking.) All multireference calculations were performed in MolPro using the 6-31G** basis. (At these relatively low levels of theory, energies and properties are fairly well converged with respect to basis set even with this relatively small basis, though a much larger basis and some inclusion of dynamic correlation would be needed to achieve quantitative results.[70]) Throughout this paper active spaces are abbreviated (*M*/*N*), where *M* and *N* are the numbers of active electrons and orbitals, respectively.

For each calculation, the symmetry of the final wavefunction is assessed by considering two properties: the dipole moment of the ground state (which should be zero for all T_d molecules) and the average energy gap between states which are expected to be degenerate by symmetry (as defined in Table 3-2). Deviation of either observable from zero indicates a symmetry-broken wavefunction. In all cases the final orbitals and CI vectors were investigated to ensure that the energy gaps are calculated between states which have approximately the same character as those identified in the reference EOM-CCSD calculations.

3.2 Results and Discussion

3.2.1 Unphysical Symmetry Breaking in SA-CASSCF and SEAS-CASCI

The results of the study described in section II.B. for the adamantane clusters and linear $Ti(OH)_4$ are presented in Figures 3-1 and 3-2, respectively. Table 3-4 summarizes the results. Note that for the three T_d clusters, full CASSCF only preserves the symmetry of the wavefunction in three out of the 23 selected active spaces, all of them for the adamantane cluster: CAS(6/4), CAS(12/7), and CAS(16/12). In all other cases the ground state dipole moment, the energy gap between would-be degenerate states, or both of these observables deviate significantly from zero due to wavefunction symmetry breaking.

The application of the full SEAS-CASCI approach gives a significant improvement. In twelve out of 23 cases the wavefunction preserves spatial symmetry. In the case of $Ti(OH)_4$ and sila-adamantane, it is mostly smaller active spaces which result in unphysical symmetry breaking. This suggests that these choices of active space may include configurations which are strongly coupled to singly excited configurations outside the active space, and thus symmetry breaking can occur by mixing between orbitals inside and outside of the active space. For example, in the case of $Ti(OH)_4$, linear combinations of the eight oxygen lone pair orbitals are split into three degenerate sets (Table 3-2). Despite the fact that the splittings between the energies of these orbitals are quite large (1.24 eV and 0.66 eV at the RHF level of theory), symmetry is not preserved at the SEAS-CASCI or CASSCF level of theory unless all three sets (16 active electrons) are included in the active space.

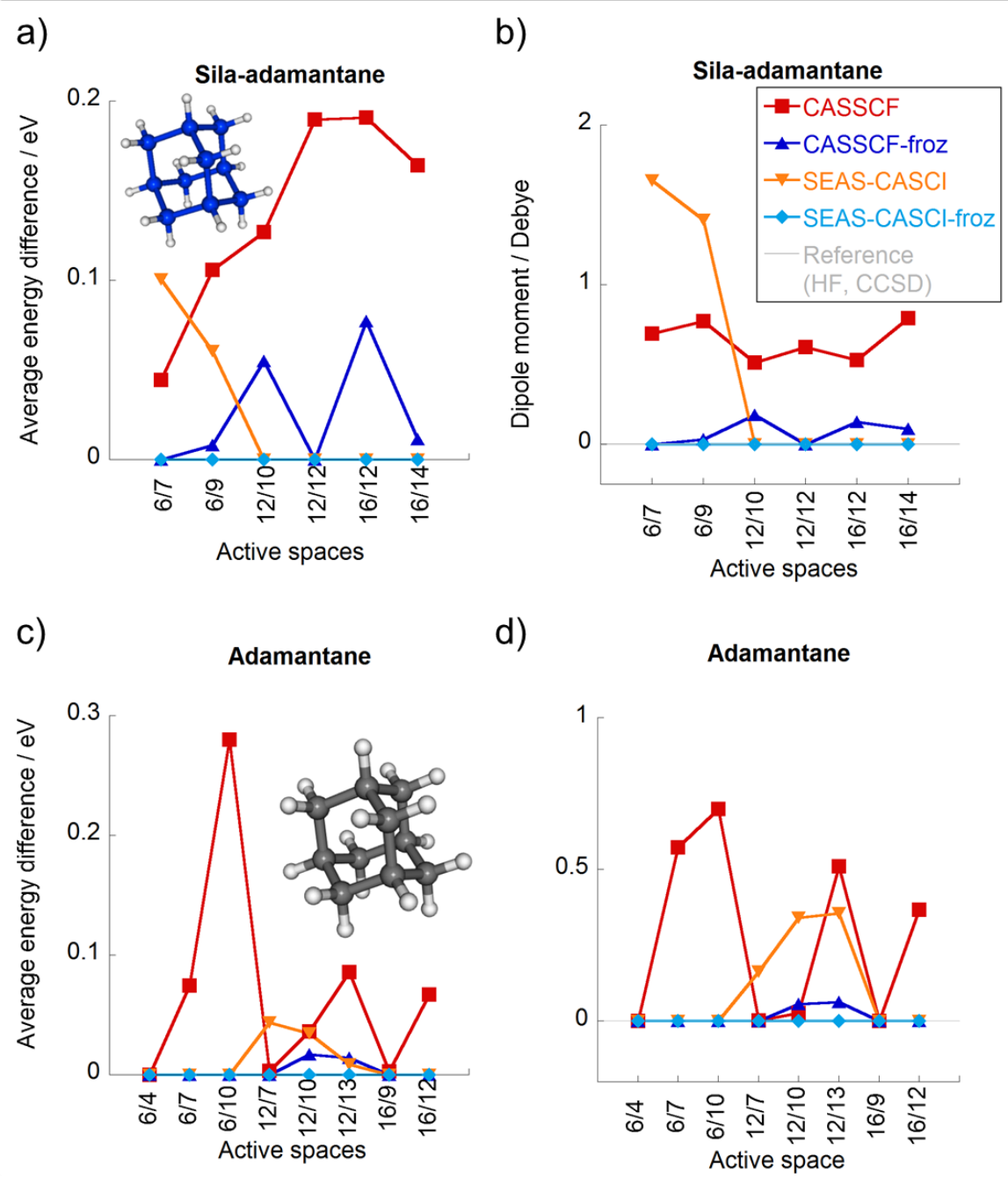


Figure 3-1. a) and c) The average energy difference between degenerate states (as defined in Table 3-2) as a function of active space for sila-adamantane (a) and adamantane (c) calculated using various multireference methods. b) and d) The S_0 dipole moment as a function of active space for sila-adamantane (b) and adamantane (d) calculated using

Figure 3-1(cont'd). the same set of multireference methods. Deviation of either value from zero indicates wavefunction symmetry breaking. SEAS-CASCI predicts a symmetry-preserving wavefunction in more cases than CASSCF. Freezing of closed orbitals in their RHF configurations also decreases the tendency toward symmetry breaking. The sila-adamantane and adamantane molecules are pictured in the insets of panels a) and c), with silicon, carbon, and hydrogen atoms represented in blue, gray, and white, respectively.

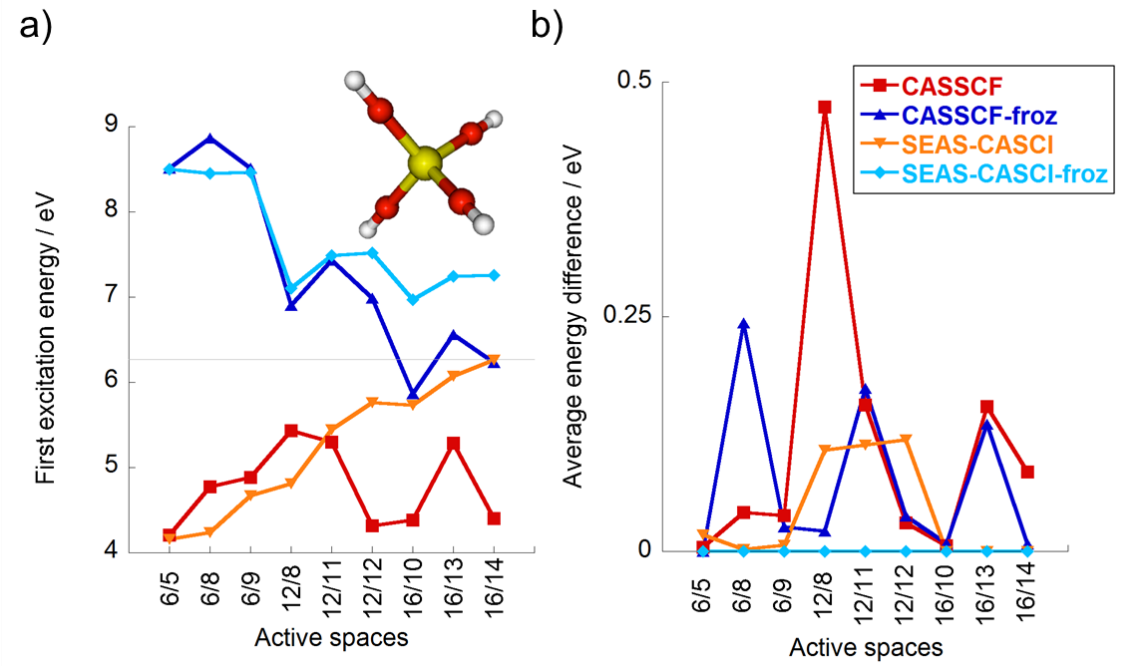


Figure 3-2. a) The average energy difference between degenerate states (as defined in Table 3-2) as a function of active space for the linear titanium hydroxide cluster calculated using various multireference methods. b) The S_0 dipole moment as a function of active space for the linear titanium hydroxide cluster using the same set of multireference methods. Deviation of either value from zero indicates wavefunction symmetry breaking. SEAS-CASCI predicts a symmetry-preserving wavefunction in more cases than CASSCF. Freezing of closed orbitals in their Hartree-Fock configurations also decreases the tendency toward symmetry breaking. The linear $Ti(OH)_4$ cluster is pictured in the inset of a), with titanium, oxygen, and hydrogen atoms represented in yellow, red, and white, respectively.

The freezing of doubly-occupied orbital coefficients at their symmetry-preserving RHF values decreases the tendency towards symmetry breaking in both CASSCF and SEAS-CASCI. When CASSCF-froz is applied, nine of 23 chosen active spaces provide symmetry-preserving solutions, while all 23 chosen active spaces provide symmetry-preserving solutions when SEAS-CASCI-froz is applied. Constraining the occupied inactive orbital coefficients to their HF values prevents them from mixing with the active orbitals, and thus discourages symmetry breaking. For example, when small active spaces (those containing less than the full set of 16 lone pair electrons) are applied to the $Ti(OH)_4$ cluster, the inactive MOs constructed from linear combinations of the lone pairs are frozen, and thus unable to mix with the active lone pair MOs, thus decreasing the variational space available to the active orbitals. In order to break symmetry, these MOs must now mix with inactive virtual orbitals, which remain unfrozen, but it appears that such mixing is unfavorable.

The inclusion of frozen RHF orbitals is similar to doing state-averaged MCSCF with the ground state weighted more heavily than the excited states, in that both approaches bias the orbitals towards those that favor the ground state. Based on this observation, it is reasonable to suggest that the dynamical state weighting approach of Deskevich, et al.[71] might be adopted to reduce the incidence of symmetry breaking in MCSCF calculations.

3.2.2 Relative accuracy of SEAS-CASCI and SA-CASSCF

To gauge the relative accuracy of SEAS-CASCI and CASSCF, the first vertical excitation energies (averaged over nominally degenerate states) of the three T_d clusters were

calculated as a function of active space and compared to reference calculations at the EOM-CCSD/6-31G** level of theory (Figure 3-3). In the two adamanatane clusters, both SEAS-CASCI and CASSCF significantly over-predict the vertical excitation energies regardless of active space. In both cases, variation between active spaces and between methods is small relative to the large deviations from the EOM-CCSD reference. It is well known that CASSCF predicts unphysically large vertical excitation energies due to the absence of dynamic electron correlation, and it is not surprising that SEAS-CASCI exhibits the same behavior.

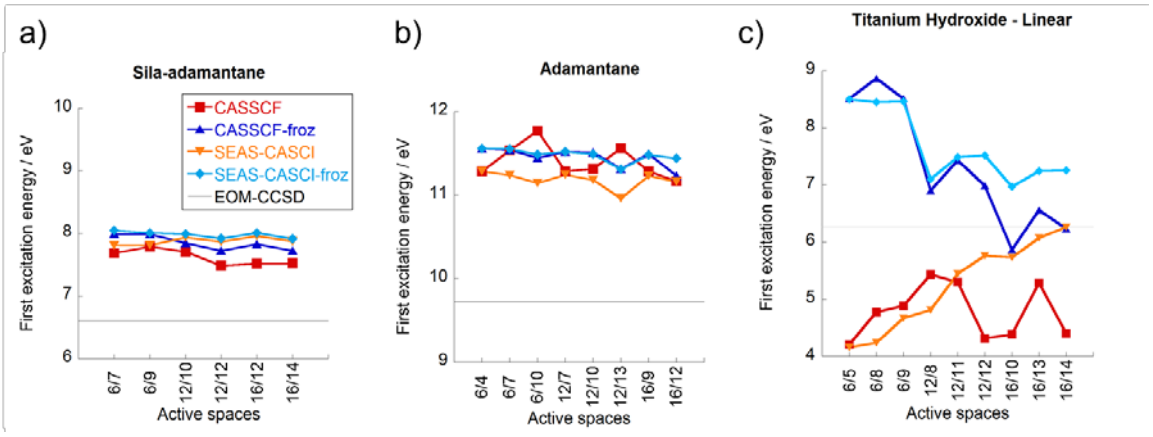


Figure 3-3. The first vertical excitation energies (averaged over nominally degenerate states) of the three T_d clusters as a function of active space calculated using several different multireference methods. A thin black line indicates the EOM-CCSD reference energies. In all cases errors relative to EOM-CCSD are significant, but SEAS-CASCI performs similarly well to CASSCF.

The vertical excitation energies of $Ti(OH)_4$ behave differently as a function of active space. In contrast to the typical overestimation of vertical excitation energies, both SA-CASSCF and SEAS-CASCI predict vertical excitation energies lower than the EOM-CCSD reference for $Ti(OH)_4$, with smaller active spaces predicting lower vertical excitation energies than larger ones. As was seen above, these smaller active spaces exhibit symmetry breaking at both the CASSCF and SEAS-CASCI levels of theory. It appears that this symmetry breaking favors the excited states relative to the ground states, and is exacerbated by the choice of a smaller active space. In contrast, CASSCF-froz and SEAS-CASCI-froz over-predict the vertical excitation energies, again with larger errors predicted by the smaller active spaces. The majority of these calculations preserve symmetry, and are likely exhibiting the usual tendency of calculations which do not treat dynamic electron correlation accurately to over-predict the vertical excitation energy. Errors approaching 3 eV were observed in some cases suggesting, again, that these smaller active spaces are insufficient to describe excitations in $Ti(OH)_4$ accurately due to the strong couplings of the oxygen lone pair orbitals with each other and with the Ti d orbitals. This case points out the extreme importance of a judicious choice of active space in applying any multireference method, whether the orbital and CI coefficient optimizations are coupled or not.

3.2.3 Unphysical wavefunction distortion in asymmetric molecules

The presence of unphysical cusps in the PES at geometries of high symmetry resulting from artifactual spatial symmetry breaking at the Hartree-Fock level of theory is well

known, and it has been predicted that multiconfigurational SCF procedures such as CASSCF should not be immune to this behavior.[30] To explore the effects of symmetry breaking at the CASSCF and SEAS-CASCI levels of theory on the PES, we first consider a simpler hydroxide cluster: $Be(OH)_2$. The ground state minimum structure of $Be(OH)_2$ was optimized enforcing C_s symmetry (all atoms in the same plane). The final optimized geometry (pictured in the inset of Figure 3-4b) was of C_{2v} symmetry to within numerical error. The PES as a function of the asymmetric Be-O stretching mode of $Be(OH)_2$ was scanned by rigidly increasing the length of one Be-O bond while decreasing the other Be-O bond by an equivalent distance, thus reducing the C_{2v} symmetry of the molecule to C_s. All other Z-matrix coordinates were held constant at their ground state optimal values. EOM-CCSD/6-31G** calculations are presented as a reference. The PES and dipole moment of S_1 as a function of this asymmetric stretching coordinate are presented in Figure 3-4a and b, respectively.

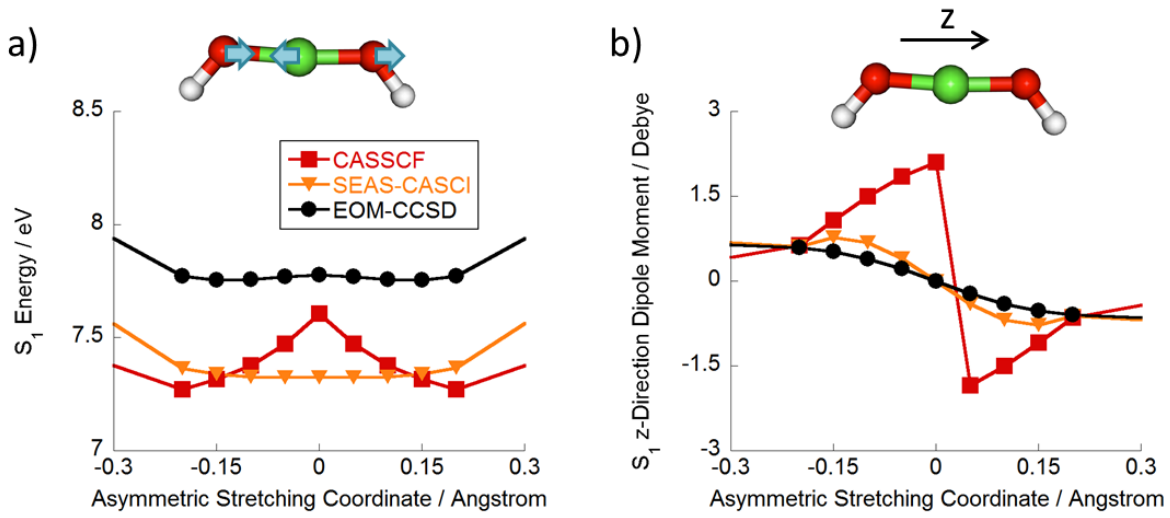


Figure 3-4. a) The PES of the first excited state of beryllium hydroxide as a function of a rigid asymmetric stretching (AS) coordinate calculates at three levels of theory. CASSCF predicts a large cusp in the PES at the C_{2v} symmetric (zero AS) geometry, while the SEAS-CASCI surface is very flat with respect to AS. The shape of the SEAS-CASCI curve is in much better agreement with the symmetry-preserving EOM-CCSD reference. b) The z component of the S_1 dipole moment of beryllium hydroxide as a function of the AS coordinate calculated at the same levels of theory. This value should be zero by symmetry at the C_{2v} structure, but CASSCF predicts a large non-zero value, indicating a symmetry-broken wavefunction. A sudden change in the dipole at zero indicates discontinuous behavior of the CASSCF wavefunction. The relatively smooth curve predicted by SEAS-CASCI is in better agreement with the EOM-CCSD reference.

An unphysical cusp in the S_1 PES is observed at the C_{2v} geometry (where the asymmetric stretching coordinate is defined to be zero) when the calculations are performed at the CASSCF (SA2-CAS(8/8)/6-31G**) level of theory.[72] Though this cusp looks similar to the PES of the lower of two states involved in a conical intersection, the next excited state is 2.4 eV above S_1, and thus there is no such intersection. Instead, this cusp results from symmetry breaking in the CASSCF wavefunction, which can be seen in the z component of the CASSCF-calculated S_1 dipole moment of the molecule (Figure 3-4b), which has an unphysical non-zero value and changes discontinuously at the C_{2v} geometry point, indicating a sudden change of wavefunction. In contrast, the SEAS-CASCI level of theory predicts a smoothly differentiable S_1 PES, and continuous dipole moment curve. The shape of both SEAS-CASCI curves agrees better with the EOM-CCSD reference than does the CASSCF curve (though EOM-CCSD predicts a very small barrier at the C_{2v} geometry where SEAS-CASCI predicts a very shallow minimum).

Similar behavior can be seen when SA-CASSCF, SEAS-CASCI, and EOM-CCSD are applied to calculate the PES of linear $Ti(OH)_4$ along an asymmetric stretching coordinate (Figure 3-5). The asymmetric stretching coordinate is defined as simultaneously shortening two Ti-O bonds and lengthening the other two by the same distance. All other internal coordinates are held constant at the values of the ground state optimized T_d structure. The large and chemically intuitive (16/13) active space, which includes all Ti d and O lone pair orbitals, was used for both the SA-CASSCF and SEAS-CASCI calculations. Both SEAS-CASCI and the reference EOM-CCSD calculations predict that the S_1, S_2, and S_3

surfaces intersect at the T_d geometry (zero asymmetric stretching) and split slowly as symmetry is broken. In both cases, the surfaces remain smooth and continuous along the asymmetric stretching coordinate. In contrast, the three SA-CASSCF surfaces exhibit unphysical cusps at the T_d structure.

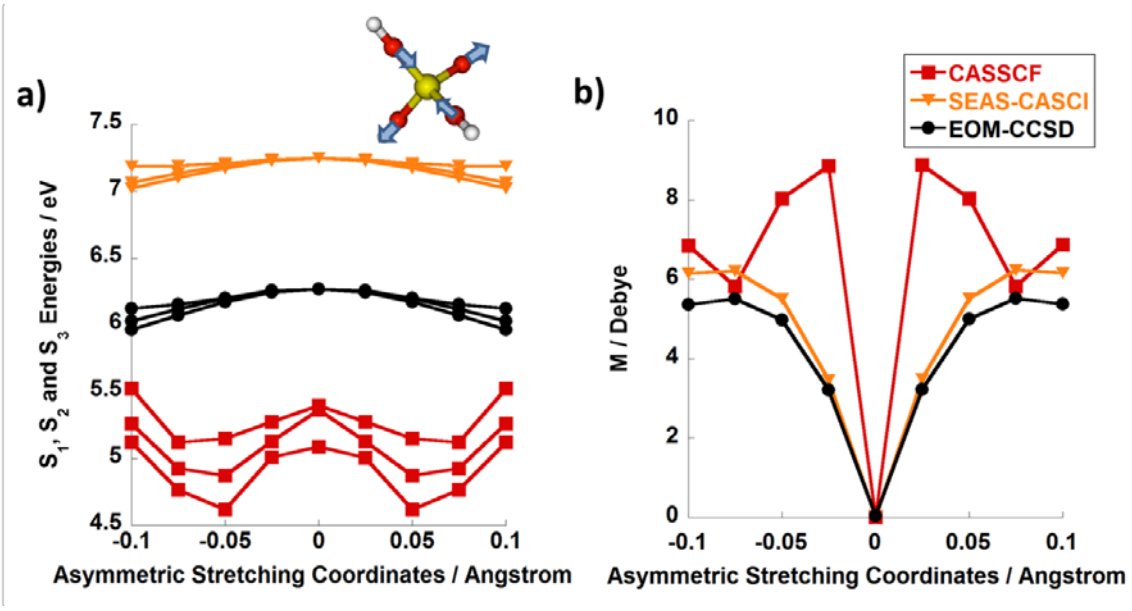


Figure 3-5. The a) PES and b) norm of the trace of the dipole moment operator over states S_1, S_2, and S_3 (M; defined in Eq.(3-4)) for the linear $Ti(OH)_4$ molecule as a function of an asymmetric stretching coordinate (illustrated in the inset). The SEAS-CASCI PES (orange) is smoothly varying and in qualitative agreement with EOM-CCSD (black), while the CASSCF surface (red) has unphysical cusps at the T_d symmetric geometry and is in poor agreement with EOM-CCSD. SEAS-CASCI predicts a dipole moment curve in good agreement with EOM-CCSD, while the CASSCF dipole moment curve changes discontinuously at the T_d geometry.

The analysis of the electronic wavefunction along the asymmetric stretching coordinate is complicated by the fact that the three lowest excited states of linear $Ti(OH)_4$ are degenerate by symmetry at the T_d structure and thus mix as a function of this coordinate. Rather than analyze the dipole moment of a single state, we consider the trace of the dipole moment operator over adiabatic states S_1, S_2, and S_3. This trace is expected to be continuous and smoothly varying as a function of nuclear geometry so long as these three states do not intersect with lower or higher electronic states. The norm of this trace, defined

$$M = \left\| \sum_{I=S_1}^{S_3} \langle I | \hat{\boldsymbol{\mu}} | I \rangle \right\| \tag{3-4}$$

is plotted as a function of the asymmetric stretching coordinate in Figure 3-5b. Here $|I\rangle$ is the wavefunction of adiabatic electronic state, *I*, and $\hat{\boldsymbol{\mu}}$ is the dipole moment operator. The dipole moments computed at the SEAS-CASCI level of theory are in excellent agreement with those predicted by EOM-CCSD, with both dipole moments varying smoothly. (The cusp at zero is due to the dipole moment vector passing through zero.) The dipole moment curve computed at the SA-CASSCF level of theory overestimates the dipole and exhibits very sharp changes near zero, suggesting discontinuities in the wavefunction attributable to artifactual spatial symmetry breaking.

Unphysical wavefunction symmetry breaking can only be clearly defined in molecules with symmetric nuclear configurations. However, the phenomena that result in such symmetry breaking (near degeneracy between configurations and a coupling matrix

element passing through zero) can exist for a wide range of nuclear geometries, and need not be required by symmetry. Just as non-Jahn-Teller conical intersections often occur in asymmetric molecules (accidental same-symmetry intersections),[73] it would not be surprising to observe symmetry-breaking-like distortions of the wavefunction of totally asymmetric molecules.

To investigate the existence of symmetry-breaking-like unphysical wavefunction distortions in asymmetric molecules, we apply the same analysis to a fully relaxed (C_0 symmetric) $Ti(OH)_4$ cluster, with results presented in Figure 3-6. Unlike the linear $Ti(OH)_4$ cluster, this system exhibits no symmetry-required orbital or state degeneracy. Both the PES and dipole moment curve computed at the SEAS-CASCI level of theory are in nearly quantitative agreement with the EOM-CCSD reference. SA-CASSCF, on the other hand, overestimates the dipole moment and predicts an unsmooth PES with an unphysically large energy gap between excited states. This suggests that the SA-CASSCF wavefunction may be suffering from symmetry-breaking-like distortions at these completely asymmetric geometries.

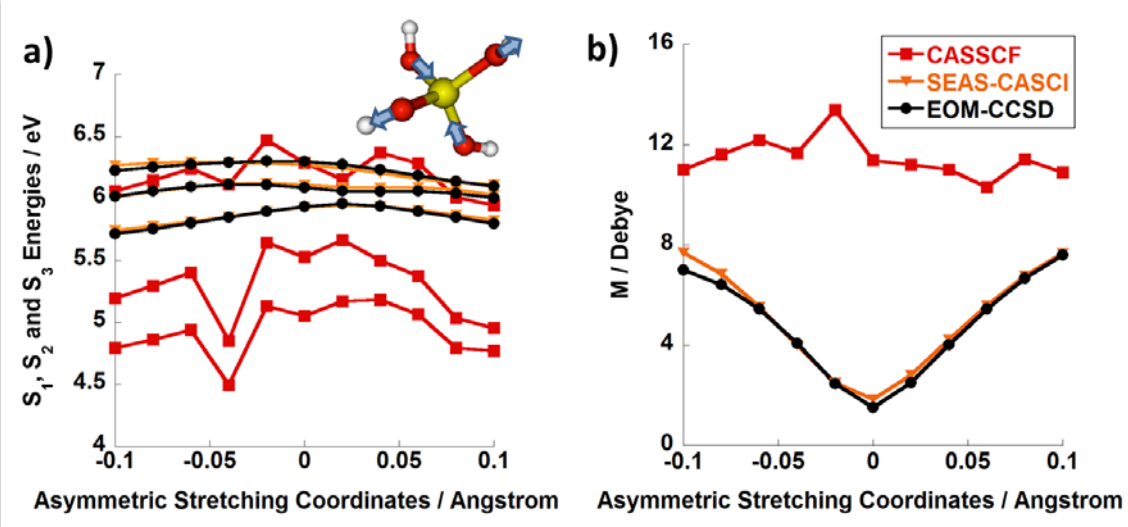


Figure 3-6. The a) PES and b) norm of the trace of the dipole moment operator over states S_1, S_2, and S_3 (*M*; defined in Eq. (3-4)) for the bent $Ti(OH)_4$ molecule as a function of an asymmetric stretching coordinate (illustrated in the inset). The SEAS-CASCI PES (orange) is smoothly varying and in near-quantitative agreement with EOM-CCSD (black), while the CASSCF surface (red) is both in poor agreement with EOM-CCSD and relatively unsmooth. SEAS-CASCI predicts a dipole moment in near-quantitative agreement with EOM-CCSD.

3.2.4 SEAS-CASCI and biradicaloid systems

Non-radiative decay mechanisms often involve biradicaloid regions of the PES—regions where two electrons are shared between two nearly degenerate molecular orbitals (MOs). In such regions three singlet configurations can be constructed with similar energies, and therefore conical intersections—points of true degeneracy between adiabatic electronic states that facilitate non-radiative decay[11]—are often present. One

reason that CASSCF has been so widely applied to these problems is that the CASSCF energy is invariant to rotations between the active orbitals, and therefore the two MOs involved in biradicaloid behavior are treated on equal footing. This is in contrast to single-reference methods based on a closed-shell RHF reference where the equations determining the orbitals are not invariant to rotations between the highest occupied (HO-) and lowest unoccupied (LU-) MOs. The SEAS orbital determination step draws a similar division between the active occupied and active unoccupied orbitals, and thus is not invariant to rotations between the frontier MOs, though the final CASCI wavefunction is invariant to such rotations. This suggests that SEAS may fail to describe the electronic wavefunction at geometries where the HOMO and LUMO become nearly degenerate (biradicaloid geometries).

To investigate the behavior of SEAS-CASCI in biradicaloid regions of the PES, we consider ethylene, a small organic molecule with a well understood non-radiative decay behavior involving a biradicaloid conical intersection.[74] Upon excitation, ethylene is believed to very quickly twist about its double bond, breaking the coupling of the π system and leading to a biradicaloid region where orthogonal p atomic orbitals on each carbon atom are nearly degenerate. Additional distortion (pyramidalization about one carbon) leads to a conical intersection which facilitates fast and efficient decay to the electronic ground state.

The PES of ethylene as a function of twisting and pyramidalization coordinates is calculated at the SEAS-CASCI and CASSCF levels of theory and presented in Figure 3-7. In

both cases, the (2/2) active space and 6-31G** basis were used, keeping with previous studies of the photodynamics of ethylene.[74] Three states were averaged in the CASSCF calculation. The minimal energy conical intersection was first optimized at both the SEAS-CASCI and CASSCF levels of theory using the multiplier penalty approach implemented in the CIOpt software package,[75] with numerical gradients employed in the case of SEAS-CASCI. Then, displacements along the twisting and pyramidalization coordinates were applied by rigid rotation of the methylene groups, leaving all other internal degrees of freedom constant.

The optimized SEAS-CASCI structure is nearly identical to the optimized CASSCF structure, as are the PESs. Though further investigation is warranted, these results suggest that SEAS-CASCI is appropriate for studying biradicaloid systems, despite the fact that the SEAS energy (Eq (1)) is not invariant to rotations between HOMO and LUMO.

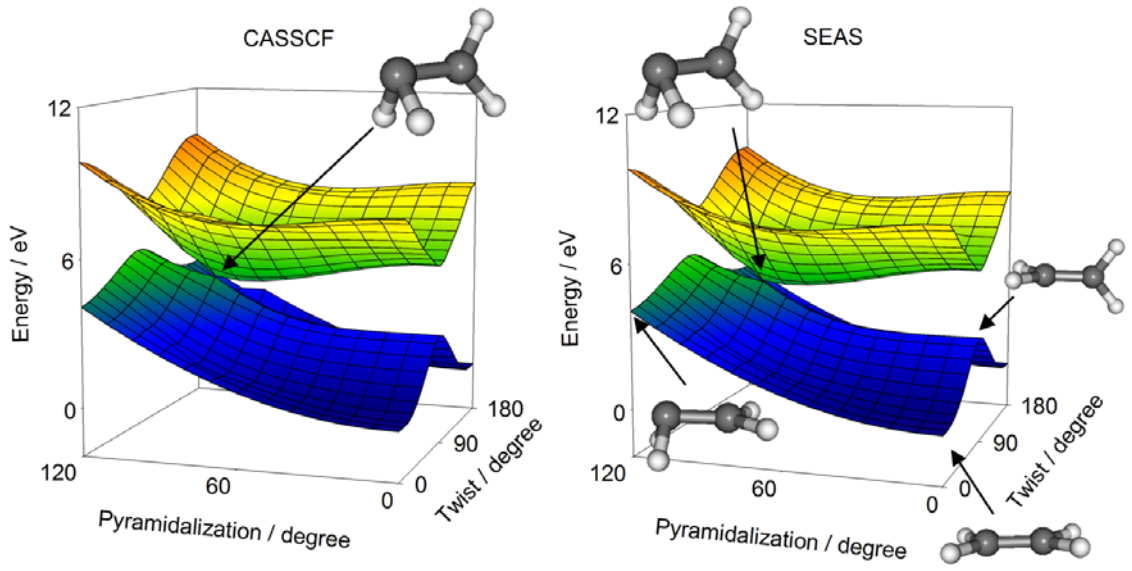


Figure 3-7. The PES of ethylene as a function of the rigid twisting and pyramidalization of fully optimized conical intersection structures (pictured at the top of each panel) calculated at the CASSCF (a) and SEAS-CASCI (b) levels of theory. The SEAS-CASCI optimized structure and PES are in excellent agreement with CASSCF, indicating the ability of SEAS-CASCI to describe biradicaloid regions of the PES, despite the lack of invariance to rotations between HOMO and LUMO in the initial SEAS SCF step.

3.3 Conclusions

A qualitatively accurate and symmetry-preserving excited state wavefunction is highly desirable in many cases, such as in AIMD, in geometry optimization, and as a reference for subsequent treatment of electron correlation via CI or perturbation theory. With these cases in mind, we developed the SEAS-CASCI approach, which decreases the propensity for unphysical wavefunction symmetry breaking by eliminating off-diagonal

Hamiltonian matrix elements from the energy expression used to variationally optimize the orbitals. Unlike existing two-step approaches, both SEAS occupied and virtual orbitals are optimized to minimize the average energies of low-lying electronic excited states.

Calculations of the low-lying excited states of a series of small molecules suggests that that symmetry breaking is less prevalent at the SEAS-CASCI level of theory than for traditional SA-CASSCF. Though the symmetry breaking behavior is still observed at the SEAS-CASCI level of theory for some choices of active space, the preservation of wavefunction symmetry can be further enforced by freezing doubly-occupied orbitals in their symmetry-preserving RHF form.

It was also shown that SEAS-CASCI reduces the artifacts of symmetry-breaking at asymmetric molecular geometries, including eliminating cusps in the PES which result from the discontinuous description of the electronic structure as a function of nuclear geometry. This suggests the possibility that symmetry-breaking-like distortions of the electronic structure can occur in asymmetric species as well, in analogy to “accidental” conical intersections in asymmetric molecules. Finally, we have demonstrated that, despite the fact that the energy optimized in the SEAS orbital determination calculation is not invariant to rotations between HOMO and LUMO, SEAS-CASCI is capable of describing the biradicaloid region of the PES of ethylene with accuracy similar to SA-CASSCF.

REFERENCES

REFERENCES

(1) Stier, W.; Prezhdo, O. V. *J. Phys. Chem. B* **2002**, *106*, 8047.

(2) Rego, L. G. C.; Batista, V. S. *J. Am. Chem. Soc.* **2003**, *125*, 7989.

(3) Abuabara, S. G.; Rego, L. G. C.; Batista, V. S. *J. Am. Chem. Soc.* **2005**, *127*, 18234.

(4) Fernandez-Alberti, S.; Kleiman, V. D.; Tretiak, S.; Roitberg, A. E. *J. Phys. Chem. A* **2009**, *113*, 7535.

(5) Long, R.; Prezhdo, O. V. *J. Am. Chem. Soc.* **2011**, *133*, 19240.

(6) Chapman, C. T.; Liang, W.; Li, X. S. *J. Phys. Chem. Lett.* **2011**, *2*, 1189.

(7) Li, J. R.; Nilsing, M.; Kondov, I.; Wang, H. B.; Persson, P.; Lunell, S.; Thoss, M. J. Phys. Chem. C 2008, 112, 12326.

(8) Foresman, J. B.; Headgordon, M.; Pople, J. A.; Frisch, M. J. J. Phys. Chem. 1992, 96, 135.

(9) Gross, E. K. U.; Kohn, W. *Phys. Rev. Lett.* **1985**, *55*, 2850.

(10) Casida, M. E.; Jamorski, C.; Casida, K. C.; Salahub, D. R. *J. Chem. Phys.* **1998**, *108*, 4439.

(11) Yarkony, D. R. *J. Phys. Chem. A* **2001**, *105*, 6277.

(12) Levine, B. G.; Ko, C.; Quenneville, J.; Martinez, T. J. *Mol. Phys.* **2006**, *104*, 1039.

(13) Roos, B. O.; Taylor, P. R. *Chem. Phys.* **1980**, *48*, 157.

(14) Lengsfield, B. H.; Saxe, P.; Yarkony, D. R. *J. Chem. Phys.* **1984**, *81*, 4549.

(15) Andersson, K.; Malmqvist, P. A.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.

(16) Tao, H. L.; Levine, B. G.; Martinez, T. J. *J Phys Chem A* **2009**, *113*, 13656.

(17) Bauschlicher, C. W.; Langhoff, S. R. *J. Chem. Phys.* **1988**, *89*, 4246.

(18) Demeras, A. S.; Lepetit, M. B.; Malrieu, J. P. *Chem Phys Lett* **1990**, *172*, 163.

(19) Zaitsevskii, A.; Malrieu, J. P. *Chem Phys Lett* **1994**, *228*, 458.

(20) Lykos, P.; Nesbet; Sinanoglu; Slater; Pratt, G. W.; Delbruck; Lowdin; Harris, G. M. Rev. Mod. Phys. 1963, 35, 496.

(21) Jackels, C. F.; Davidson, E. R. *J. Chem. Phys.* **1976**, *64*, 2909.

(22) Davidson, E. R.; Borden, W. T. *J. Phys. Chem.* **1983**, *87*, 4783.

(23) Engelbrecht, L.; Liu, B. *J. Chem. Phys.* **1983**, *78*, 3097.

(24) McLean, A. D.; Lengsfield, B. H.; Pacansky, J.; Ellinger, Y. *J. Chem. Phys.* **1985**, *83*, 3567.

(25) Scuseria, G. E.; Schaefer, H. F. *Chem. Phys. Lett.* **1987**, *142*, 354.

(26) Burton, N. A.; Yamaguchi, Y.; Alberts, I. L.; Schaefer, H. F. *J. Chem. Phys.* **1991**, *95*, 7466.

(27) Delbene, J. E.; Kim, K. S.; Shavitt, I. *Can. J. Chem.* **1991**, *69*, 246.

(28) Stanton, J. F.; Gauss, J.; Bartlett, R. J. *J. Chem. Phys.* **1992**, *97*, 5554.

(29) Barnes, L. A.; Lindh, R. *Chem. Phys. Lett.* **1994**, *223*, 207.

(30) Crawford, T. D.; Stanton, J. F.; Allen, W. D.; Schaefer, H. F. J. Chem. Phys. 1997, 107, 10626.

(31) McGrady, J. E.; Lovell, T.; Stranger, R. *Inorg. Chem.* **1997**, *36*, 3242.

(32) Ayala, P. Y.; Schlegel, H. B. *J. Chem. Phys.* **1998**, *108*, 7560.

(33) Sherrill, C. D.; Krylov, A. I.; Byrd, E. F. C.; Head-Gordon, M. J. Chem. Phys. 1998, 109, 4171.

(34) Sherrill, C. D.; Lee, M. S.; Head-Gordon, M. *Chem. Phys. Lett.* **1999**, *302*, 425.

(35) Crawford, T. D.; Stanton, J. F. *J. Chem. Phys.* **2000**, *112*, 7873.

(36) Eisfeld, W.; Morokuma, K. *J. Chem. Phys.* **2000**, *113*, 5587.

(37) Filatov, M.; Shaik, S. *Chem. Phys. Lett.* **2000**, *332*, 409.

(38) Braida, B.; Lauvergnat, D.; Hiberty, P. C. *J. Chem. Phys.* **2001**, *115*, 90.

(39) Cohen, R. D.; Sherrill, C. D. *J. Chem. Phys.* **2001**, *114*, 8257.

(40) Crawford, T. D.; Kraka, E.; Stanton, J. F.; Cremer, D. *J. Chem. Phys.* **2001**, *114*, 10638.

(41) Dunietz, B. D.; Head-Gordon, M. *J. Phys. Chem. A* **2003**, *107*, 9160.

(42) Beran, G. J. O.; Austin, B.; Sodt, A.; Head-Gordon, M. *J. Phys. Chem. A* **2005**, *109*, 9183.

(43) Sato, T.; Tokunaga, K.; Tanaka, K. *J. Chem. Phys.* **2006**, *124*, 024314.

(44) Halvick, P. *Chem. Phys.* **2007**, *340*, 79.

(45) Lawler, K. V.; Beran, G. J. O.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 024107.

(46) Sorkin, A.; Iron, M. A.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 307.

(47) Li, X. Z.; Paldus, J. *J. Chem. Phys.* **2009**, *130*, 084110.

(48) Li, X. Z.; Paldus, J. *Int. J. Quantum Chem.* **2009**, *109*, 1756.

(49) Mintz, B.; Crawford, T. D. *Phys. Chem. Chem. Phys.* **2010**, *12*, 15459.

(50) Brueckner, K. A.; Wada, W. *Phys. Rev.* **1956**, *103*, 1008.

(51) Nesbet, R. K. *Phys. Rev.* **1958**, *109*, 1632.

(52) Edgecombe, K. E.; Becke, A. D. *Chem. Phys. Lett.* **1995**, *244*, 427.

(53) Jimenez-Hoyos, C. A.; Henderson, T. M.; Scuseria, G. E. J. Chem. Theory Comput. 2011, 7, 2667.

(54) Bersuker, I. *The Jahn-Teller Effect*; Cabridge University Press: Cambridge, 2006.

(55) See Supplementary Materials for MCSCF calculations confirming PJTE-like symmetry breaking in the helium dimer cation.

(56) Potts, D. M.; Taylor, C. M.; Chaudhuri, R. K.; Freed, K. F. *J. Chem. Phys.* **2001**, *114*, 2592.

(57) Bofill, J. M.; Pulay, P. *J. Chem. Phys.* **1989**, *90*, 3637.

(58) Abrams, M. L.; Sherrill, C. D. *Chem. Phys. Lett.* **2004**, *395*, 227.

(59) Lu, Z.; Matsika, S. *J. Chem. Theory Comput.* **2012**, *8*, 509.

(60) Granucci, G.; Persico, M.; Toniolo, A. *J. Chem. Phys.* **2001**, *114*, 10608.

(61) Slavicek, P.; Martinez, T. J. *J. Chem. Phys.* **2010**, *132*, 234102.

(62) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schutz, M.; Celani, P.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Koppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; O'Neill, D. P.; Palmieri, P.; Pfluger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stol, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M.; Wolf, A. MOLPRO, version 2010.1, a package of ab initio programs, 2010.

(63) M.W.Schmidt; K.K.Baldridge; J.A.Boatz; S.T.Elbert; M.S.Gordon; J.H.Jensen; S.Koseki; N.Matsunaga; K.A.Nguyen; S.J.Su; T.L.Windus; M.Dupuis; J.A.Montgomery. J. Comput. Chem. 1993, 14, 1347.

(64) Anpo, M.; Yamashita, H.; Matsuoka, M.; Park, D. R.; Shul, Y. G.; Park, S. E. J. Ind. Eng. Chem. 2000, 6, 59.

(65) Li, G.; Dimitrijevic, N. M.; Chen, L.; Nichols, J. M.; Rajh, T.; Gray, K. A. J. Am. Chem. Soc. 2008, 130, 5402.

(66) Pople, J. A.; Binkley, J. S.; Seeger, R. *Int. J. Quantum Chem.* **1976**, *10*, 1.

(67) Harihara, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.

(68) Stanton, J. F.; Bartlett, R. J. *J. Chem. Phys.* **1993**, *98*, 7029.

(69) See Supplementary Materials for additional information on the reference EOM-CCSD calculations.

(70) See Supplementary Materials for additional information on the reference EOM-CCSD calculations.

(71) Deskevich, M. P.; Nesbitt, D. J.; Werner, H. J. *J. Chem. Phys.* **2004**, *120*, 7281.

(72) See Supplementary Materials for conformation of the existence of this cusp at the full-valence CASSCF level.

(73) Yarkony, D. R. *Rev. Mod. Phys.* **1996**, *68*, 985.

(74) Ben-Nun, M.; Quenneville, J.; Martinez, T. J. *J. Phys. Chem. A* **2000**, *104*, 5161.

(75) Levine, B. G.; Coe, J. D.; Martinez, T. J. *J Phys Chem B* **2008**, *112*, 405.

Chapter 4

Configuration Interaction Singles Natural Orbitals: An Orbital Basis for an Efficient and Size Intensive Multireference Description of Electronic Excited States

"The Best way to have a good idea is to have a lot of ideas"

-Linus C. Pauling (1901-1994)

This chapter is reproduced from paper: Configuration interaction singles natural orbitals: an orbital basis for an efficient and size intensive multireference description of electronic excited states, Y. Shu, E. G. Hohenstein, and B. G. Levine, J. Chem. Phys., 142, 024102 (2015).

Multireference electronic structure theories are widely used to model the electronic wavefunctions and potential energy surfaces of systems for which a traditional Hartree-Fock or Kohn-Sham reference is poorly suited, such as those undergoing bond breaking or non-radiative decay via conical intersection. Multiconfigurational self-consistent field (MCSCF) methods, such as the widely used complete active space self-consistent field (CASSCF) method,[1] provide a reference wavefunction ansatz flexible enough to describe the salient features of the wavefunction in these regions of the PES and can be

systematically improved by adding dynamic electron correlation via configuration interaction (CI) or perturbation theory.[2-9] Though CASSCF is extremely successful in general, its use brings challenges. The results of CASSCF calculations depend strongly and unpredictably on the choice of the space of active orbitals. When employing the state-averaged variant of CASSCF (SA-CASSCF), which is often used in the treatment of electronic excited states, the results can also depend strongly on the choice of the state average ensemble. The computational cost of CASSCF calculations scales factorially with the number of active orbitals, and thus can be significantly higher than the cost of Hartree-Fock even when a relatively modest active space is chosen. The computational cost is increased by the complex nonlinear nature of the variational solution of the CASSCF wavefunction, which requires the simultaneous optimization of the CI and orbital coefficients to self-consistency. The nonlinear nature of this optimization problem can also lead to multiple solutions, discontinuities in the PES, convergence difficulties, and unphysical spatial symmetry breaking of the wavefunction.[10-13]

The computational challenges associated with the non-linear optimization of the CASSCF wavefunction can be eased in several ways. By reducing the size of the configuration space in an intelligent way, decreased scaling with respect to active space can be achieved.[14-19] Similarly, the cost can be reduced by employing a two-step complete active space CI (CASCI) method which separates the determination of the orbitals from the solution of the CI problem. In such an approach the orbitals are determined first, and these orbitals are then used to build and diagonalize a complete

active space (CAS) Hamiltonian. Because the CI and orbital coefficients are not optimized to self-consistency, CASCI wavefunctions are variationally inferior to those computed at the CASSCF level of theory, but the computational cost associated with a CASCI calculation can be as much as an order of magnitude less than that of a similar CASSCF calculation.

Various CASCI approaches exist, differing in their orbital determination procedures. The accuracy of these methods depends strongly on this choice. Since Koopmans it has been known that canonical Hartree-Fock orbitals are optimal for describing electron attachment and detachment processes. Thus these orbitals do not provide a very efficient orbital basis for a CASCI expansion of a bond breaking or excited state system. The natural orbitals of unrestricted Hartree-Fock calculations[20] and correlated single reference methods,[21] on the other hand, form a basis that is well suited for the efficient expansion of the electronic wavefunction of a system undergoing bond breaking. Useful expansions of the low-lying excited state wavefunctions of molecules based on improved Hartree-Fock virtual orbitals (IVO),[22-27] floating occupation molecular orbitals (FOMO),[28,29] high multiplicity natural orbitals (HMNO),[30,31] and orbitals which minimize the average energy of all singly excited electronic configurations within the active space (SEAS)[13] have all been shown to provide an efficient and accurate alternative to fully optimized SA-CASSCF orbitals. The elimination of the nonlinear simultaneous optimization of CI and orbital coefficients can lead not only to a reduction in computational cost, but also to an improved qualitative description of the PES. For example, two of the authors have

recently reported that the two-step SEAS-CASCI approach reduces the propensity for unphysical spatial symmetry breaking of the wavefunction relative to SA-CASSCF, leading to a smoother description of the PES in some cases.13

In this study we investigate a novel two-step CASCI approach based on the natural orbitals of single-reference calculations of electronic excited states: the configuration interaction singles natural orbital (CISNO-) CASCI method. The state-averaged natural orbitals of configuration interaction singles calculations have long been used as an initial guess for SA-CASSCF calculations,[32] but their explicit use without further optimization has not been investigated. After detailing the CISNO-CASCI wavefunction ansatz, we report a study of the accuracy of this method, paying specific attention to vertical excitation energies of closed-shell molecules computed both with and without additional treatment of dynamic electron correlation, the appropriateness of CISNO-CASCI for the description of large systems composed of one or more excitable subsystems, and the accuracy of CISNO-CASCI near crossings between potential energy surfaces. Comparisons will be made to SA-CASSCF and other two-step CASCI methods.

4.1 Method

4.1.1 CISNO-CASCI method

In the present work, we consider only singlet systems, though CISNO-CASCI could be extended to high spin systems, in principle. The orbitals used in the CISNO-CASCI method are determined by first performing a restricted Hartree-Fock (HF) calculation and

a subsequent configuration interaction singles (CIS)[33] calculation to provide an approximation to the lowest *N* singlet excited states. The state averaged (SA-) CIS first order reduced density matrix is constructed by averaging the first order reduced density matrices of the ground (HF) and low-lying excited (CIS) states according to

$$\gamma_{\mu\upsilon}^{\text{SA-CIS}} = \frac{1}{N+1}\left(\gamma_{\mu\upsilon}^{\text{HF}} + \sum_{I=1}^{N} \gamma_{\mu\upsilon}^{\text{CIS}(I)} \right) \tag{4-1}$$

where $\gamma_{\mu\upsilon}^{\text{HF}}$ and $\gamma_{\mu\upsilon}^{\text{CIS}(I)}$ are elements of the first order reduced density matrices of the HF ground state and the *I*th CIS singlet excited state, respectively, and lowercase Greek letters index basis functions. The CIS first order reduced density matrix is defined

$$\gamma_{\mu\upsilon}^{CIS(I)} = \left\langle \Psi_{CIS(I)} \middle| a_{\mu}^{?} a_{\upsilon} \middle| \Psi_{CIS(I)} \right\rangle \tag{4-2}$$

where $\left|\Psi_{CIS(I)}\right\rangle$ is the wavefunction of the *I*th CIS excited state.

Because $\boldsymbol{\gamma}^{\text{SA-CIS}}$ represents an approximate average first order reduced density matrix for the low-lying excited states of interest, diagonalizing it provides an efficient natural orbital basis in which to expand those states. In fact, $\boldsymbol{\gamma}^{\text{SA-CIS}}$ can be thought of as an approximation to the state averaged density matrix from a SA-CASSCF calculation and its NOs as an approximation to the state averaged NOs of a SA-CASSCF ensemble. Upon diagonalization of $\boldsymbol{\gamma}^{\text{SA-CIS}}$, the natural orbitals are ordered according to their occupation numbers, and those orbitals with occupation numbers nearest to one are chosen to be the active orbitals. It is up to the user to determine reasonable cutoffs for the active space. A standard CASCI Hamiltonian (including all excitations within the active space) is then constructed and diagonalized yielding the CISNO-CASCI states and

energies. Note that, despite the fact that the orbitals are determined using a single-reference procedure, the CASCI states should provide all of the benefits of a multireference treatment of the electronic structure, as has been demonstrated in analogous CASCI methods for describing bond breaking.[20,21]

CISNO-CASCI offers several conceptual advantages over previously developed two-step CASCI methods. In the CISNO approach, both the virtual and occupied orbitals are optimized to efficiently describe the low-lying excited states of the system, unlike the IVO method in which the HF occupied orbitals remain unchanged, and thus are optimal for electron detachment. The ensemble of states used to determine the CISNOs includes only ground and low-lying excited states of a single charge state, unlike the FOMO method which is based on an ensemble that includes not only the low-lying states of a single charge state, but also high energy multiply-excited states and states which are ionized relative to the reference. Unlike the IVO, FOMO, and SEAS approaches, the CIS states used to determine CISNOs include strong electron correlation effects introduced by the mixing of different singly excited configurations, such as are present in the L_a and L_b states of benzene. However, these conceptual advantages come at a significant practical cost; the determination of energy gradients at the CISNO-CASCI level of theory requires solution of the coupled perturbed CI equations for the CIS step. Thus in the present work only energy calculations and geometry optimizations performed with numerical gradients are reported.

It is noteworthy that a relaxed density matrix, rather than the one-electron

reduced density matrix used here, is employed in the computation of properties and gradients at the CIS level of theory.[33] The relaxed density matrix is the density matrix corresponding to a CIS energy that is stationary to first-order with respect to changes in both the CI vector as well as the molecular orbitals; with the one electron reduced density matrix, the CIS energy is only stationary with respect to changes in the CI vector. We have also attempted to employ this relaxed density matrix in CISNO-CASCI calculations, and notable differences between the results will be discussed below.

4.1.2 Computational details

With an exception noted below, CISNOs were determined using the CIS and matrix operation (MATROP) subprograms of the MolPro 2010 electronic structure package[34] to generate and diagonalize $\gamma^{\text{SA-CIS}}$. Determination of the CISNOs is followed by computation of the final CISNO-CASCI energies and wavefunctions, which is done by construction and diagonalization of a CASCI Hamiltonian using the MCSCF function in MolPro. CISNO-CASCI calculations presented in section C were instead conducted using a GPU implementation of this method in the TeraChem software package,[35] to be detailed in another publication.

We have tested this method on a wide range of molecules with interesting excited state properties. Organic molecules in our test set include pyrazine ($C_4N_2H_4$), thymine ($C_5N_2O_2H_6$, a DNA base), pentadiene iminium ($C_5NH_8^+$, a retinal analog), and benzene (C_6H_6). Inorganic compounds in our test set include trihydroxy(oxo) vanadium ($VO(OH)_3$, a cluster model of a single-site photocatalyst),[36] ferrocyanide ($Fe(CN)_6^{4-}$, a molecule

capable of photoejection of electrons),[37-41] tetracyanonickelate ($Ni(CN)_4^{2-}$, a molecule with unusual pressure sensitive spectroscopic properties)[40,42,43], iron pentacarbonyl ($Fe(CO)_5$, an inorganic compound with well-studied photodissociation)[40] and sila-adamantane ($Si_{10}H_{16}$, a cluster model of the silicon surface). Note that in the present work we limit our test set to molecules with closed-shell ground states, but we will investigate the applicability of CISNO-CASCI to open-shell systems in a future study.

All geometries are optimized at the MP2 theoretical level with the 6-31G** basis set. Except when noted, all CASSCF and CASCI calculations were done with the cc-pVTZ basis set for transition metal atoms[44] and 6-31G** for all non-transition metal atoms. In this work, the active spaces were chosen with reference to the occupation numbers of the CISNOs, and the numbers of states to be averaged in SA-CASSCF and CISNO-CASCI calculations were chosen based on the results of preliminary equation of motion coupled clusters singles and doubles (EOM-CCSD)[45] calculations performed with the same basis. In principle, the user of CISNO-CASCI could choose the active space and state average by any of the various approaches that are standard in the field. The largest computationally convenient active space has been chosen for all systems. Multistate complete active space second order perturbation theory (MS-CASPT2)[5] results serve as a reference against which to judge the accuracy of the CASCI methods. Similar multistate Rayleigh-Schrodinger perturbation theory calculations were performed with the CISNO-CASCI reference (CISNO-CASCI-RS2) to evaluate the utility of the CISNO-CASCI wavefunction as a reference for higher level correlated methods. Smaller active spaces are chosen for the

perturbative calculations for computational convenience.

We abbreviate our chosen active spaces (n_{elec} / n_{orb}), where n_{elec} is the number of active electrons and n_{orb} is the number of active orbitals. SA-CASSCF, SEAS-CASCI, canonical HF orbital (HF-) CASCI, EOM-CCSD, MS-CASPT2, and CISNO-CASCI-RS2 calculations were performed in MolPro,[6,34,46-49] while IVO- and FOMO-CASCI calculations were performed in GAMESS.[50,51] Except where noted, SA-CASSCF and SEAS-CASCI calculations presented here were performed with CISNOs as the initial orbital guess.

4.2 Results and Discussion

4.2.1 Accuracy of vertical excitation energies

To gauge the accuracy of the CISNO-CASCI method, we calculated the vertical excitation energies of the lowest-lying excited states of our test set molecules at the CISNO-CASCI and SA-CASSCF theoretical levels, comparing to energies computed at the highly accurate MS-CASPT2 level. Results are reported in Table 4-1. Adiabatic state labels (S_1, S_2, and so on) have been assigned according to the ordering of states as computed at the SA-CASSCF level of theory, and the states computed at other levels of theory are labeled so that states of the same chemical character share the same adiabatic state label, and therefore may not be ordered by energy. Orbitals and CI vectors for these states are presented in Tables S1-S18 of Supplementary Materials.[53] The general chemical character of all states is also reported in Table 4-1, with metal centered, metal-to-ligand charge transfer, and ligand-to-metal charge transfer states abbreviated MC,

MLCT, and LMCT, respectively.

Both SA-CASSCF and CISNO-CASCI typically overestimate excitation energies relative to MS-CASPT2, with mean signed errors (averaged over all computed excited states) of 0.59 and 0.47 eV, respectively. It is well known that SA-CASSCF overestimates vertical excitation energies due to the absence of dynamics correlation, and it is not surprising that CISNO-CASCI exhibits similar behavior, with only a slightly smaller propensity to overestimate the excitation energies. Mean unsigned errors of 0.61 and 0.49 eV for SA-CASSCF and CISNO-CASCI suggest that SA-CASSCF and CISNO-CASCI are similarly accurate when compared with MS-CASPT2.

In most cases SA-CASSCF and CISNO-CASCI provide very similar excitation energies, but $Ni(CN)_4^{2-}$ and $Fe(CN)_6^{4-}$ are notable exceptions. For $Ni(CN)_4^{2-}$, CISNO-CASCI underestimates the MS-CASPT2 vertical excitation energies of the lowest-lying states by 0.06-0.25 eV, while SA-CASSCF overestimates them by 1.07-1.30 eV. We note that by reducing the size of the active space from (14/11) to (6/7) (see Table 4-2), the large errors in the SA-CASSCF vertical excitation energies can be reduced, suggesting that the larger active space provides an unbalanced treatment of electron correlation that favors the S_0 state. Similar behavior is observed for $Fe(CN)_6^{4-}$. It is noteworthy that CISNO-CASCI avoids this pitfall. The larger active space SA-CASSCF calculations may suffer because the inclusion of additional orbitals in the active space leads to stronger correlation in the ground state than in the excited states, and thus to the unpredictable dependence of excitation energies on active space. Because CISNOs are computed from a wavefunction

ansatz with a more limited ability to describe electron correlation, it may be that this procedure is less prone to the inclusion of active orbitals which strongly correlate a single state. The present data is not enough to draw such a conclusion, but this is an interesting hypothesis for future investigation.

To investigate the suitability of CISNO-CASCI as a reference for a more accurate treatment of dynamic electron correlation, we applied Rayleigh-Schrodinger perturbation theory on top of the CISNO-CASCI wave function. Table 4-2 lists the excitation energies of the lowest-lying excited states of our model molecules computed at the CISNO-CASCI, CISNO-CASCI-RS2, SA-CASSCF, and MS-CASPT2 levels of theory, with the corresponding active space as well as number of correlated orbitals. Note that for computational convenience smaller active spaces were used for these calculations than for the CASSCF and CISNO-CASCI calculations reported in Table 4-1. When gas phase experimental results are available, they are listed for comparison.

Table 4-1. Excitation energies of the low-lying excited states of the test set molecules computed at the state averaged CASSCF and CISNO-CASCI levels of theory. The active spaces and numbers of states averaged for these calculations are also presented. The number of states averaged includes the HF ground state in the case of CISNO-CASCI. MS-CASPT2 results are shown for comparison. The smaller active spaces and numbers of correlated orbitals used for the MS-CASPT2 calculations are shown in Table 4-2. Adiabatic state labels are assigned according to the state ordering at the CASSCF level of theory. The labels of the states computed by all other methods are assigned by comparison of chemical character (orbitals and CI vectors) to those of CASSCF. The error is defined as the difference between the excitation energy computed at the CASSCF or CISNO-CASCI level of theory, and that computed at the MS-CASPT2 level.

Molecule	Active Space	States Averaged	CASSCF State	State Character	CASSCF results / eV	Error / eV	CISNO-CASCI results / eV	Error / eV	MS-CASPT2 / eV
Pyrazine	14/10	4	S_1	n→π*	4.574	0.400	4.566	0.392	4.174
			S_2	π→π*	4.952	0.718	6.158[a]	1.924	4.234
			S_3	n→π*	5.783	0.200	6.013	0.430	5.583

Table 4-1 (cont'd)

Molecule	Active Space	States Averaged	CASSCF State	State Character	CASSCF results / eV	Error / eV	CISNO-CASCI results / eV	Error / eV	MS-CASPT2 / eV
Thymine	10/8	3	S_1	n→π*	5.000	-0.310	5.761	0.451	5.310
			S_2	π→π*	6.943	0.900	7.164	1.121	6.043
Pentadiene	10/9	4	S_1	π→π*	4.765	0.616	5.232	1.083	4.149
Iminium			S_2	π→π*	5.587	0.027	6.103	0.543	5.560
			S_3	σ→π*	7.314	0.375	7.661	0.722	6.939
Benzene	16/13	3	S_1	π→π*	5.051	0.075	5.133	0.157	4.976
			S_2	π→π*	7.989	1.389	7.656	1.056	6.600
$VO(OH)_3$	16/13	4	S_1	LMCT	4.747	0.173	4.851	0.277	4.574
			S_2	LMCT	4.859	0.150	4.950	0.241	4.709
			S_3	LMCT	4.859	0.150	4.950	0.241	4.709
$Fe(CN)_6^{4-}$	18/14	4	S_1	MC	5.197	1.540	3.896	0.239	3.657[b]
			S_2	MC	5.197	1.583	3.897	0.283	3.614
			S_3	MC	5.197	1.267	3.897	-0.033	3.930[b]

Table 4-1 (cont'd)

Molecule	Active Space	States Averaged	CASSCF State	State Character	CASSCF results / eV	Error / eV	CISNO-CASCI results / eV	Error / eV	MS-CASPT2 / eV
$Ni(CN)_4^{2-}$	14/11	4	S_1	MLCT	5.202	1.071	4.074	-0.057	4.131
			S_2	MLCT	5.211	1.08	4.075	-0.056	4.131
			S_3	MLCT	5.440	1.3	3.882	-0.258	4.140
[c]$Fe(CO)_5$	14/14	5	S_1	MC	5.210	0.238	5.241[c]	0.269	4.972
			S_2	MC	5.646	0.388	5.548[c]	0.29	5.258
			S_3	MC	5.754	0.519	5.549[c]	0.314	5.235
			S_4	MC	6.062	0.645	5.823[c]	0.406	5.417
[b,c]Sila-adamantane	16/14	5	S_1	σ→σ*	7.171	0.222	7.998[c]	1.049	6.949
			S_2	σ→σ*	7.341	0.244	7.637[c]	0.54	7.097[b]
			S_3	σ→σ*	7.629	0.485	7.637[c]	0.493	7.144
			S_4	σ→σ*	7.671	0.472	7.637[c]	0.438	7.199[b]

a. This state is the fourth excited state at the CISNO-CASCI level.
b. These MS-CASPT2 states have mixed character compared to SA-CASSCF.
c. These CISNO-CASCI states have mixed characters compared to SA-CASSCF.

Table 4-2. Excitation energies of the low-lying excited states of the test set molecules computed at the CISNO-CASCI, CISNO-CASCI-RS2, SA-CASSCF and MS-CASPT2 theoretical levels. Note that smaller active spaces are used than in Table 4-1 for computational convenience. All adiabatic state labels are assigned according to the characters of the CASSCF states in Table 4-1. Where possible, gas phase experimental data has been presented for comparison.

Molecule	Active Space	# of correlated orbitals	CASSCF State	CISNO-CASCI / eV	CISNO-CASCI-RS2 / eV	SA-CASSCF / eV	MS-CASPT2 / eV	Exptl results / eV
Pyrazine	8/7	15	S_1	4.903	3.906	4.618	4.174	3.83[77]
			S_2	6.162	4.588	5.801	4.234	
			S_3	6.303	5.443	6.033	5.583	
Thymine	8/7	19	S_1	5.791	4.825	4.883	5.310	4.95[78]
			S_2	7.166	5.438	6.954	6.043	
Pentadiene Iminium	6/7	16	S_1	5.280	3.986	4.775	4.149	
			S_2	6.113	5.356	5.589	5.560	
			S_3	8.211	6.342	7.955	6.936	
Benzene	10/10	15	S_1	5.105	5.043	5.048	4.976	4.9[79]
			S_2	7.661	6.658	8.025	6.600	
$VO(OH)_3$	6/8	16	S_1	4.902	4.342	4.130	4.575	
			S_2	4.999[a]	4.384[a]	4.242[a]	4.709[a]	
			S_3	5.000[a]	4.385[a]	4.242[a]	4.709[a]	
$Fe(CN)_6^{4-}$	6/5	16	S_1	3.519	3.744[a]	3.425	3.657[a]	

Table 4-2 (cont'd).

Molecule	Active Space	# of correlated orbitals	CASSCF State	CISNO-CASCI / eV	CISNO-CASCI-RS2 / eV	SA-CASSCF / eV	MS-CASPT2 / eV	Exptl results / eV
$Fe(CN)_6^{4-}$	6/5	16	S_2	3.519	3.607	3.426	3.614	
			S_3	3.519	3.947[a]	3.426	3.930[a]	
$Ni(CN)_4^{2-}$	6/7	16	S_1	4.065	4.011	4.347	4.131	
			S_2	4.066	4.012	4.348	4.131	
			S_3	3.922	4.178	4.210	4.140	
$Fe(CO)_5$	4/6	15	S_1	4.498[a]	5.044[a]	4.244[a]	4.972[a]	
			S_2	5.003[a]	5.312[a]	4.728[a]	5.235[a]	
			S_3	5.004[a]	5.449[a]	4.731[a]	5.258[a]	
			S_4	5.315[a]	5.575[a]	5.040[a]	5.417[a]	
Sila-adamantane	6/7	15	S_1	8.350[a]	6.838[a]	7.560[a]	6.949[a]	
			S_2	8.224[a]	6.868[a]	7.648[a]	7.097[a]	
			S_3	8.224[a]	6.865[a]	7.640[a]	7.144[a]	
			S_4	8.224[a]	6.871[a]	7.708[a]	7.199[a]	

a. These states are of mixed character relative to the CASSCF states reported in Table 4-1.

The CISNO-CASCI-RS2 excitation energies are, on average, 0.13 eV below those computed at the MS-CASPT2 level of theory. This is consistent with the fact that CISNO-CASCI-predicted excitation energies are, on average, slightly lower than those predicted by SA-CASSCF. The mean unsigned difference between the excitation energies computed at the MS-CASPT2 and CISNO-CASCI-RS2 levels of theory is 0.21 eV. These

very small differences, which are on the order of the errors expected for MS-CASPT2 itself, suggest that CISNO-CASCI-RS2 is a reliable alternative to CASPT2. The excellent agreement between the CISNO-CASCI-RS2 results and the available experimental data confirm the reliability of this approach.

The discussion above is limited to states which are singly excited relative to the closed-shell ground state. Some systems, such as the linear polyenes, exhibit low-lying doubly excited states of photochemical significance. Given that CISNOs arise from the description of singly excited states only, it is reasonable to wonder whether CISNO-CASCI can accurately describe these states. With this question in mind, we have employed SA-CASSCF and CISNO-CASCI to compute the vertical excitation energy to the doubly excited (2^1A_g) state of trans-1,3-butadiene. We use a four electron/four orbital active space, an average over three states, and the 6-31g** basis. In the case of CISNO-CASCI, several σ orbitals were removed from the active space by hand in favor of π orbitals. The vertical excitation energy to the doubly excited state is found to be 6.67 and 6.82 eV at the SA-CASSCF and CISNO-CASCI levels of theory, respectively. The coefficient of the dominant doubly excited configuration (in which the LUMO is doubly occupied and the HOMO is unoccupied) is 0.57 in both cases. The application of multireference configuration interaction with single and double excitations (MRCISD) from these respective references results in vertical excitation energies of 6.75 and 6.80 eV. Thus, relative to SA-CASSCF, CISNO-CASCI provides a similarly accurate description of this doubly excited state.

4.2.2 Comparison of variational efficiency of orbitals

A desirable set of orbitals for a CASCI expansion would provide fast and efficient convergence towards the exact ground and low-lying excited state energies with increasing active space size. CI expansions based on natural orbitals (NOs), first proposed by Löwdin,[52] are known to exhibit such fast convergence in other contexts. Thus we expect the CISNO-CASCI approach, which is based on the NOs of the low-lying excited states of interest, to converge more rapidly to the exact energies of the ground and low-lying excited states than CASCI approached based on other sets of orbitals. As such, here we compare the average absolute energies of the lowest *N* states computed at the SA*N*-CASSCF, CISNO-CASCI, IVO-CASCI, and SEAS-CASCI levels of theory, with results presented in Figure 4-1. By definition, SA-CASSCF provides the variationally optimal solution for each active space, and thus we judge the remaining CASCI calculations by how closely their energies approach that of SA-CASSCF.

In most cases the CISNO-CASCI method gives a lower average absolute energy relative to the other tested two-step methods, especially when employed with large active spaces. Note that for the largest active spaces studied, CISNO-CASCI is variationally superior to all other CASCI methods in eight out of nine molecules in our test set. Pentadiene iminium is the sole exception, with IVO-CASCI demonstrating very slight variational superiority for the (10/9) active space. Considering smaller active spaces, CISNO-CASCI is variationally superior or nearly the same as the other CASCI methods for all active spaces tested in five out of nine cases, spanning both organic and inorganic systems: thiamine, benzene,

ferrocyanide, tetracyanonickelate, and iron pentacabonyl.

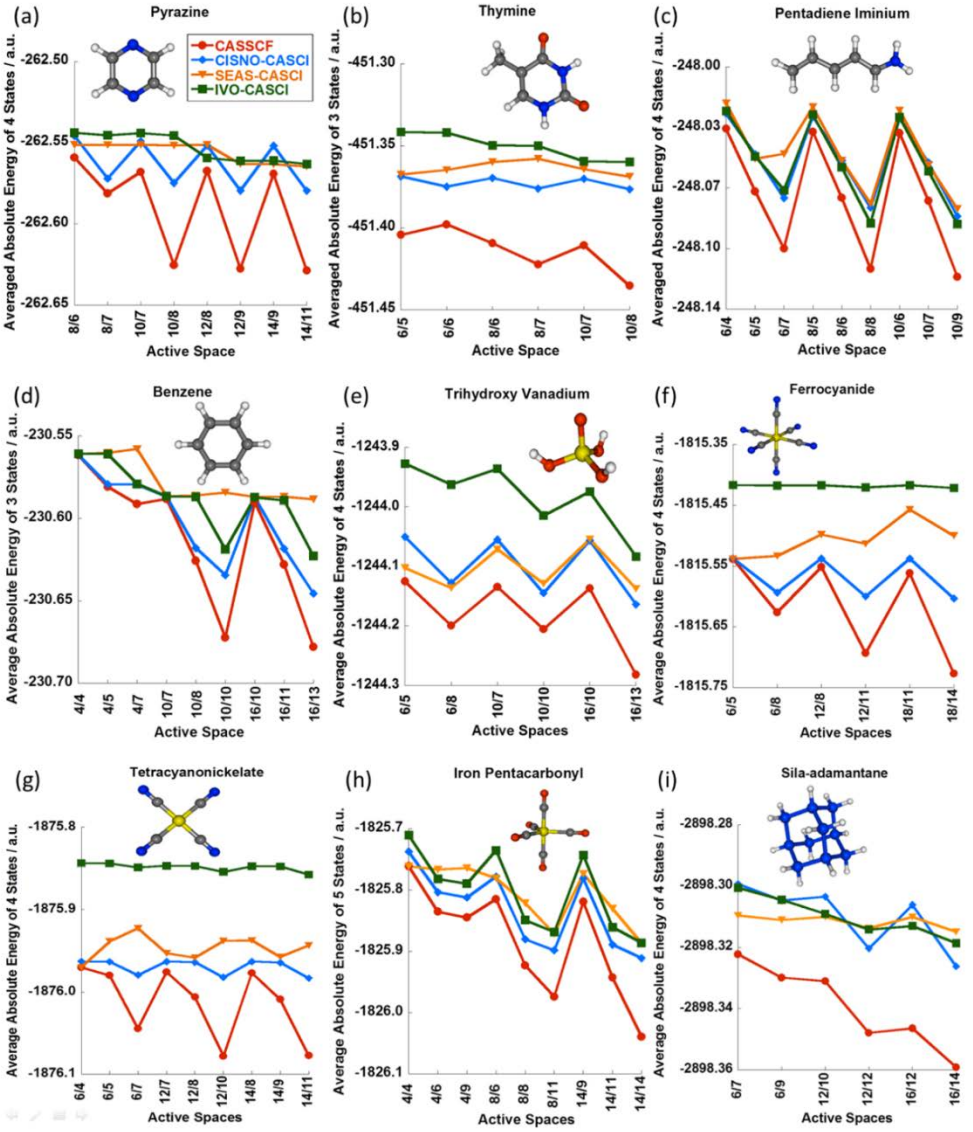


Figure 4-1. The state-averaged absolute energies of (a) pyrazine, (b) thymine, (c) pentadiene iminium, (d) benzene, (e) trihydroxy(oxo) vanadium, (f) ferrocyanide, (g) tetracyanonickelate, (h) iron pentacarbonyl, and (i) sila-adamantane molecules computed at the SA-CASSCF, CISNO-CASCI, SEAS-CASCI, and IVO-CASCI theoretical levels as a function of active space.

The oscillatory behavior observed in many of the graphs in Figure 4-1 is due to the ordering of the active spaces on the x-axis. For example, consider the CISNO-CASCI curve of Figure 4-1h. The first three active spaces all contain the same two active occupied orbitals, where by active occupied orbitals we mean those active orbitals which would be occupied in the Hartree-Fock configuration. The next three active spaces contain four active occupied orbital, while the final three contain seven active occupied orbitals. Within each group of three, the first active space contains two active virtual orbitals, the second contains four, and the third contains seven. Remember that these orbitals are identical regardless of active space in the case of the two-step methods. Thus, going from (4/4) to (4/6) we are adding the same two orbitals to the active space as when we go from (8/6) to (8/8) and from (14/9) to (14/11). The fact that the total energy decreases by roughly 0.1 a.u. in all three cases indicates that these two orbitals make an important contribution to the total energy in this case. Adding the next set of three virtual orbitals to the active space (e.g. increasing from (4/6) to (4/9)) has a considerably smaller effect on the total energy, suggesting that these orbitals are less important, at least in a variational sense. This visualization of the change in the total energy as a function of active space size can be helpful in identifying the most important orbitals to include in the active space for a particular system.

It is also noteworthy that all of the SEAS-CASCI and SA-CASSCF calculations presented above employed CISNOs as initial guess orbitals. Figures S2 and S3 in the Supplementary Materials show the same state-averaged energies of our test set

molecules computed at the SA-CASSCF and SEAS-CASCI levels of theory as a function of active space with two different initial guesses: CISNOs and canonical HF orbitals.[53] When the results differ, the use of CISNOs as the initial guess generally results in a variationally superior solution compared with HF. In fact, in some cases the CISNO-CASCI variational energy itself is lower than that of SA-CASSCF calculations utilizing Hartree-Fock orbitals as the initial guess, even though SA-CASSCF is, by definition, expected to be variationally superior. These results support the popular practice of using CISNOs as an initial guess for SA-CASSCF calculations to reduce the probability of the variational optimization becoming stuck in a local minimum.[32]

We have also investigated the dependence of the variational energy and the lowest excitation energy on the number of states averaged over at the CISNO-CASCI level of theory. Small but noticeable differences are observed, with variational energies varying by as much as 1 eV and excitation energies by up to 0.3 eV. (See Figures S4 and S5 for details.[53])

4.2.3 Description of localized excited states

A CASCI expansion is particularly useful for describing spatially localized electronic excited states. Such a situation arises anytime the excited states under consideration are centered on a single molecule or chromophore and are decoupled from the surrounding environment. Some examples of localized excited states include dye molecules in solution, the chromophores of many photoactive proteins, and excitable defects in materials. Here, we will consider unsaturated hydrocarbons (specifically, ethylene, butadiene, hexatriene,

and benzene) microsolvated by 250 methanol molecules. In this case, an accurate treatment would result in low-lying excited states which are spatially localized on the unsaturated molecule. Methods that are invariant to rotations of occupied orbitals and rotations of virtual orbitals, such as CIS (or, similarly, time-dependent density functional theory (TDDFT)[54,55] and EOM-CC[45]), will find these excited states and the excited states will localize, naturally, on the unsaturated molecule. A method based on a CASCI expansion, however, can only describe these localized excited states if the active orbitals are localized on the chromophore to begin with. The orbital optimization procedure in SA-CASSCF will localize the orbitals on the chromophore, but orbital delocalization can become problematic for methods where the orbital optimization is decoupled from the optimization of the CI coefficients. This problem is exacerbated if the orbital optimization scheme contains no information about the excited electronic states under consideration. The worst case is HF-CASCI, where the orbitals are often so delocalized that they are not even suitable as a starting guess for SA-CASSCF. If a one-electron density matrix describes localized electronic states, then the corresponding natural orbitals tend to localize on the chromophore. This implies that CISNO-CASCI should be a suitable method for describing localized excitations.

To illustrate the locality of SA-CASSCF (state-averaged natural) orbitals and CISNOs, we plot the π^* orbital of ethylene in a cluster of 250 methanol molecules (see Figure 4-2a and b). To the eye, both π^* orbitals are identical and localized on the ethylene molecule. For comparison, canonical HF virtual orbitals are also shown in Figure 4-2c, d, and e; the π^*

orbital of ethylene mixes with the other virtual orbitals and, in this case, three virtual orbitals of the cluster contain significant contributions from the π* orbital of ethylene. In order to perform HF-CASCI with a 2/2 active space containing the π and π* orbitals, a 2/4 active space would be required in practice if the canonical HF orbitals are used (so that the entire orbital can be included). Additionally, while the LUMO of isolated ethylene computed at the HF level corresponds to the π* orbital, the LUMO of the microsolvated system does not have π* character. Instead, pieces of it are scattered across the LUMO+10, LUMO+11, and LUMO+12 of the ethylene/methanol cluster, and thus ten lower energy virtual orbitals would need to be removed from the active space to efficiently describe the π→π* excitation. The well-known shortcomings of HF virtual orbitals are thus exaggerated when explicit solvent is included in the computation. CISNOs provide a photochemically motivated rotation of the canonical HF orbitals that pick out the important contributions to the low-lying excited states. In Table 4-3, we present the excitation energies of our test systems computed with CISNO-CASCI/6-31G* and SA-CASSCF/6-31G*. The differences between the two methods for these clusters are comparable to those shown in Tables 4-1 and 4-2; this shows that the quality of the CISNO-CASCI excitation energies are comparable to those of SA-CASSCF for these much larger systems. It is also worth noting that these clusters contain roughly 1500 atoms and 9500 basis functions, demonstrating that the CISNO-CASCI method can be applied to extended systems.

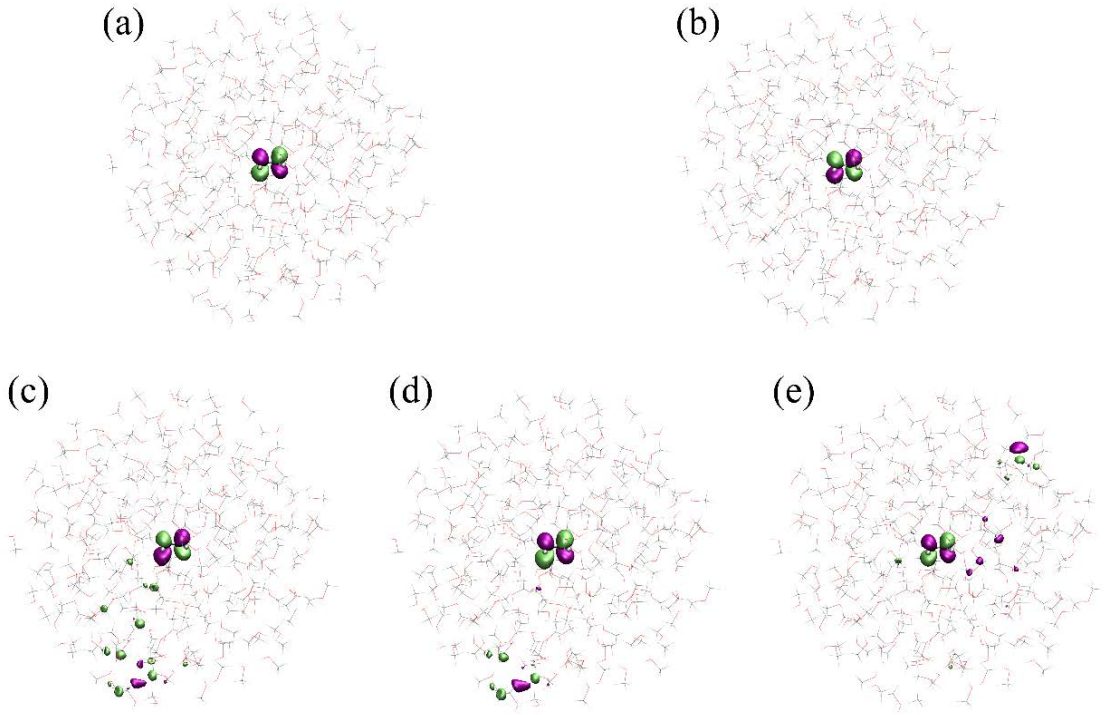


Figure 4-2. The π* orbital of ethylene microsolvated by 250 methanol molecules plotted for different orbital choices: (a) a SA-2-CASSCF natural orbital, (b) a CISNO averaged over the ground state and one excited state, (c) a HF virtual orbital LUMO+10 of the cluster, (d) a HF virtual orbital LUMO+11 of the cluster, (e) a HF virtual orbital LUMO+12 of the cluster. The CISNO and SA-CASSCF orbitals are plotted with an isovalue of 0.01. The HF virtual orbitals are plotted with an isovalue of 0.025. All computations described in this Figure were performed in a 6-31G basis set.

Table 4-3. Excitations energies computed at the SA-CASSCF/6-31G* and CISNO-CASCI/6-31G* levels of theory for ethylene, butadiene, hexatriene, and benzene microsolvated by 250 methanol molecules. The SA-CASSCF computations average over the ground state and excited states reported below. The CISNOs include only the ground state and lowest excited state in the averaging. The active spaces include all π and π* orbitals located on the chromophore.

Molecule	Active Space	CASSCF State	CISNO-CASCI / eV	SA-CASSCF / eV
Ethylene	2/2	S_1	10.19	9.88
Butadiene	4/4	S_1	7.37	6.95
		S_2	8.28	8.59
Hexatriene	6/6	S_1	6.49	5.90
		S_2	8.22[a]	7.02
		S_3	6.99[b]	7.83
Benzene	6/6	S_1	5.10	5.03
		S_2	8.15	8.29

a. This is the third excited state at the CISNO-CASCI level.
b. This is the second excited state at the CISNO-CASCI level.
c. This is the third excited state at the CISNO-CASCI level.

4.2.4 Size intensivity and size consistency

With recent advances in computing technology, applying multireference quantum chemical methods to nanoscale systems has become realistic. Therefore, it is important to consider the scaling of computed energies with system size. In particular, we are interested in the size dependence of the excitation energy in systems composed of multiple excitable subsystems, like conjugated polymers, organic semiconductors, or defective nanoparticles. It may be surprising to discuss this topic, as the complete active space CI expansion is rigorously size extensive,[56] and much work has been done in the application of cluster approximations to limit configuration spaces used in active space calculations without losing this property.[57-61] It is important to note, however, that the scaling of energies with system size depends not only on the CI or cluster expansion employed, but also on the choice of orbitals. As such, we will investigate whether SA-CASSCF and the various CASCI methods exhibit three properties: ground state size consistency, excited state size consistency, and size intensivity.[62,63]

First we define these terms as they are used here. Consider two non-interacting systems, *A* and *B*, which can be in either the ground or excited state. Systems in the excited state will be denoted by *. As is well known, a method, *X*, is ground state size consistent if the energy for the union of *A* and *B* in their ground states computed by method *X*, $E_X(A \cup B)$ is equal to the sum of the energies of *A* and *B* calculated independently by method *X*,

$$E_X(A \cup B) = E_X(A) + E_X(B) \tag{4-3}$$

Similarly, a method is considered to be excited state size consistent if the same property holds when both *A* and *B* are in their excited state,

$$E_X(A^* \cup B^*) = E_X(A^*) + E_X(B^*) \tag{4-4}$$

Finally, a method is size intensive when the excitation energy of one subsystem, *A*, does not depend on the presence of an additional non-interacting subsystem, *B*, in its ground state.

$$E_X(A^* \cup B) - E_X(A \cup B) = E_X(A^*) - E_X(A) \tag{4-5}$$

Note that in our tests and proof we consider only systems such that subsystems *A* and *B* are singlet systems representable by closed-shell HF wavefunctions in their reference states. This leads to a weak definition of size consistency, which usually requires that non-interacting open-shell systems (such as the non-interacting H atoms of a fully dissociated H_2 molecule) also obey Eq. (4-3).[63] In fact, by the standard we employ in this work, RHF is considered ground state size consistent, though by the most widely used definition of size consistency, it is not.

To assess which CASCI methods are size intensive, we applied several two-step methods as well as SA-CASSCF to compute the lowest excitation energies of systems composed of a varying number of non-interacting H_2 and ethylene molecules. We have computed the excitation energies of monomer-localized singly excited states at the SA-CASSCF, HF-CASCI, SEAS-CASCI, CISNO-CASCI, IVO-CASCI, and FOMO-CASCI theoretical levels as a function of the number of non-interacting molecules. EOM-CCSD, which is known to be size

intensive,[45] provides a reference for comparison. A size intensive method will show no change in the excitation energy as a function of the number of non-interacting molecules. Results for sets of non-interacting H_2 molecules are shown in Figure 4-3a and b, while those for sets of non-interacting ethylene molecules are presented in Figure 4-3c. In all cases, we use an active space of (2*M*/2*M*) where *M* is the number of non-interacting molecules in the system. We choose to average over (*M*+1) states in the SA-CASSCF and CISNO-CASCI calculations. More discussion of this choice and several alternatives will follow presentation of the results.

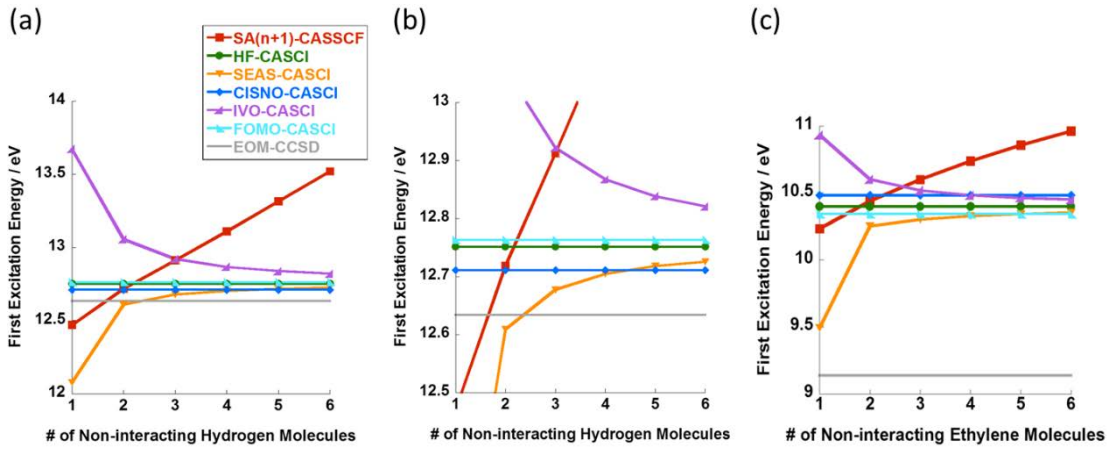


Figure 4-3. (a) The excitation energies of the monomer-localized singly excited states of systems of non-interacting H_2 molecules computed at the SA-CASSCF, HF-CASCI, SEAS-CASCI, CISNO-CASCI, IVO-CASCI, FOMO-CASCI, and EOM-CCSD theoretical levels as a function of the number of non-interacting subsystems are shown in red, green, orange, blue, purple, cyan, and gray, respectively. (b) Same as (a), but with a narrower energy scale for clarity. (c) Same as (a), but for systems of non-interacting ethylene molecules.

As can be seen in Figure 4-3, the CISNO-, FOMO-, and HF-CASCI methods are size intensive. In contrast, the SEAS- and IVO-CASCI methods are not. Notably, the SA-CASSCF excitation energies depend strongly on the number of non-interacting systems, as well. Though the scaling of the SA-CASSCF excitation energy appears nearly linear with the number of non-interacting subsystems in Figure 4-3, further analysis shows that it is actually asymptotically approaching the excitation energy that one would obtain by using the orbitals from a state-specific (SS-) CASSCF calculation of the ground state to compute CASCI excited states. These orbitals are optimized for description of static electron correlation, and therefore yield a poor description of electronic excitation, resulting in the unphysically large excitation energies observe for larger *M*.

It is easy to explain how this behavior arises. The state-average ensemble of (*M*+1) states contains one state in which all monomers are in their electronic ground states, and *M* states in which a single monomer is in its electronic excited state and the remaining (*M*-1) monomers are in their ground states. The percentage of states in the ensemble in which any particular monomer is in its excited state is therefore $1/(M+1)$, which decreases rapidly to zero with increasing *M*. Thus, as *M* increases, the orbitals approach those that are optimal for describing the ground state of each monomer.

One may argue that this problem can be fixed by increasing the size of the state-average ensemble to include multiply-excited states corresponding to the simultaneous excitation of multiple monomers, but it is not generally possible to restore size intensivity this way. One can, for example, average over 2^n states, which could in principle include

all states in which each monomer is either in its ground or first excited state. However, this ensemble is not achieved in practice, because states involving charge transfer between monomers often have lower energies than the states that involve the simultaneous single excitation of multiple monomers. These charge transfer states pollute the state average ensemble, leading, again, to a lack of size intensivity. This is shown in practice in Figure S6, which presents the excitation energy of systems of *M* non-interacting H_2 molecules for different state averaging schemes.[53]

To our knowledge, no general, size intensive scheme for state averaging can be devised in which all states are equally weighted. As described in the Supplementary Materials, however, state averaging schemes in which different weights are assigned to the ground and excited states of the system can restore size intensivity.[53]

However, these schemes may make use of negative weights for the ground state, and in general, it is not possible to accurately describe the PES in the vicinity of intersections between states with different weights, so we do not consider this possibility further here. Active space decomposition approaches[19] that eliminate the offending charge transfer states may offer another route to solve this problem in the context of MCSCF. Alternatively, it may be possible to devise a dynamic state weighting scheme[64,65] which can simultaneously yield size intensive excitation energies and a proper description of conical intersections.

Another interesting result of the calculations in Figure 4-3 is that two of the two-step CASCI methods strongly favor charge transfer between monomers. Most of the

studied methods predict the monomer-localized singly excited states to be the lowest excited states (S_1) of the total system, but for FOMO-CASCI and IVO-CASCI, the lowest-lying excited states involve charge transfer between monomers. Therefore in these two cases the energies ploted in Figure 4-3 are not those of S_1 but of higher excited states corresponding to monomer-local singly excited states. A priori arguments can be made that both of these methods should be expected to favor charge transfer. In FOMO-CASCI, the orbitals are chosen to minimize the energy of an ensemble of states that includes not only excited states with the same charge as the reference, but also states that are ionized relative to the reference. In IVO-CASCI, the occupied RHF canonical orbitals remain unchanged, and it is well known that these orbitals are optimal for electron detachment. Thus, one would expect both IVO- and FOMO-CASCI to favor states in which the individual monomers are ionized relative to the reference, while SA-CASSCF, SEAS-CASCI, and CISNO-CASCI would not be expected to favor such states.

We have also investigated the ground state size consistency of the various methods. We have computed the ground state energies of the same systems of *M* non-interacting H_2 and ethylene molecules, and in Tables 4-4 and 4-5 we present the ratio

$$y = E_X(MA) / E_X(A) \tag{4-6}$$

where *MA* represents a system of *M* non-interacting copies of molecule *A*. For a ground state size consistent method, *y* should be equal to *M*. Not surprisingly, the three methods which exhibit size intensivity, CISNO-CASCI, HF-CASCI, and FOMO-CASCI, also exhibit ground state size consistency, while the remaining methods do not.

Table 4-4. Absolute ground state energy factors (*y* as defined in Eq. (4-6)) computed at the SA-CASSCF, HF-CASCI, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, and FOMO-CASCI theoretical levels for a series of systems comprising different numbers of non-interacting hydrogen molecules.

Number of Non-Interacting Hydrogen Molecules	SA -CASSCF / a.u.	HF -CASCI / a.u.	CISNO -CASCI / a.u.	SEAS -CASCI / a.u.	IVO -CASCI / a.u.	FOMO -CASCI / a.u.
1	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
2	2.006812	2.000000	2.000001	2.02761	1.993232	2.000000
3	3.015771	3.000000	3.000002	3.043491	2.986634	3.000000
4	4.026894	4.000000	4.000000	4.058893	3.980035	4.000000
5	5.039824	5.000000	5.000000	5.074142	4.973437	5.000000
6	6.054101	6.000000	6.000000	6.089331	5.966838	6.000000

Table 4-5. Absolute ground state energy factors (*y* as defined in Eq. (4-6)) computed at the SA-CASSCF, HF-CASCI, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, and FOMO-CASCI theoretical levels for a series of systems comprising different numbers of non-interacting ethylene molecules.

Number of Non-Interacting Ethylene Molecules	SA -CASSCF / a.u.	HF -CASCI / a.u.	CISNO -CASCI / a.u.	SEAS -CASCI / a.u.	IVO -CASCI / a.u.	FOMO -CASCI / a.u.
1	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
2	2.000080	2.000000	2.000000	2.000571	1.999937	2.000000
3	3.000185	3.000000	3.000000	3.000883	2.999606	3.000000
4	4.000304	4.000000	4.000000	4.001193	3.999275	4.000000
5	5.000432	5.000000	5.000000	5.001502	4.998944	5.000000
6	6.000564	6.000000	6.000000	6.001810	5.998613	6.000000

In addition to size intensivity and size consistency, CISNO-CASCI is also excited state size consistent. Figure 4-4 shows the excitation energies of monomers and non-interacting dimers of hydrogen and ethylene. The reported excitation energies of the monomers correspond to the singly excited state, and those of the dimer correspond to the doubly excited state in which each non-interacting subsystem is singly excited. It can be seen that the excitation energy of the dimer is double that of the monomer, i.e.

$$E_X(A^* \cup A'^*) - E_X(A \cup A') = 2\left(E_X(A^*) - E_X(A)\right) \tag{4-7}$$

Given that CISNO-CASCI is ground state size consistent, this condition is equivalent to Eq. (4-4), and thus these results demonstrate excited state size consistency.

CISNO-CASCI has the above discussed desirable properties because the CISNOs of a system do not depend on the presence of additional non-interacting subsystems. We will now prove this by considering the form of the SA-CIS first order reduced density matrix for a system comprising *M* excitable non-interacting systems. For simplicity, in the below we consider only systems composed of identical non-interacting excitable subsystems, though these arguments can be extended to systems composed of non-identical subsystems. As noted above, we only consider systems in which the ground states of the non-interacting subsystems are describable by a closed-shell RHF wavefunctions.

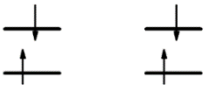
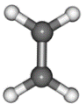
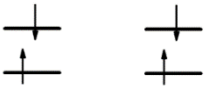
System	Represented Excited State	Excitation Energy / eV
Hydrogen Molecule Monomer 		12.71
Hydrogen Molecule Dimer (non-interacting) 		25.42
Ethylene Molecule Monomer 		10.48
Ethylene Molecule Dimer (non-interacting) 		20.96

Figure 4-4. Excitation energies of the singly excited state of monomers of hydrogen and ethylene, and of the doubly excited states of non-interacting dimers of these systems in which each monomer is singly excited. Notably, the excitation energies of the dimers are exactly twice those of the monomers, indicating excited state size consistency.

Let us first define our notation. For a method, *X*, the reduced first order density matrix of the entire system (comprising *M* subsystems) will be written $\gamma^{X,M}$. The block of $\gamma^{X,M}$ associated with subsystem *m* will be abbreviated $\gamma^{X,m,M}$. Note that the superscript *M* indicates that $\gamma^{X,m,M}$ is represented in the basis of the full system encompassing all *M* subsystems. The density matrix of a single subsystem in the absence of any non-interacting subsystem is abbreviated $\gamma^{X,m,M=1}$. Note that for HF, for

example, $\boldsymbol{\gamma}^{X,m,M=1}$ is a diagonal block of $\boldsymbol{\gamma}^{X,M}$, but this is only the case because the HF wavefunction of a subsystem is not affected by the presence of additional non-interacting subsystems. When *X* is CIS, the density matrix will be abbreviated $\boldsymbol{\gamma}^{CIS(m',I),M}$, where m' is the index of the excited subsystem, and *I* is the index of the excited state of that subsystem. Consider the SA-CIS first order reduced density matrix elements for a single subsystem:

$$\gamma_{\mu\upsilon}^{\text{SA-CIS},m,M=1} = \frac{1}{N+1}\left(\gamma_{\mu\upsilon}^{\text{HF},m,M=1} + \sum_{I=1}^{N} \gamma_{\mu\upsilon}^{\text{CIS}(m'=m,I),m,M=1} \right) \tag{4-8}$$

Now we will construct the SA-CIS density matrix of the full system of *M* non-interacting subsystems. Represented in the full basis of the system (including all subsystems), the HF and CIS density matrices for the electrons of a single subsystem are written

$$\gamma_{\mu\upsilon}^{X,m,M} = \begin{cases} \gamma_{\mu\upsilon}^{X,m,M=1} & \text{if } \mu \text{ and } \upsilon \text{ are both associated with subsystem } m \\ 0 & \text{otherwise} \end{cases} \tag{4-9}$$

where *X* represented either HF or CIS(m',*I*). Note that equation (9) is the root of the size consistency of method *X*, and does not necessarily hold for methods which are not size consistent. We can construct the HF reduced one-electron density matrix of the entire system from those of the *M* subsystems according to

$$\gamma_{\mu\upsilon}^{\text{HF},M} = \sum_{m=1}^{M} \gamma_{\mu\upsilon}^{\text{HF},m,M} \tag{4-10}$$

Similarly, the CIS reduced density matrices for the entire system can be constructed

according to

$$\gamma_{\mu\upsilon}^{\mathrm{CIS}(m',I),M} = \gamma_{\mu\upsilon}^{\mathrm{CIS}(m',I),m=m',M} + \sum_{m\neq m'} \gamma_{\mu\upsilon}^{\mathrm{HF},m,M} \qquad (4\text{-}11)$$

Illustrations of the definitions in equations (10) and (11) for the case M = 2 are presented in Figure 4-5. Note that Eq. (4-11) is true because the CIS wavefunction itself does not depend on the presence of non-interacting subsystems, so long as each subsystem is representable by a closed-shell RHF reference. Note also that both of the above density matrices are block diagonal, with all elements for which μ and υ reside on different monomers equal to zero.

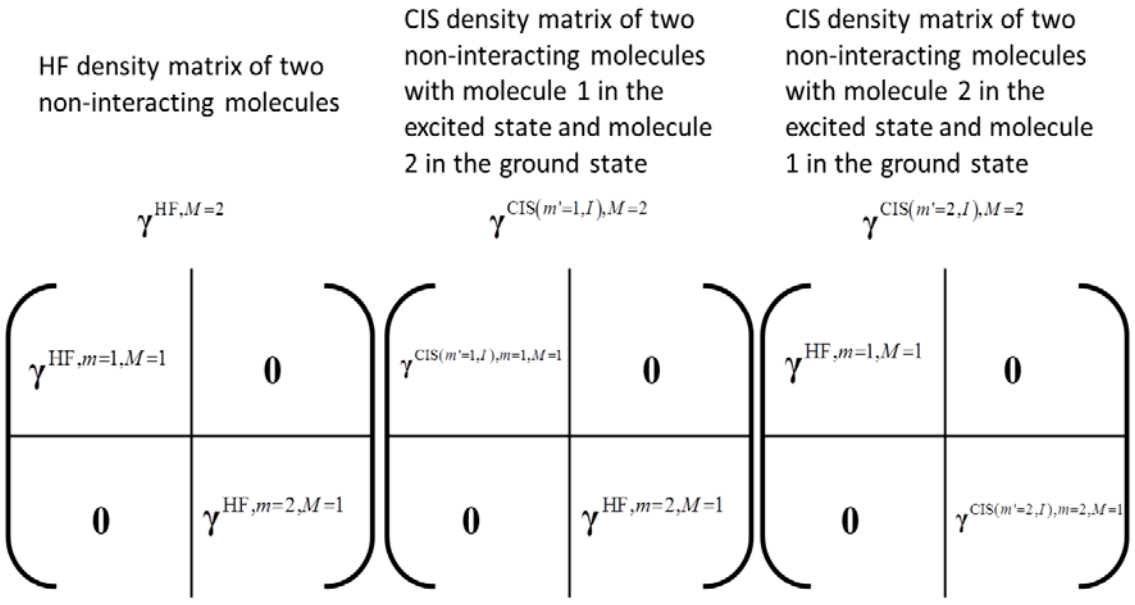


Figure 4-5. Illustration of the density matrices defined in equations (10) and (11) for case M = 2.

The SA-CIS density matrix of the full system can now be constructed according to:

$$\gamma_{\mu\upsilon}^{\text{SA-CIS},M} = \frac{1}{MN+1}\left(\gamma_{\mu\upsilon}^{\text{HF},M} + \sum_{m'=1}^{M}\sum_{I=1}^{N}\gamma_{\mu\upsilon}^{\text{CIS}(m',I),M}\right) \tag{4-12}$$

Using equations (9), (10), (11), and (12), each diagonal block of $\boldsymbol{\gamma}^{\text{SA-CIS},m,M}$ can be written

$$\gamma_{\mu\upsilon}^{\text{SA-CIS},m,M} = \frac{1}{MN+1}\left(\left((M-1)N+1\right)\gamma_{\mu\upsilon}^{\text{HF},m,M=1} + \sum_{I=1}^{N}\gamma_{\mu\upsilon}^{\text{CIS}(m'=m,I),m,M=1}\right) \tag{4-13}$$

$$\gamma_{\mu\upsilon}^{\text{SA-CIS},m,M} = \frac{1}{MN+1}\left((M-1)N\gamma_{\mu\upsilon}^{\text{HF},m,M=1} + (N+1)\gamma_{\mu\upsilon=1}^{\text{SA-CIS},m,M=1}\right) \tag{4-14}$$

Because the eigenvectors of a block diagonal matrix are equal to the eigenvectors of the individual blocks, we can prove that the presence of non-interacting subsystems does not change the CISNOs by showing that the one-electron reduced density matrices of the isolated subsystem, $\boldsymbol{\gamma}^{\text{SA-CIS},m,M=1}$ as defined in Eq. (4-8), and that of this subsystem in the presence of the remaining non-interacting subsystems, $\boldsymbol{\gamma}^{\text{SA-CIS},m,M}$ as defined in Eq. (4-13), are the same. Equivalently, we will show that $\boldsymbol{\gamma}^{\text{SA-CIS},m,M=1}$ and $\boldsymbol{\gamma}^{\text{SA-CIS},m,M}$ commute.

Because $\boldsymbol{\gamma}^{\text{SA-CIS},m,M}$ can be represented as a linear combination of $\boldsymbol{\gamma}^{\text{SA-CIS},m,M=1}$ and $\boldsymbol{\gamma}^{\text{HF},m,M=1}$, proving the necessary commutation relation requires

only proving that $\gamma^{\text{SA-CIS},m,M=1}$ and $\gamma^{\text{HF},m,M=1}$ commute. Similarly, because $\gamma^{\text{SA-CIS},m,M=1}$ can be expressed as a linear combination of $\gamma^{\text{HF},m,M=1}$ and $\gamma^{\text{CIS}(m'=m,I),m,M=1}$, we need only show that $\gamma^{\text{HF},m,M=1}$ and $\gamma^{\text{CIS}(m'=m,I),m,M=1}$ commute, which is trivial; when represented in the basis of HF canonical orbitals, the virtual-occupied (VO) block of both matrices is zero, the occupied-occupied block of $\gamma^{\text{HF},m,M=1}$ is twice the identity matrix, and thus commutes with any matrix, and the virtual-virtual block of $\gamma^{\text{HF},m,M=1}$ is the zero matrix, and thus also commutes with any matrix. Thus $\gamma^{\text{HF},m,M=1}$ will commute with any density matrix for which the VO block is zero, and $\gamma^{\text{CIS}(m'=m,I),m,M=1}$ has this property. Therefore, the CISNOs of a given subsystem do not depend on the presence of non-interacting subsystems. Numerical evidence of this is presented in Tables S19, S20, and S21.[53]

Note that the above argument is premised on two facts: that there are no coherences between blocks of the density matrix corresponding to non-interacting subsystems (Eq. (4-9)) and that the VO block is zero. Thus, a size intensive/consistent CASCI method can be developed from the natural orbitals of the state averaged density matrix of any blackbox method which meets these criteria, e.g. TDDFT or the random phase approximation.[66-71] Note also that the VO block of the relaxed CIS density matrix used to compute CIS gradients and properties is not zero, and thus using the relaxed CIS density matrix in place of the reduced one-electron density matrix does not result in a size intensive or size consistent method. This is why we have chosen to use the reduced one-electron density matrix for the CISNO-CASCI method. However, anecdotally we find that

the use of the relaxed density matrix provides more accurate results for some charge transfer states, as shown in the case of titanium (IV) hydroxide in Figure S7 of Supplementary Materials.[53] That orbital relaxation is extremely important for accurately describing charge transfer states with CIS has recently been demonstrated.[72,73] For most systems, though, differences between the excitation energies computed using the relaxed and unrelaxed density matrices are small, as can be seen in Table S22.[53] What discrepancies are present can be alleviated by including orbital relaxation and dynamical correlation into the wave function through multireference CI or perturbation theory.

4.2.5 CISNO-CASCI and biradicaloid conical intersections.

Conical intersections often arise in biradicaloid regions of the PES, where two electrons are shared between two nearly degenerate molecular orbitals.[74] Thus the accurate description of such regions is essential to the proper theoretical treatment of photochemistry. One reason that CASSCF is widely used in computational photochemistry is that, unlike single reference methods such as restricted Hartree-Fock, the CASSCF energy expression is invariant to the rotations of active orbitals. Though the same is true of CASCI, the equations that determine the CISNOs do not treat the RHF occupied and virtual spaces on the same footing, and thus there is some reason to be concerned that CISNO-CASCI may not be an appropriate method for describing biradicaloid photochemistry.

Thus, we investigate the ability of CISNO-CASCI to describe the biradicaloid region of the PES of ethylene, which undergoes non-radiative decay via a well-known conical

intersection: the twisted pyramidalized intersection.[75] Following excitation, ethylene is known to twist about its double bond and subsequently pyramidalize about one of its carbon atoms. Figures 4-6a and b shows the PESs of ethylene as a function of rigid twisting and pyramidalization coordinates computed at the SA-CASSCF and CISNO-CASCI theoretical levels, respectively. These two surfaces are centered at minimal energy conical intersections (MECI) which have been optimized at the SA-CASSCF and CISNO-CASCI theoretical levels. These calculations use the 6-31G** basis set and (2/2) active space, and the MECI optimizations were performed using MolPro in conjunction with the CIOpt software package.[76]

The SA-CASSCF and CISNO-CASCI PESs are qualitatively very similar, with the SA-CASSCF- and CISNO-CASCI-optimized MECIs 5.88 and 5.58 eV above the energy of the ground state minimum, respectively. The optimized MECI geometries are also very similar. Four of five bond lengths differ by 0.03 Å or less in the optimized MECI structures, with one of the C-H bond lengths being different by 0.10 Å. Angles differ by 1-10º, but the structures remain qualitatively very similar. These results demonstrate that CISNO-CASCI is well-suited to describe the biradicaloid region of the PES in this case, though further investigation of conical intersections in larger systems is underway.

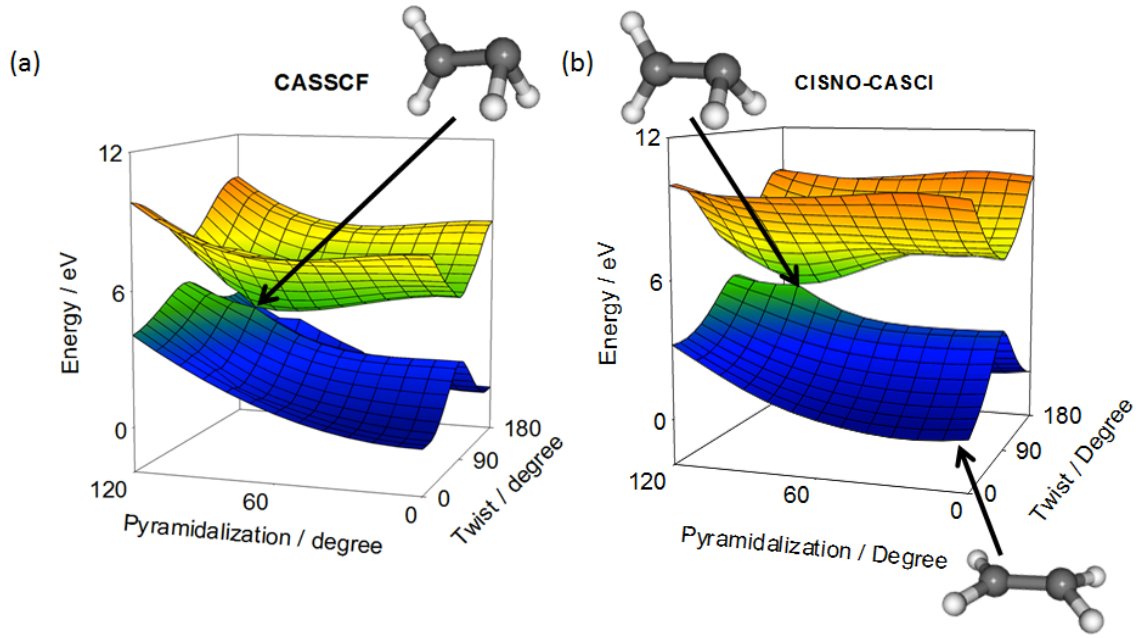


Figure 4-6. PESs of ethylene as a function of rigid twisting and pyramidalization coordinates computed at the (a) SA-CASSCF and (b) CISNO-CASCI theoretical levels. The intersection points in (a) and (b) reflect the MECIs optimized at their respective levels of theory.

4.3 Conclusions

We have shown that CIS natural orbitals can serve as an efficient orbital basis from which to build a CASCI representation of the low-lying excited states of both organic and inorganic molecules. The computed vertical excitation energies are in most cases very similar to those obtained by the variationally superior SA-CASSCF method. When discrepancies exist, they can be alleviated by inclusion of dynamic electron correlation via perturbation theory. We have also shown that CISNO-CASCI provides wavefunctions which are often variationally superior to those computed via other two-step CASCI

methods and that CISNO-CASCI is appropriate for describing biradicaloid photochemistry. The applicability of CISNO-CASCI to large chemical systems was also investigated. It was found that CISNO-CASCI describes localized excitation in systems comprising a single excitable molecule in a large sphere of explicit solvent with accuracy similar to that of SA-CASSCF. Additionally, we have explored the challenges of treating systems with multiple excitable subsystems at the SA-CASSCF level of theory. For such systems, it is possible that no state averaging scheme exists in which all states are weighted equally and the resulting vertical excitation energies are size intensive. CISNO-CASCI, on the other hand, is found to provide vertical excitation energies which are size intensive and size consistent. Along with its decreased computational cost compared with SA-CASSCF, this method is an appealing choice for treating larger, nanoscale systems composed of multiple excitable units, such as light harvesting systems, surface-modified nanocrystals, and conjugated polymers.

REFERENCES

REFERENCES

(1) Roos, B. O.; Taylor, P. R. *Chem. Phys.* **1980**, *48*, 157.

(2) Hirao, K. *Chem. Phys. Lett.* **1992**, *190*, 374.

(3) Andersson, K.; Malmqvist, P. A.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.

(4) Nakano, H. *J. Chem. Phys.* **1993**, *99*, 7983.

(5) Finley, J.; Malmqvist, P. A.; Roos, B. O.; Serrano-Andres, L. *Chem. Phys. Lett.* **1998**, *288*, 299.

(6) Celani, P.; Werner, H. J. *J. Chem. Phys.* **2000**, *112*, 5546.

(7) Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J. P. J. Chem. Phys. 2001, 114, 10252.

(8) Khait, Y. G.; Song, J.; Hoffmann, M. R. *J. Chem. Phys.* **2002**, *117*, 4133.

(9) Shiozaki, T.; Gyorffy, W.; Celani, P.; Werner, H. J. *J. Chem. Phys.* **2011**, *135*.

(10) Bauschlicher, C. W.; Langhoff, S. R. *J. Chem. Phys.* **1988**, *89*, 4246.

(11) Demeras, A. S.; Lepetit, M. B.; Malrieu, J. P. *Chem. Phys. Lett.* **1990**, *172*, 163.

(12) Zaitsevskii, A.; Malrieu, J. P. *Chem. Phys. Lett.* **1994**, *228*, 458.

(13) Shu, Y.; Levine, B. G. *J. Chem. Phys.* **2013**, *139*, 074102.

(14) Olsen, J.; Roos, B. O.; Jorgensen, P.; Jensen, H. J. A. *J. Chem. Phys.* **1988**, *89*, 2185.

(15) Nakano, H.; Hirao, K. *Chem. Phys. Lett.* **2000**, *317*, 90.

(16) Ivanic, J. *J. Chem. Phys.* **2003**, *119*, 9364.

(17) Ma, D. X.; Manni, G. L.; Gagliardi, L. *J. Chem. Phys.* **2011**, *135*.

(18) Manni, G. L.; Ma, D. X.; Aquiliante, F.; Olsen, J.; Gagliardi, L. J. Chem. Theory Comput.

2013, 9, 3375.

(19) Parker, S. M.; Seideman, T.; Ratner, M. A.; Shiozaki, T. *J. Chem. Phys.* **2013**, *139*.

(20) Bofill, J. M.; Pulay, P. *J. Chem. Phys.* **1989**, *90*, 3637.

(21) Abrams, M. L.; Sherrill, C. D. *Chem. Phys. Lett.* **2004**, *395*, 227.

(22) Hunt, W. J.; Goddard, W. A. *Chem. Phys. Lett.* **1969**, *3*, 414.

(23) Huzinaga, S.; Arnau, C. *Phys. Rev. A* **1970**, *1*, 1285.

(24) Huzinaga, S.; Arnau, C. *J. Chem. Phys.* **1971**, *54*, 1948.

(25) Beebe, N. H. F. *Int. J. Quantum Chem.* **1979**, *15*, 589.

(26) Potts, D. M.; Taylor, C. M.; Chaudhuri, R. K.; Freed, K. F. *J. Chem. Phys.* **2001**, *114*, 2592.

(27) Chattopadhyay, S.; Chaudhuri, R. K.; Freed, K. F. *J. Phys. Chem. A* **2011**, *115*, 3665.

(28) Granucci, G.; Persico, M.; Toniolo, A. *J. Chem. Phys.* **2001**, *114*, 10608.

(29) Slavicek, P.; Martinez, T. J. *J. Chem. Phys.* **2010**, *132*, 234102.

(30) Lu, Z.; Matsika, S. *J. Chem. Theory Comput.* **2012**, *8*, 509.

(31) Lu, Z.; Matsika, S. *J. Phys. Chem. A* **2013**, *117*, 7421.

(32) Blancafort, L.; Ogliaro, F.; Olivucci, M.; Robb, M. A.; Bearpark, M. J.; Sinicropi, A. Computational Investigation of Photochemical Reaction Mechanisms. In Computational Methods in Photochemistry; Kutateladze, A., Ed.; CRC Press: Boca Raton, 2005; pp 31.

(33) Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J. J. Phys. Chem. 1992, 96, 135.

(34) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schuetz, M. WIREs Comput. Mol. Sci. 2012, 2, 242.

(35) Ufimtsev, I. S.; Martinez, T. J. *J. Chem. Theory Comput.* **2009**, *5*, 2619.

(36) Anpo, M.; Yamashita, H.; Matsuoka, M.; Park, D. R.; Shul, Y. G.; Park, S. E. *J. Ind. Eng.*

Chem. **2000**, *6*, 59.

(37) Shirom, M.; Tomkiewicz, M. *J. Chem. Phys.* **1972**, *56*, 2731.

(38) Shirom, M.; Stein, G. *J. Chem. Phys.* **1971**, *55*, 3372.

(39) Shirom, M.; Stein, G. *J. Chem. Phys.* **1971**, *55*, 3379.

(40) Adamson, A. W.; Waltz, W. L.; Zinato, E.; Watts, D. W.; Fleischauer, P. D.; Lindholm, R. D. Chem. Rev. 1968, 68, 541.

(41) De Angelis, F.; Tilocca, A.; Selloni, A. *J. Am. Chem. Soc.* **2004**, *126*, 15024.

(42) Mantz, Y. A.; Musselman, R. L. *Inorg. Chem.* **2002**, *41*, 5770.

(43) Osathaphan, K.; Ruengruehan, K.; Yngard, R. A.; Sharma, V. K. Water Air Soil Poll. 2013, 224, 1647.

(44) Balabanov, N. B.; Peterson, K. A. *J. Chem. Phys.* **2005**, *123*.

(45) Stanton, J. F.; Bartlett, R. J. *J. Chem. Phys.* **1993**, *98*, 7029.

(46) Knowles, P. J.; Werner, H. J. *Chem. Phys. Lett.* **1985**, *115*, 259.

(47) Werner, H. J.; Knowles, P. J. *J. Chem. Phys.* **1985**, *82*, 5053.

(48) Werner, H. J. *Mol. Phys.* **1996**, *89*, 645.

(49) Korona, T.; Werner, H. J. *J. Chem. Phys.* **2003**, *118*, 3006.

(50) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. J. Comput. Chem. 1993, 14, 1347.

(51) Gordon, M. S.; Schmidt, M. W. Advances in electronic structure theory: GAMESS a decade later. In Theory and Applications of Computational Chemistry: the First Forty Years; Dykstra, C. E., Frenking, G., Kim, K. S., Scuseria, G. E., Eds.; Elsevier: Amsterdam, 2005; pp 1167.

(52) Lowdin, P. O. *Phys. Rev.* **1955**, *97*, 1474.

(53) See Supplementary Material Document for natural orbitals, CI vectors, select orbital coefficients, CISNO-CASCI energies computed using the relaxed density matrix, SA-CASSCF energies computed with different initial guesses, and CISNO-CASCI and SA-CASSCF results computed with different state average ensembles. For information on Supplementary Material, see http://www.aip.org/pubservs/epaps.hteml

(54) Gross, E. K. U.; Kohn, W. *Phys. Rev. Lett.* **1985**, *55*, 2850.

(55) Casida, M. E.; Jamorski, C.; Casida, K. C.; Salahub, D. R. *J. Chem. Phys.* **1998**, *108*, 4439.

(56) Simons, J. *J. Phys. Chem.* **1989**, *93*, 626.

(57) Krylov, A. I.; Sherrill, C. D.; Byrd, E. F. C.; Head-Gordon, M. J. Chem. Phys. 1998, 109, 10669.

(58) Li, X. Z.; Paldus, J. *J. Chem. Phys.* **2008**, *128*.

(59) Li, X. Z.; Paldus, J. *J. Chem. Phys.* **2008**, *129*.

(60) Li, X. Z.; Paldus, J. *J. Chem. Phys.* **2008**, *129*.

(61) Parkhill, J. A.; Lawler, K.; Head-Gordon, M. *J. Chem. Phys.* **2009**, *130*.

(62) Koch, H.; Jensen, H. J. A.; Jorgensen, P.; Helgaker, T. *J. Chem. Phys.* **1990**, *93*, 3345.

(63) Duch, W.; Diercksen, G. H. F. *J. Chem. Phys.* **1994**, *101*, 3018.

(64) Deskevich, M. P.; Nesbitt, D. J.; Werner, H. J. *J. Chem. Phys.* **2004**, *120*, 7281.

(65) Glover, W. J. *J. Chem. Phys.* **2014**, *141*, 171102.

(66) Langreth, D. C.; Perdew, J. P. *Phys. Rev. B* **1977**, *15*, 2884.

(67) Furche, F. *Phys. Rev. B* **2001**, *64*.

(68) Miyake, T.; Aryasetiawan, F.; Kotani, T.; van Schilfgaarde, M.; Usuda, M.; Terakura, K. Phys. Rev. B 2002, 66.

(69) Furche, F.; Van Voorhis, T. *J. Chem. Phys.* **2005**, *122*.

(70) Marini, A.; Garcia-Gonzalez, P.; Rubio, A. *Phys. Rev. Lett.* **2006**, *96*.

(71) Ren, X.; Tkatchenko, A.; Rinke, P.; Scheffler, M. *Phys. Rev. Lett.* **2011**, *106*.

(72) Liu, X. L.; Fatehi, S.; Shao, Y. H.; Veldkamp, B. S.; Subotnik, J. E. J. Chem. Phys. 2012, 136.

(73) Liu, X. L.; Subotnik, J. E. *J. Chem. Theory Comput.* **2014**, *10*, 1004.

(74) Klessinger, M.; Michl, J. Excited States and Photochemistry of Organic Molecules; VCH Publishers, Inc.: New York, New York, 1995.

(75) Ben-Nun, M.; Quenneville, J.; Martinez, T. J. *J. Phys. Chem. A* **2000**, *104*, 5161.

(76) Levine, B. G.; Coe, J. D.; Martinez, T. J. *J. Phys. Chem. B* **2008**, *112*, 405.

(77) Bolovinos, A.; Tsekeris, P.; Philis, J.; Pantos, E.; Andritsopoulos, G. J. Mol. Spectrosc. 1984, 103, 240.

(78) Abouaf, R.; Pommier, J.; Dunet, H. *Chem. Phys. Lett.* **2003**, *381*, 486.

(79) Lassettre, E. N.; Skerbele, A.; Dillon, M. A.; Ross, K. J. *J. Chem. Phys.* **1968**, *48*, 5066.

Chapter 5

Recovering The Coupling Between Doubly Excited and Closed Shell Hartree-Fock Determinants On Single Excitation Level: The Class of Virtual-Occupied Coupled Orbital Configuration Interaction Singles (VOCO-CIS) Methods

"How wonderful that we have met with paradox. Now we have hope of making progress"

-Niels H.D. Bohr (1885-1962)

With the increasing global energy consumption and associated environmental problems brought by the fossil fuels, utilizing clean alternative energies is becoming significant. As the total solar energy is four orders of magnitude larger than global human energy consumption, developing efficient solar cells has been an active field of research for decades. However, a comprehensive atom-by-atom understanding of a series of complex photophysical processes of optoelectronic materials upon excitation like excitation energy transfer,[1,2] non-radiative recombination[3-5] is still lacking. From the theoretical perspective, on one hand, this requires sophisticated electronic structure theories that can correctly describe the complex wavefunctions in various regions of

potential energy surfaces (PESs). On the other hand, extending our ability to simulate real size materials requires the methodology being computationally efficient. With the recent advances in graphical processing units (GPU) computing[6-8], full *ab-initio* quantum chemistry simulations of large molecules up to small proteins[9-12] and nanoparticles[5] is becoming practical.

In our group, we are specifically interested in understanding the non-radiative recombination mechanisms of optoelectronic materials. This requires the knowledge of PESs near conical intersections (CIs). CIs are a set of molecular geometries where two or more Born-Oppenheimer(BO) PESs are degenerate and facilitate efficient electronic population transfer from one state to the other.[13-17] Hence, CI plays a significant role in the non-radiative recombination process. The conical intersection region of the BO PESs has a double cone local topology.[13,14] For non-relativistic Hamiltonian, same spin multiplicity and same symmetry,[18] the degeneracy of the local BO PESs will be lifted linearly by the so-called g and h vectors as defined in eq. (5-1) and eq.(5-2),

$$g^{IJ} = \left\langle \psi_I(\mathbf{r};\mathbf{R}) \left| \frac{\partial \hat{H}_e}{\partial \mathbf{R}} \right| \psi_I(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} - \left\langle \psi_J(\mathbf{r};\mathbf{R}) \left| \frac{\partial \hat{H}_e}{\partial \mathbf{R}} \right| \psi_J(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} \quad (5\text{-}1)$$

$$h^{IJ} = \left\langle \psi_I(\mathbf{r};\mathbf{R}) \left| \frac{\partial \hat{H}_e}{\partial \mathbf{R}} \right| \psi_J(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} \quad (5\text{-}2)$$

where *I* and *J* are the indices for electronic states, $\mathbf{r}$ and $\mathbf{R}$ are the electronic and nuclear coordinates, $\psi_J(\mathbf{r};\mathbf{R})$ is the electronic wavefunction as *J*th eigenfunction of electronic time independent Hamiltonian $\hat{H}_e$

$$\hat{H}_e \psi_J(\mathbf{r};\mathbf{R}) = \varepsilon_J \psi_J(\mathbf{r};\mathbf{R}) \quad (5\text{-}3)$$

Hence, a proper electronic structure theory employed to study the non-radiative recombination mechanism should have the capability of describing both vectors and PESs in the vicinity of CI correctly.

Among various wave-function-based and electron-density-based electronic structure theories, configuration interaction singles (CIS) and time-dependent density functional theory (TDDFT) with Tamm-Dancoff approximation (TDA) are the two widely used, efficient and conceptually simple black box methods to obtain excited state energies and properties.[19,20] TDDFT with TDA has been more widely applied due to the dynamical correlations provided by the functionals and hence the calculated ground, and valence excited states have better agreement with experiments as compared to CIS.[21,22] However, it is known that both methods fail at excited and ground states CI region of the PESs.[23] This is because h vector is always zero at conical intersection geometry as will be derived following. At the conical intersection geometry, the h vector can be re-written as eq. (5-4). The second equality is true because of eq. (5-5). Due to the Brillouin's theorem[24] and TDA[25], $\langle \psi_I(\mathbf{r};\mathbf{R}) | \hat{H}_e | \psi_J(\mathbf{r};\mathbf{R}) \rangle$ at CIS and TDDFT with TDA theoretical levels are always zero when either *I* or *J* equal 0 (meaning either *I* or *J* state represents the ground state). And hence, $h^{IJ} = 0$ at excited and ground states conical intersection geometry.

$$
\begin{aligned}
h^{IJ} &= \left\langle \psi_I(\mathbf{r};\mathbf{R}) \left| \frac{\partial \hat{H}_e}{\partial \mathbf{R}} \right| \psi_J(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} \\
&= \left\langle \frac{\partial \psi_I(\mathbf{r};\mathbf{R})}{\partial \mathbf{R}} \middle| \hat{H}_e \middle| \psi_J(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} + \left\langle \psi_I(\mathbf{r};\mathbf{R}) \middle| \frac{\partial \hat{H}_e}{\partial \mathbf{R}} \middle| \psi_J(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} \\
&\quad + \left\langle \psi_I(\mathbf{r};\mathbf{R}) \middle| \hat{H}_e \middle| \frac{\partial \psi_J(\mathbf{r};\mathbf{R})}{\partial \mathbf{R}} \right\rangle_{\mathbf{r}} \\
&= \frac{\partial \left\langle \psi_I(\mathbf{r};\mathbf{R}) \middle| \hat{H}_e \middle| \psi_J(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}}}{\partial \mathbf{R}}
\end{aligned}
\tag{5-4}
$$

$$
\begin{aligned}
&\left\langle \frac{\partial \psi_I(\mathbf{r};\mathbf{R})}{\partial \mathbf{R}} \middle| \hat{H}_e \middle| \psi_J(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} + \left\langle \psi_I(\mathbf{r};\mathbf{R}) \middle| \hat{H}_e \middle| \frac{\partial \psi_J(\mathbf{r};\mathbf{R})}{\partial \mathbf{R}} \right\rangle_{\mathbf{r}} \\
&= \left\langle \frac{\partial \psi_I(\mathbf{r};\mathbf{R})}{\partial \mathbf{R}} \middle| \varepsilon_J \psi_J(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} + \left\langle \varepsilon_I \psi_I(\mathbf{r};\mathbf{R}) \middle| \frac{\partial \psi_J(\mathbf{r};\mathbf{R})}{\partial \mathbf{R}} \right\rangle_{\mathbf{r}} \\
&\Downarrow \text{at CI, states are degenerate and hence } \varepsilon_J = \varepsilon_I = \varepsilon \\
&= \varepsilon \left[\left\langle \frac{\partial \psi_I(\mathbf{r};\mathbf{R})}{\partial \mathbf{R}} \middle| \psi_J(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}} + \left\langle \psi_I(\mathbf{r};\mathbf{R}) \middle| \frac{\partial \psi_J(\mathbf{r};\mathbf{R})}{\partial \mathbf{R}} \right\rangle_{\mathbf{r}} \right] \\
&= \varepsilon \frac{\partial \left\langle \psi_I(\mathbf{r};\mathbf{R}) \middle| \psi_J(\mathbf{r};\mathbf{R}) \right\rangle_{\mathbf{r}}}{\partial \mathbf{R}} \\
&\Downarrow \text{states are orthorgonal} \\
&= 0
\end{aligned}
\tag{5-5}
$$

It is thus concluded that double excitations are necessary in order to correctly describe the excited and ground states conical intersection region of the PESs.[23] Post CIS single reference methods to include double excitation characters into the wavefunction can roughly be divided into two categories: 1). Considering double excitations perturbatively, e.g. configuration interaction singles with perturbative doubles (CIS(D))[26], 2). Including the double excitations variationally, e.g. configuration interaction singles and doubles (CISD), spin-flip (SF-) CIS.[27,28]

The intuitive variational method CISD is computational infeasible nowadays when

applied to large molecules like nano-sized particles. Besides, including only single and double excitation character into the calculation gives an unbalanced treatment of correlation energies to the ground and excited states.[26] Involving higher excitations could alleviate this problem, however, it is becoming more computationally expensive. In addition, truncated configuration interaction method except CIS always suffers size extensive problem.[29]

On the other side, various efficient zeroth order multi-reference active space methods[30-40] have been developed and successfully applied to compute the PESs in the vicinity of CI. These multi-reference methods include double excitation characters in the wavefunction, however, it usually requires experienced users to choose a reasonable active space in order to get qualitatively correct description of the PESs.

Hence, the current electronic structure theories with double excitation characters in the wavefunction can be either easy to use but computationally expensive or requires practical experiences in order to choose a reasonable active space. One exception is the SF-CIS method which is computationally efficient and applicable to di-radical systems with moderate spin contamination problems sometimes.[41,42] Li and Truhlar et al.[43] have developed configuration interaction corrected (CIC-) TDA method with the similar spirit as the approach going to be introduced here by explicitly introducing the couplings between KS reference state and CIS excited determinants into the TDDFT with TDA Hamiltonian. To model non-radiative recombination mechanism for nano-sized materials, it is important to develop new method that is both efficient and easy to use.

The goal of this chapter is to develop new electronic structure method that is efficient, hopefully, black box and has the capability of describing PESs in the vicinity of ground and excited states CI. I will show an idea that by employing the orbitals optimized beyond Hartree-Fock (HF) wavefunctions, the double excitation characters of the wavefunction can be recovered without explicitly including doubly excited determinants into the wavefunction. The importance of the orbitals employed in the wavefunction in describing the CI region of the PES will be emphasized. A new conceptual understanding of the iterative orbital optimization with excited determinants will also be proposed at the end.

5.1 Method

5.1.1 To go beyond the Brillouin's theorem

Consider a two-state model Hamiltonian at CIS level,

$$\begin{bmatrix} \left\langle \varphi_{HF}\left(\mathbf{r};\mathbf{R}\right)\middle|\hat{H}_e\middle|\varphi_{HF}\left(\mathbf{r};\mathbf{R}\right)\right\rangle & \left\langle \varphi_{HF}\left(\mathbf{r};\mathbf{R}\right)\middle|\hat{H}_e\middle|\varphi_{CIS}\left(\mathbf{r};\mathbf{R}\right)\right\rangle \\ \left\langle \varphi_{CIS}\left(\mathbf{r};\mathbf{R}\right)\middle|\hat{H}_e\middle|\varphi_{HF}\left(\mathbf{r};\mathbf{R}\right)\right\rangle & \left\langle \varphi_{CIS}\left(\mathbf{r};\mathbf{R}\right)\middle|\hat{H}_e\middle|\varphi_{CIS}\left(\mathbf{r};\mathbf{R}\right)\right\rangle \end{bmatrix} \tag{5-6}$$

where $\varphi_{HF}\left(\mathbf{r};\mathbf{R}\right)$ is the HF determinant and $\varphi_{CIS}\left(\mathbf{r};\mathbf{R}\right)=\varphi_{HF}\left(\mathbf{r};\mathbf{R}\right)_i^a$ is the singly excited determinant, with *i* and *a* are the occupied and virtual orbitals in HF reference state. $\left|\varphi_{CIS}\left(\mathbf{r};\mathbf{R}\right)_i^a\right\rangle = a^+ i\left|\varphi_{HF}\left(\mathbf{r};\mathbf{R}\right)\right\rangle$, where a^+ and i are creation and annihilation operators respectively. For the sake of simplicity, we will drop the wavefunction explicit dependence of electronic and nuclear coordinates shown as $\left(\mathbf{r};\mathbf{R}\right)$ in the rest of the discussion. And only considering closed-shell case here with *i*, *j*, … represent occupied

orbitals and *a*, *b*, … represent virtual orbitals in the reference determinant. Brillouin's theorem states that at any nuclear geometry $\mathbf{R}$,

$$\left\langle \varphi_{HF} \left| \hat{H}_e \right| \varphi_{HF\,i}^{\ a} \right\rangle = \left\langle i \left| f \right| a \right\rangle = \left\langle i \left| h \right| a \right\rangle + \sum_j \left\langle ij \right\| aj \right\rangle \tag{5-7}$$

where f is the fock operator. Due to the fact that *i* and *a* are eigenvectors of fock operator,

$$\left\langle \varphi_{HF} \left| \hat{H}_e \right| \varphi_{HF\,i}^{\ a} \right\rangle = \left\langle i \left| f \right| a \right\rangle = 0 \tag{5-8}$$

And hence the two-state model Hamiltonian can be written as,

$$\begin{bmatrix} \left\langle \varphi_{HF} \left| \hat{H}_e \right| \varphi_{HF} \right\rangle & 0 \\ 0 & \left\langle \varphi_{HF\,i}^{\ a} \left| \hat{H}_e \right| \varphi_{HF\,i}^{\ a} \right\rangle \end{bmatrix} \tag{5-9}$$

with the off-diagonal elements being zero independent of $\mathbf{R}$.

However, when the orbitals are not eigenvectors of fock operator, at nuclear geometry $\mathbf{R} \neq \mathbf{R}_{CI}$,

$$\left\langle \varphi_0 \left| \hat{H}_e \right| \varphi_{i'}^{a'} \right\rangle = \left\langle i \left| f \right| a \right\rangle \neq 0 \tag{5-10}$$

where $\mathbf{R}_{CI}$ indicates the CI geometry, $\left|\varphi_0\right\rangle = \left|1', 2', \cdots, i', \cdots, a', b', \cdots\right\rangle$ and $\left|\varphi_{i'}^{a'}\right\rangle = a'^{+}\, i' \left|\varphi_0\right\rangle$. Hence the couplings can be introduced with any orbitals $1', 2', \cdots, i', \cdots, a', b', \cdots$ optimized beyond closed shell HF theoretical level as long as they have no linear dependence of fock operator eigenvectors. And the two-state model Hamiltonian becomes,

$$\begin{bmatrix} \left\langle \varphi_0 \left| \hat{H}_e \right| \varphi_0 \right\rangle & \left\langle \varphi_0 \left| \hat{H}_e \right| \varphi_{i'}^{a'} \right\rangle \\ \left\langle \varphi_{i'}^{a'} \left| \hat{H}_e \right| \varphi_0 \right\rangle & \left\langle \varphi_{i'}^{a'} \left| \hat{H}_e \right| \varphi_{i'}^{a'} \right\rangle \end{bmatrix} \tag{5-11}$$

with the gradient $\frac{\partial\left\langle\varphi_0(\mathbf{r};\mathbf{R})\middle|\hat{H}_e\middle|\varphi_{i'}^{a'}(\mathbf{r};\mathbf{R})\right\rangle_{\mathbf{r}}}{\partial\mathbf{R}}\neq 0$.

Due to the explicit dependence of the wavefunctions on nuclear coordinates, the h vector at CI geometry calculated from this two-state model (5-11),

$$h^{IJ}=\frac{\partial\left\langle\psi_I(\mathbf{r};\mathbf{R})\middle|\hat{H}_e\middle|\psi_J(\mathbf{r};\mathbf{R})\right\rangle_{\mathbf{r}}}{\partial\mathbf{R}}=\frac{\partial\left\langle\varphi_0(\mathbf{r};\mathbf{R})\middle|\hat{H}_e\middle|\varphi_{i'}^{a'}(\mathbf{r};\mathbf{R})\right\rangle_{\mathbf{r}}}{\partial\mathbf{R}}\neq 0 \qquad (5\text{-}12)$$

with $\left|\psi_{I(J)}(\mathbf{r};\mathbf{R})\right\rangle=c_{0,I(J)}\left|\varphi_0(\mathbf{r};\mathbf{R})\right\rangle+c_{i',I(J)}^{a'}\left|\varphi_{i'}^{a'}(\mathbf{r};\mathbf{R})\right\rangle$ are the two eigenvectors of model Hamiltonian (5-11) and c are the configuration interaction coefficients. In CIC-TDA method, the coupling is introduced by treating $\left|\varphi_0\right\rangle=\left|\varphi_{KS}\right\rangle$ and $\left|\varphi_{i'}^{a'}\right\rangle=\left|\varphi_{HF\,i}^{\ a}\right\rangle$.[43]

5.1.2 Practical recipe

As has been demonstrated in the previous section, introducing couplings between reference determinant and singly excited determinant can be achieved by optimizing orbitals of reference determinant beyond HF theoretical level.

One example of such orbital with simple concept can be state averaged CIS self-consistent field (SA-CISSCF) orbitals. The SA-CISSCF orbital is obtained by iteratively optimizing both configuration interaction coefficients and orbital coefficients to minimize the state averaged expectation value of the Hamiltonian,

$$E=\sum_{I=1}^{N_{state}}w_I\frac{\left\langle\Phi_I\middle|\hat{H}_e\middle|\Phi_I\right\rangle}{\left\langle\Phi_I\middle|\Phi_I\right\rangle} \qquad (5\text{-}13)$$

where *I* is the index of state, N_{state} is the number of states being averaged in SA-CISSCF procedure, w_I is the weight of state *I*. The multi-configuration wavefunction Φ_I is expanded as,

$$\left|\Phi_I\right\rangle = c^I{}_0 \left|\varphi_0\right\rangle + \sum_i^{N_{occ}} \sum_a^{N_{vir}} c^{Ia}_i \left|\varphi_i^a\right\rangle \quad (5\text{-}14)$$

N_{occ} *and* N_{vir} are the total number of occupied and virtual orbitals in closed shell HF determinant, $c^I{}_0$ and c^{Ia}_i are the configuration interaction coefficients.

Other examples of such orbitals can be employed from those previous developed in the two-step methods. The two-step method defined here specifically means an approach with orbital optimization followed by configuration interaction expansion, we will denote it as *orbital optimization method-configuration interaction method*. The current two-step methods are developed in a framework as orbital optimization followed by complete active space configuration interaction (CASCI) approach. As compared to complete active space self-consistent field (CASSCF) method,[30] two-step method decouples the orbital optimization and configuration interaction coefficients optimization. There exists a set of such orbitals in two-step methods, for example, improved Hartree-Fock virtual orbitals (IVO),[34] unrestricted Hartree-Fock natural orbitals (UNO),[44] natural orbitals of correlated wavefunction calculations,[45] natural orbitals of high-multiplicity states,[46] floating occupation molecular orbitals (FOMO),[47,48] singly excited active space self-consistent field (SEAS) orbitals,[37] CIS natural orbitals (CISNO) and its variant with relaxed density matrix (CISNOr).[38] By employing these orbitals (except CISNO as it shares the same orbital space with HF[38]) to form the reference determinant, the coupling between the reference and singly excited determinants will have non-zero gradient. For example, both SEAS-CIS and CISNOr-CIS are believed to have such capability and is suitable

for describing the PES near CI with non-zero h vector despite the fact that no doubly excited determinant is included in the wavefunction.

The current method has close spirit as CIC-TDA.[43] The differences are, 1) It is a pure wavefunction based method, and hence the accuracy can be systematically improved. 2) The method can be simplified to orbital optimization-CIS approach, and hence utilizes various properties from those orbitals developed in current two-step methods with low computational cost. For example, SEAS orbitals alleviate the symmetry breaking issue of SA-CASSCF PES; it is expected such property will be retained in SEAS-CIS approach. 3). The occupied and virtual orbitals in HF reference state can be coupled by some of the optimized orbitals, for example, FOMO, SEAS and CISNOr orbitals. The virtual occupied coupling has quantitative importance in describing the PES in the vicinity of CI as will be discussed in the following sections.

The current approach also has close relation to the orbital-optimized methods which was firstly proposed by Purvis and Barlett in 1982,[49] and have been incorporated mostly in the many-body perturbation theory framework.[50-60] Because orbital-optimized method rotates the orbitals to minimize the energy of the correlated wavefunction, the wavefunction obeys the Hellmann-Feynman theorem for orbital rotation parameters and hence orbital coefficients gradient is not required when to evaluate analytic first derivative.[55,59] Previously developed orbital-optimized methods focus on obtaining quantitatively better ground and excited state energies,[49,50,58] thermochemistry parameters[54,59] and solving spatial and spin symmetry-breaking problems[51,56,57] at the

Frank-Condon (FC) region or along the bond dissociation path of the PESs. However, the goal of the current method is to recover the double excitation characters of the wavefunction without explicitly including the doubly excited determinants into the wavefunction and also to explore its potential use in photochemistry especially to capture the features of PES in the vicinity of CI geometry.

The SF ansatz has inspired the current approach as well. SF method employs a triplet reference state[27,28], by doing that, the optimized orbitals in the reference state are no longer eigenvectors of the closed-shell fock operator. In addition, even at single excitation level, for example, SF-CIS wavefunction contains doubly excited determinants as well.[27,28] Hence, not surprisingly, the SF-CIS and SF-TDDFT has been successfully applied to various problems in photochemistry.[27,28,61-68]

5.2 Results and Discussion

5.2.1 Orbital optimization-CIS approach

In this section, we will test two example approaches provided in section 5.1.2, SA-CISSCF and CISNOr-CIS. For both methods, no doubly excited determinants are included in the wavefunction. As we will show, they are able to lift the degeneracy of the PESs near the ground and excited states CI geometry with non-zero h and g vectors.

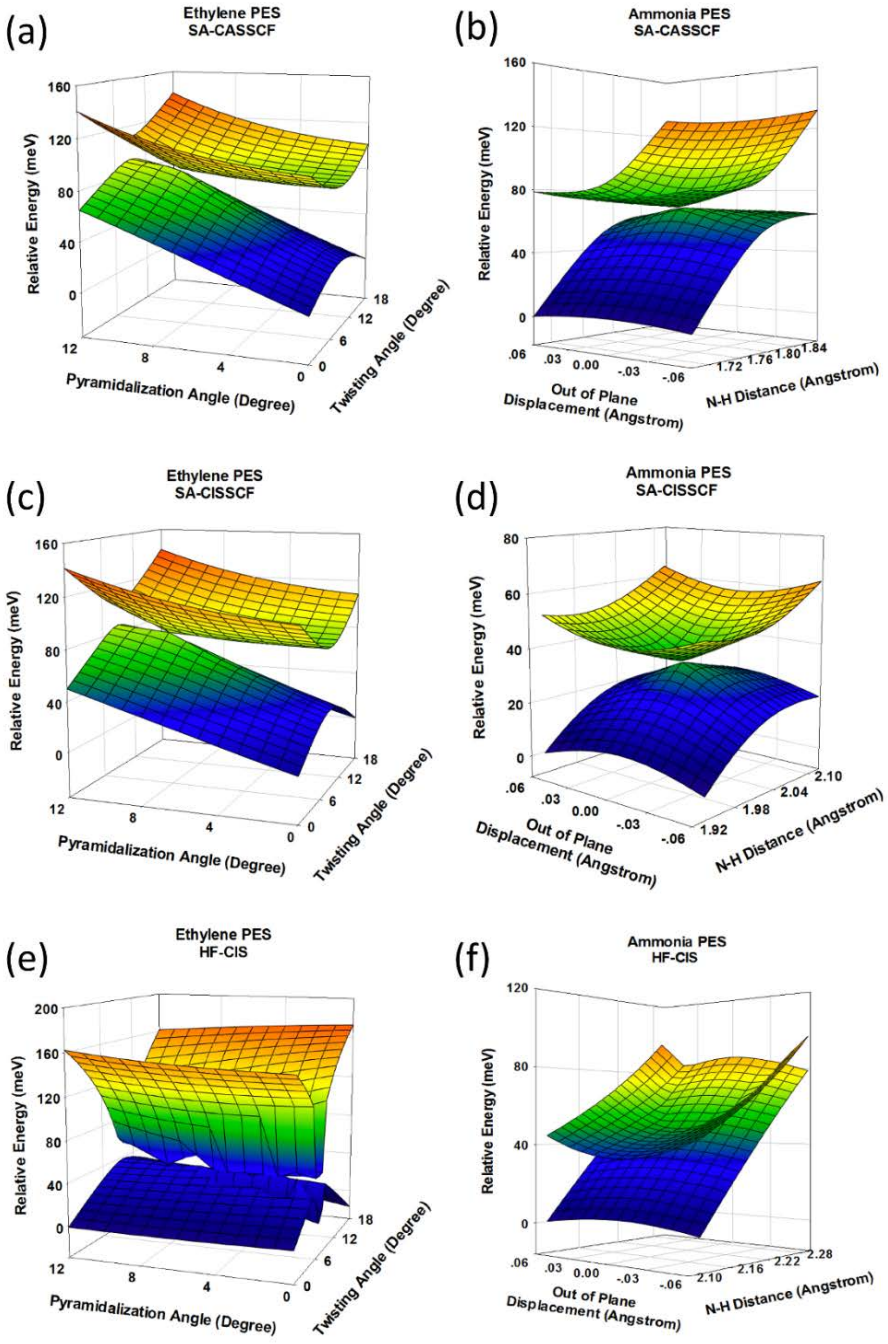


Figure 5-1. S_1 and S_0 PESs PESs near conical intersection geometries of (a) ethylene molecule, calculated on SA-CASSCF theoretical level, (b) ammonia molecule, calculated on SA-CASSCF theoretical level, (c) ethylene molecule, calculated on SA-CISSCF theoretical level, (d) ammonia molecule, calculated on SA-CISSCF theoretical level, (e) ethylene

Figure 5-1 (cont'd). molecule, calculated on HF-CIS theoretical level, (f) ammonia molecule, calculated on HF-CIS theoretical level.

SA-CISSCF optimizes orbital coefficients and configuration interaction coefficients simultaneously at CIS theoretical level within the full orbital space as compared to SA-CASSCF, which optimizes orbital coefficients and configuration interaction coefficients simultaneously at full CI theoretical level within the active orbital space. The SA-CISSCF orbitals are optimal orbitals toward both ground and excited states at CIS level. SA-CASSCF is current the state of the art method in describing conical intersection region of PESs as well as widely employed in non-adiabatic dynamics simulation.[69-71] Hence, it is reasonable to compare with current approaches with SA-CASSCF results. Figure 5-1 (a), (c) and (e) show the S_0 and S_1 PESs near pyramidalized CI region of ethylene molecule. The pyramidalized S_1/S_0 minimal energy CI (MECI) geometry is optimized at SA-CASSCF, SA-CISSCF and HF-CIS levels of theory with 6-31G basis set in (a), (c) and (e) respectively. The same level of theory is applied to compute the PESs. The PESs are calculated along two directions that are known to lift the S_1-S_0 energy gap near the MECI region of the PESs of ethylene molecule. These two directions correspond to the g and h vectors, one is twisting coordinate; the other is pyramidalization coordinate.[69,72] Compare Figure 5-1(a) and (c), it is clear that the S_1-S_0 energy gap is lifted up when the nuclear coordinates displaced away from the MECI geometry along these two directions. This proves the non-zero h and g vectors in SA-CISSCF approach. However, as shown in Figure 15-1(e),

on the HF-CIS PES, the S_1-S_0 energy gap can only be increased along the twisting coordinate. It is known that the twisting coordinate corresponds to the g vector, while the h vector is represented by the pyramidalization coordinate.[69] This is consistent with what we have derived in equation (5-4) and (5-5) which predicts the h vector is always zero at conical intersection geometry at HF-CIS level of theory. Figure 5-1 (b), (d) and (f) show the S_0 and S_1 PESs near S_1/S_0 MECI geometry of ammonia molecule. The ammonia S_1/S_0 MECI geometry is optimized at SA-CASSCF, SA-CISSCF and HF-CIS levels of theory with 6-31G basis set in (b), (d) and (f) respectively with the same level of theory computing the PESs. The PESs are calculated along two directions, one corresponds to N-H bond dissociation, the other corresponds to the vertical displacement of the N atom (vertical here means treating the three hydrogen atoms as a plane, and displace N atom by moving toward or away from the plane vertically). Again, we see the SA-CASSCF and SA-CISSCF PESs are in qualitative agreement as shown in (b) and (d), with two directions open up the S_1-S_0 energy gap. The HF-CIS predicts the S_1 and S_0 PESs without breaking the degeneracy when geometries are distorted from MECI geometry.

In another example, we will employ the CISNOr-CIS approach. CISNOr is obtained by diagonalizing the state averaged (SA-) CIS first order reduced density matrix which is constructed by averaging the first order reduced density matrices of the ground (HF) and low-lying excited (CIS) states,

$$\boldsymbol{\gamma}^{\text{SA-CIS}} = \frac{1}{N+1}\left(\boldsymbol{\gamma}^{\text{HF}} + \sum_{I=1}^{N} \boldsymbol{\gamma}^{\text{CISr}(I)}\right) \tag{5-14}$$

where γ^{HF} and $\gamma^{CISr(I)}$ are the first order reduced density matrix of the HF ground state and the first order relaxed reduced density matrix of the *I*th CIS singlet excited state.[73,74] The S_1 and S_0 PESs around MECI geometry of ammonia is computed at CISNOr-CIS theoretical level with 6-31G basis set as shown in Figure 5-2(a). The S_1/S_0 MECI geometry is optimized at the same level of theory employed in calculating PESs. Nuclear geometry is displaced away from CI geometry along with the N-H bond dissociation direction and N vertical displacement direction to generate the PESs. S_1-S_0 energy gap is illustrated as a contour plot shown in Figure 5-2(b). As can be seen from Figure 5-2(b), this gap is lifted in both directions. This result indicates the qualitative correct double cone topology of the PESs near CI geometry. All electronic structure calculations were performed with the MolPro software package[75-78], MECI geometry optimizations utilize CIOpt[71] in conjunction with MolPro.

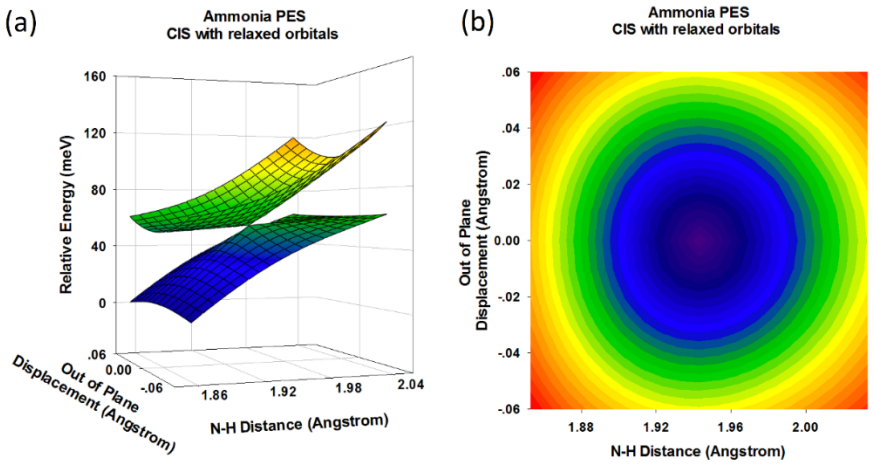


Figure 5-2. (a) S_1 and S_0 PESs of ammonia molecule near conical intersection geometry calculated on CISNOr-CIS theoretical level. (b), a corresponding contour plot of the S_1-S_0 energy gap.

5.2.2 The importance of virtual occupied orbital coupling

We have investigated the possibility of employ orbital optimization-CIS approach in describing the PES in the vicinity of CI in the previous sections. The differences of the varieties of reference determinants lie at the orbitals constitute it. And hence it is important to choose the suitable set of orbitals.

It is noteworthy that any zeroth order configuration interaction method can be treated as a two-step approach, for example, SA-CASSCF can be thought as a reference determinant with SA-CASSCF orbital plus a CASCI expansion. The zeroth order here means the orbital optimization step may involve multi-configurations, while in the final configuration interaction expansion, there is only one reference determinant, e.g., multi-reference configuration interaction (MRCI)[79] is not a zeroth order method. And hence, it is reasonable to argue that the orbitals are of great importance or the reference determinant is of great importance. A zeroth order configuration interaction method with suitable single reference determinant followed by a CI expansion may provide quantitative good agreement with trust worthy higher order wavefunction method like multi-reference CI with Davison correction (MRCI(Q)).[79]

Here, we investigate the quantitative behaviors of some of the orbitals in describing the PES around CI region. To avoid the zero h vector problem for HF-CIS method, we investigate various orbitals behaviors through orbital optimization-CASCI approach. Table 5-1 shows the various single point calculations performed on ethylene Frank-Condon (FC) geometry and pyramidalized MECI geometries optimized at different

theoretical levels. Each column represents single point calculations at different theoretical levels performed on the same geometry. And hence each row with S_0, S_1 energies and S_1-S_0 energy gaps are calculated on the same theoretical level at different geometries. With all the S_1, S_0 energies at the same level of theory are computed by treating ground state of FC geometry as reference. The FC geometry is optimized at MP2/6-31G** theoretical level, MECI geometries are optimized at SA-CASSCF, CISNO-CASCI, CISNOr-CASCI and SEAS-CASCI levels with two electron two orbital active space and 6-31G** basis set. Two states are being averaged in SA-CASSCF theoretical level. The same parameters were chosen for single point calculations.

At SA-CASSCF optimized MECI geometry, MRCI(Q) predicts 0.164 eV gap which indicates SA-CASSCF predicted MECI geometry should be closed to that of MRCI(Q) predicted. The 0.164 eV gap is not perfect, however, it can be treated as semi-quantitative accuracy. At the same geometry, CISNOr-CASCI and SEAS-CASCI predicts small S_1-S_0 energy gaps (0.253 and 0.021 eV) as well. While CISNO-CASCI and HF-CASCI predicts 1.578 and 1.711 eV energy gaps at SA-CASSCF optimized MECI. This large discrepancy between MRCI(Q) and CISNO-CASCI or HF-CASCI indicates the qualitative disagreement between these methods. The semi-quantitative accuracy of SA-CASSCF, SEAS-CASCI and CISNOr-CASCI as compared with MRCI(Q) in predicting MECI geometry is further investigated by performing single point calculations on CISNOr-CASCI and SEAS-CASCI optimized MECI geometries. At CISNOr-CASCI optimized geometry, SA-CASSCF, SEAS-CASCI and MRCI(Q) predicts 0.272, 0.272 and 0.072 eV S_1-S_0 energy gap respectively.

While CISNO-CASCI and HF-CASCI shows 1.337 and 1.466 eV energy gap. The similar behavior is observed at SEAS-CASCI optimized MECI as SA-CASSCF, CISNOr-CASCI and MRCI(Q) predicts 0.006, 0.268 and 0.177 eV energy gap. It is not surprising based on the previous behaviors that SA-CASSCF, SEAS-CASCI, CISNOr-CASCI and MRCI(Q) all predicts large S_1-S_0 energy gap at MECI geometries optimized either at CISNO-CASCI or HF-CASCI theoretical levels as shown in Table 5-1.

Hence, among these five different theoretical levels optimized ethylene pyramidalized MECI geometries, we can conclude that SA-CASSCF, SEAS-CASCI and CISNOr-CASCI have semi-quantitative accuracy, while HF-CASCI and CISNO-CASCI have qualitative disagreement with MRCI(Q) results for this specific system. In the following text, we will call SA-CASSCF, SEAS and CISNOr orbitals as semi-quantitative orbitals. As we applied same active space in the CASCI expansion for these methods, the quantitative and qualitative differences among these method can be reasonably attributed to the differences of the orbitals or the reference determinant. The CASCI expansion allows doubly excited determinant being coupled with the reference determinant and hence all these two-step methods have the capability in describing the CI region PES and provide a correct double cone local topology description. However, it can be seen from the above calculations that these CI geometries may have qualitative disagreement with those predicted from more trustworthy calculations. This again, emphasizes the importance of the orbitals in the reference determinant.

Table 5-1. The S_1, S_0 and S_1-S_0 energy gap (all in the unit of eV) computed with SA-CASSCF, CISNO-CASC, CISNOr-CASCI, SEAS-CASCI, HF-CASCI and MRCI(Q) methods are shown at FC geometry and MECI geometries optimized at SA-CASSCF, CISNO-CASCI, CISNOr-CASCI, SEAS-CASCI theoretical levels.

	Geometry		FC	MECI			
	Optimization Method		MP2/6-31G**	SA-CASSCF /6-31G**	CISNO-CASCI /6-31G**	CISNOr-CASCI /6-31G**	SEAS-CASCI /6-31G**
Single Point Energy Calculations	SA-CASSCF	S_0	0	5.883	4.301	5.622	5.907
		S_1	10.233	5.904	5.944	5.894	5.913
		S_1-S_0 gap	10.233	0.021	1.643	0.272	0.006
	CISNO-CASCI	S_0	0	5.543	5.574	5.531	5.552
		S_1	10.48228	7.121	5.578	6.868	7.144
		S_1-S_0 gap	10.48228	1.578	0.004	1.337	1.592
	CISNOr-CASCI	S_0	0	5.841	4.475	5.830	5.850
		S_1	10.338	6.094	5.925	5.834	6.118
		S_1-S_0 gap	10.338	0.253	1.450	0.004	0.268
	SEAS-CASCI	S_0	0	5.480	3.897	5.219	5.503
		S_1	9.701	5.501	5.540	5.491	5.510
		S_1-S_0 gap	9.701	0.021	1.643	0.272	0.007
	HF-CASCI	S_0	0	5.500	5.534	5.488	5.509
		S_1	10.399	7.211	5.646	6.954	7.233
		S_1-S_0 gap	10.399	1.711	0.112	1.466	1.724

Table 5-1 (cont'd)

Geometry		FC	MECI			
Optimization Method		MP2/6-31G**	SA-CASSCF /6-31G**	CISNO-CASCI /6-31G**	CISNOr-CASCI /6-31G**	SEAS-CASCI /6-31G**
MRCI(Q)	S_0	0	5.029	3.827	4.968	5.037
	S_1	8.9730	5.193	5.245	5.040	5.214
	S_1-S_0 gap	8.9730	0.164	1.418	0.072	0.177

Table 5-2. The S_1, S_0 and S_1-S_0 energy gap (all in the unit of eV) computed with SA-CASSCF, CISNO-CASC, CISNOr-CASCI, SEAS-CASCI and HF-CASCI methods without doubly excited determinants in the CASCI Hamiltonian are shown at the MECI geometries optimized at SA-CASSCF, CISNO-CASCI, CISNOr-CASCI, SEAS-CASCI theoretical levels.

	Geometry Optimization Method		MECI SA-CASSCF /6-31G**	MECI CISNO-CASCI /6-31G**	MECI CISNOr-CASCI /6-31G**	MECI SEAS-CASCI /6-31G**
Single Point Energy Calculations	SA-CASSCF	S_0	5.883	4.301	5.622	5.907
		S_1	5.906	5.945	5.896	5.915
		S_1-S_0 gap	0.023	1.644	0.274	0.008
	CISNO-CASCI	S_0	5.545	5.575	5.533	5.554
		S_1	7.121	5.579	6.868	7.144
		S_1-S_0 gap	1.576	0.004	1.335	1.590
	CISNOr-CASCI	S_0	5.843	5.561	5.830	5.852
		S_1	6.094	6.104	5.836	6.118
		S_1-S_0 gap	0.251	0.543	0.006	0.266
	SEAS-CASCI	S_0	5.480	3.897	5.219	5.503
		S_1	5.503	5.541	5.492	5.512
		S_1-S_0 gap	0.023	1.644	0.273	0.009
	HF-CASCI	S_0	5.502	5.535	5.489	5.511
		S_1	7.211	5.646	6.954	7.233
		S_1-S_0 gap	1.709	0.111	1.465	1.722

So, what are the differences of those semi-quantitative orbitals compared with HF orbitals or CISNO? One important difference comes from whether the optimized occupied and virtual orbitals are being coupled by Fock operator. In SA-CASSCF and SEAS-CASCI methods, both orbitals are optimized iteratively with excited determinants, and hence the occupied and virtual orbitals in HF reference state are coupled. In CISNOr-CASCI, the orbitals are relaxed based on the excited state charge redistribution.[73] However, both HF and CISNO orbitals do not have such coupling. This can be further illustrated by comparing CISNO and CISNOr. The difference in these two orbitals comes from the CIS density matrices employed in the SA-CIS density matrix. CISNO employs non-relaxed density matrix, CISNOr employs the relaxed density matrix. The only difference of relaxed and non-relaxed CIS density matrices lies at the occupied-virtual coupling block, which is referred as OV block in the previous paper.[73] And hence, it can be reasonably believed that the coupling between occupied and virtual orbitals are important in describing PES near CI region in a quantitative sense. We will denote this type of coupled orbital as virtual-occupied coupled orbital (VOCO). The VO coupling (VOC) can be written as

$$\mathrm{VOC} = \langle i | f | a \rangle \qquad (5\text{-}15)$$

where f is the fock operator, *i* and *a* are occupied and virtual orbitals. The occupied and virtual orbitals at various orbital optimization methods are defined by ordering the optimized orbital first and assigning 2 or 0 occupation numbers onto each special orbitals based on the total number of electrons. And then those orbitals with occupation

number 2 are occupied orbitals and those with occupation number 0 are virtual orbitals. For example, SEAS orbitals are ordered according to the diagonal elements of effective fock operator; CISNOr are ordered according to the occupation numbers.

This is only one case we show here to notify the importance of VOC, according to our experience, SA-CASSCF, CISNOr-CASCI and MRCI(Q) usually predicts larger S_1-S_0 energy gap at CISNO-CASCI optimized MECI geometries, while this gap is obviously reduced at CISNOr-CASCI optimized MECI geometries. However, a more systematic investigation and rigorous mathematical proof of the importance of VOC in describing PES near CI are required in the future.

The significance of orbitals in affecting the accuracy of the methods in the vicinity of conical intersection geometry is further investigated by applying active space doubly excited determinants excluded SA-CASSCF, CISNO-CASCI, CISNOr-CASCI, SEAS-CASCI and HF-CASCI single point calculations on the pyramidalized ethylene MECI geometries optimized at SA-CASSCF, CISNO-CASCI, CISNOr-CASCI and SEAS-CASCI theoretical levels. We have summarized the results in Table 5-2. It is noteworthy here that the first column shows the various single point calculations methods in Table 5-2, however, these methods are referring to the corresponding method without doubly excited determinants in the active space, for example, in HF-CASCI, the orbitals are optimized with HF method, while the CASCI Hamiltonian does not contain doubly excited determinants. Compare Table 5-2 and Table 5-1, most of the differences between the methods with and without doubly excited determinants in the active space is negligible. For example, SA-CASSCF with and

without doubly excited determinants predict 0.021 and 0.023 eV S_1-S_0 energy gap at SA-CASSCF optimized ethylene pyramidalized MECI geometry. The only one that shows noticeable difference is the single point calculation at CISNOr-CASCI theoretical level at CISNO-CASCI optimized geometry. Including doubly excited determinant in the CASCI Hamiltonian or not predicts 1.450 or 0.543 eV S_1-S_0 energy gap. However, 0.543 eV is still a relatively large gap and supports the importance of VOCO.

5.3 A New Conceptual Understanding of Iterative Orbital Optimization with Excited Determinants

As we show in the previous section, SA-CISSCF can be treated as a two-step method with SA-CISSCF orbitals followed by CIS expansion. However, it is noteworthy here that this optimization started with HF orbitals as the initial guess. According to Brillouin's theorem, the coupling between HF determinant and the singly excited determinant is zero. This sounds contradictory statement provides us another way in looking at the iterative orbital optimization process with excited determinants.

Let us start with iteration by iteration looking at the SA-CISSCF orbital optimization process. In the first iteration, it starts with HF orbitals and the energy of the total wavefunction is minimized by obtaining the optimal configuration interaction coefficients. And then, to minimize the state averaged expectation value of the Hamiltonian, the orbitals will rotate to form a new reference determinant. This new reference determinant contains singly excited state information. In the next iteration, the singly

excited determinant is built on the new reference determinant instead of HF determinant. And hence, the new reference determinant contains doubly excited information (think this as twice single excitation process) after the second iteration. This iterative process is converged until the optimal orbitals is obtained. Hence, the orbital optimization with HF and singly excited determinants can be treated as introducing a multi-level excitation characters into HF orbitals. This multi-level excitation brings the coupling between the new reference determinant and singly excited determinant.

5.4 Conclusions

In this work, we have demonstrated that the h vector vanishes for HF-CIS method based on the previous work by Levine and Martinez et al.[23] Hence it is not suitable to apply HF-CIS in studying photochemistry involving non-radiative decay process. This can be solved by either introducing doubly excited determinants into the wavefunction as has been discussed in reference 23 or applying orbital optimization scheme to change the reference determinant. The orbitals constitute the reference determinant can be generated without explicitly introducing the doubly excited determinants into the orbital optimization step, such as SA-CISSCF, CISNOr-CIS and SEAS-CIS methods. epThis inspired us a new conceptual understanding of the iterative orbital optimization with excited determinants. An iterative orbital optimization provides multi-level excitation information into the orbitals, despite the fact that only singly excited determinants is included in orbital optimization step. We have also shown the importance of employing

VOCO in achieving quantitative accuracy in describing PES in the vicinity of CI geometry. And hence, with some choices of VOCO (e.g., CISNOr, FOMO), a VOCO-CIS two-step approach may provide us an efficient, black-box and semi-quantitative accuracy method in investigating various photophysical processes for large systems, especially with the capability of describing non-radiative recombination. The VOCO-CIS can be an alternative to SF-CIS and CIC-TDA methods.

APPENDIX

APPENDIX

Table 5-S1. The absolute energies of the FC and various optimized geometries of the ethylene pyramidalized MECI at the CASSCF level of theory (with active space and basis sets as described in the text).

CASSCF	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point (MP2)	-78.05703418	-77.68094703
SA-CASSCF optimized S_0/S_1 MECI	-77.84081888	-77.84005934
CISNO-CASCI optimized S_0/S_1 MECI	-77.89896961	-77.83860311
CISNOr-CASCI optimized S_0/S_1 MECI	-77.85041527	-77.84043831
SEAS-CASCI optimized S_0/S_1 MECI	-77.83995698	-77.83972273

Table 5-S2. The absolute energies of the FC and various optimized geometries of the ethylene pyramidalized MECI at the CISNO-CASCI level of theory (with active space and basis sets as described in the text).

CISNO-CASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point (MP2)	-78.06003566	-77.67481363
SA-CASSCF optimized S_0/S_1 MECI	-77.85632343	-77.79834894
CISNO-CASCI optimized S_0/S_1 MECI	-77.85517898	-77.85505857
CISNOr-CASCI optimized S_0/S_1 MECI	-77.85677913	-77.80763704
SEAS-CASCI optimized S_0/S_1 MECI	-77.855993	-77.79747738

Table 5-S3. The absolute energies of the FC and various optimized geometries of the ethylene pyramidalized MECI at the CISNOr-CASCI level of theory (with active space and basis sets as described in the text).

CISNOr-CASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point (MP2)	-78.0586937	-77.67876846
SA-CASSCF optimized S_0/S_1 MECI	-77.8440249	-77.83473242
CISNO-CASCI optimized S_0/S_1 MECI	-77.89424515	-77.84093798
CISNOr-CASCI optimized S_0/S_1 MECI	-77.8444473	-77.84431057
SEAS-CASCI optimized S_0/S_1 MECI	-77.84369962	-77.83384566

Table 5-S4. The absolute energies of the FC and various optimized geometries of the ethylene pyramidalized MECI at the SEAS-CASCI level of theory (with active space and basis sets as described in the text).

SEAS-CASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point (MP2)	-78.04220787	-77.68569908
SA-CASSCF optimized S_0/S_1 MECI	-77.84083366	-77.84004441
CISNO-CASCI optimized S_0/S_1 MECI	-77.89897642	-77.83859623
CISNOr-CASCI optimized S_0/S_1 MECI	-77.85042852	-77.84042493
SEAS-CASCI optimized S_0/S_1 MECI	-77.83997103	-77.83970854

Table 5-S5. The absolute energies of the FC and various optimized geometries of the ethylene pyramidalized MECI at the HF-CASCI level of theory (with active space and basis sets as described in the text).

HF-CASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point (MP2)	-78.05843623	-77.67626909
SA-CASSCF optimized S_0/S_1 MECI	-77.85629742	-77.79343888
CISNO-CASCI optimized S_0/S_1 MECI	-77.85506799	-77.85094018
CISNOr-CASCI optimized S_0/S_1 MECI	-77.85675609	-77.80287794
SEAS-CASCI optimized S_0/S_1 MECI	-77.85596782	-77.79263795

Table 5-S6. The absolute energies of the FC and various optimized geometries of the ethylene pyramidalized MECI at the MRCI(Q) level of theory (with active space and basis sets as described in the text).

MRCI(Q)	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point (MP2)	-78.35714492	-78.02738884
SA-CASSCF optimized S_0/S_1 MECI	-78.1723252	-78.16628463
CISNO-CASCI optimized S_0/S_1 MECI	-78.21650255	-78.1643818
CISNOr-CASCI optimized S_0/S_1 MECI	-78.17456476	-78.17191744
SEAS-CASCI optimized S_0/S_1 MECI	-78.17203281	-78.16552976

Table 5-S7. The absolute energies of the various optimized geometries of the ethylene pyramidalized MECI at the SA-CASSCF without doubly excited determinant level of theory (with active space and basis sets as described in the text).

SA-CASSCF without doubly excited determinant	S_0 Energy / a.u.	S_1 Energy / a.u.
SA-CASSCF optimized S_0/S_1 MECI	-77.84083366	-77.83998988
CISNO-CASCI optimized S_0/S_1 MECI	-77.89897642	-77.83855984
CISNOr-CASCI optimized S_0/S_1 MECI	-77.85042851	-77.84037446
SEAS-CASCI optimized S_0/S_1 MECI	-77.83996993	-77.83965589

Table 5-S8. The absolute energies of the various optimized geometries of the ethylene pyramidalized MECI at the CISNO-CASCI without doubly excited determinant level of theory (with active space and basis sets as described in the text).

CISNO-CASCI without doubly excited determinant	S_0 Energy / a.u.	S_1 Energy / a.u.
SA-CASSCF optimized S_0/S_1 MECI	-77.85624444	-77.79834894
CISNO-CASCI optimized S_0/S_1 MECI	-77.85516395	-77.85502419
CISNOr-CASCI optimized S_0/S_1 MECI	-77.85670693	-77.80763456
SEAS-CASCI optimized S_0/S_1 MECI	-77.85591491	-77.79747725

Table 5-S9. The absolute energies of the various optimized geometries of the ethylene pyramidalized MECI at the CISNOr-CASCI without doubly excited determinant level of theory (with active space and basis sets as described in the text).

CISNOr-CASCI without doubly excited determinant	S_0 Energy / a.u.	S_1 Energy / a.u.
SA-CASSCF optimized S_0/S_1 MECI	-77.84394907	-77.83473242
CISNO-CASCI optimized S_0/S_1 MECI	-77.85434374	-77.83436084
CISNOr-CASCI optimized S_0/S_1 MECI	-77.84444645	-77.84422034
SEAS-CASCI optimized S_0/S_1 MECI	-77.84362471	-77.83384461

Table 5-S10. The absolute energies of the various optimized geometries of the ethylene pyramidalized MECI at the SEAS-CASCI without doubly excited determinant level of theory (with active space and basis sets as described in the text).

SEAS-CASCI without doubly excited determinant	S_0 Energy / a.u.	S_1 Energy / a.u.
SA-CASSCF optimized S_0/S_1 MECI	-77.84083366	-77.83998988
CISNO-CASCI optimized S_0/S_1 MECI	-77.89897642	-77.83855984
CISNOr-CASCI optimized S_0/S_1 MECI	-77.85042851	-77.84037446
SEAS-CASCI optimized S_0/S_1 MECI	-77.83996993	-77.83965589

Table 5-S11. The absolute energies of the various optimized geometries of the ethylene pyramidalized MECI at the HF-CASCI without doubly excited determinant level of theory (with active space and basis sets as described in the text).

HF-CASCI without doubly excited determinant	S_0 Energy / a.u.	S_1 Energy / a.u.
SA-CASSCF optimized S_0/S_1 MECI	-77.85624444	-77.79343888
CISNO-CASCI optimized S_0/S_1 MECI	-77.85503405	-77.85093637
CISNOr-CASCI optimized S_0/S_1 MECI	-77.85670693	-77.802875
SEAS-CASCI optimized S_0/S_1 MECI	-77.85591491	-77.79263794

Optimized Geometries (In Angstrom) corresponding to Table 5-1 and Table 5-2

Ethylene FC (MP2/6-31G):**

C	0.148292706	-0.283999332	0.020259819
C	-0.032007135	-0.104646114	1.332307992
H	0.469459619	0.523857688	-0.62266638
H	-0.020402737	-1.243858592	-0.448227155
H	0.136554012	0.855112319	1.800941576
H	-0.353376006	-0.912596076	1.975088141

SA-CASSCF optimized ethylene pyramidalized MECI

C	-0.106381178	0.184249095	-0.115054396
C	-0.018075946	0.031307731	1.249482257
H	0.953712157	-0.167227809	-0.326130821
H	-0.332030645	-0.909553896	-0.326142379
H	-0.438118334	0.758849461	1.941974922
H	0.454253940	-0.786789584	1.805874596

CISNO-CASCI optimized ethylene pyramidalized MECI

C	-0.045883474	0.061146723	-0.079399600
C	-0.012209934	0.023288184	1.320228185
H	1.022858953	-0.063763300	-0.364944011
H	-0.430687418	-0.923892999	-0.388124518
H	-0.438241782	0.792326794	1.947941069
H	0.465003415	-0.754707927	1.925881015

CISNOr-CASCI optimized ethylene pyramidalized MECI

C	-0.086631650	0.182856668	-0.111255016
C	0.004574180	0.038920874	1.258006911
H	0.974982954	-0.158575814	-0.338186002
H	-0.358271831	-0.885697941	-0.341268650
H	-0.405732529	0.777167331	1.950838879
H	0.450853723	-0.785936498	1.828370829

SEAS-CASCI optimized ethylene pyramidalized MECI

C	-0.093261239	0.186835558	-0.108884085
C	-0.001672758	0.030260178	1.254914392
H	0.975053245	-0.162005348	-0.324185265
H	-0.328437573	-0.902937498	-0.311282617
H	-0.428677342	0.767680072	1.935143216
H	0.464223072	-0.777948074	1.838513038

Optimized Geometries (In Angstrom) corresponding to Figure 5-1 and Figure 5-2

SA-CASSCF optimized ethylene pyramidalized MECI (6-31g basis set)

C	-0.129346547	0.223994102	-0.104980429
C	-0.023073650	0.039961179	1.260588685
H	0.941365236	-0.210714774	-0.331070936
H	-0.288181420	-0.920614423	-0.331104657
H	-0.440248033	0.762571499	1.953582136
H	0.452844403	-0.784362587	1.782989494

SA-CISSCF optimized ethylene pyramidalized MECI (6-31g basis set)

C	-0.069591531	0.185318729	-0.112208197
C	0.032962180	0.034342146	1.292910383
H	1.020332907	-0.139038493	-0.338980036
H	-0.343932455	-0.910419099	-0.331250594
H	-0.402368127	0.798282820	1.940960085
H	0.490804118	-0.775316045	1.874823500

SA-CASSCF optimized ammonia MECI (6-31g basis set)

N	0.081962286	0.063857242	0.005334361
H	0.055266217	-0.010672555	1.014414081
H	1.048092103	0.174992877	-0.285339756
H	-1.536874241	-0.036354209	-0.750802100

SA-CISSCF optimized ammonia MECI (6-31g basis set)

N	0.120262320	0.087903731	0.068804637
H	0.107561947	0.164813130	1.070804871
H	1.026175489	0.052274228	-0.368255652
H	-1.789273118	0.035443933	-0.609747107

CISNOr-CIS optimized ammonia MECI (6-31g basis set)

N	0.130450476	0.004887635	-0.003149241
H	0.178937944	0.096788226	1.002403383
H	1.095981335	-0.134394406	-0.313195632
H	-1.724199157	0.158617274	-0.563958969

REFERENCES

REFERENCES

(1) Tiwari, V.; Peters, W. K.; Jonas, D. M. *P Natl Acad Sci USA* **2013**, *110*, 1203.

(2) Musser, A. J.; Liebel, M.; Schnedermann, C.; Wende, T.; Kehoe, T. B.; Rao, A.; Kukura, P. *Nat Phys* **2015**, *11*, 352.

(3) Bernardi, F.; Olivucci, M.; Robb, M. A. *Chemical Society Reviews* **1996**, *25*, 321.

(4) Galland, C.; Ghosh, Y.; Steinbruck, A.; Sykora, M.; Hollingsworth, J. A.; Klimov, V. I.; Htoon, H. Nature 2011, 479, 203.

(5) Shu, Y. N.; Fales, B. S.; Levine, B. G. *Nano Letters* **2015**, *15*, 6247.

(6) Ufimtsev, I. S.; Martinez, T. J. Journal of Chemical Theory and Computation 2008, 4, 222.

(7) Ufimtsev, I. S.; Martinez, T. J. Journal of Chemical Theory and Computation 2009, 5, 3138.

(8) Ufimtsev, I. S.; Martinez, T. J. Journal of Chemical Theory and Computation 2009, 5, 2619.

(9) Ufimtsev, I. S.; Luehr, N.; Martinez, T. J. *J Phys Chem Lett* **2011**, *2*, 1789.

(10) Isborn, C. M.; Luehr, N.; Ufimtsev, I. S.; Martinez, T. J. Journal of Chemical Theory and Computation 2011, 7, 1814.

(11) Kulik, H. J.; Luehr, N.; Ufimtsev, I. S.; Martinez, T. J. *J Phys Chem B* **2012**, *116*, 12501.

(12) Isborn, C. M.; Gotz, A. W.; Clark, M. A.; Walker, R. C.; Martinez, T. J. J Chem Theory Comput 2012, 8, 5092.

(13) Yarkony, D. R. *Accounts Chem Res* **1998**, *31*, 511.

(14) Yarkony, D. R. *Reviews of Modern Physics* **1996**, *68*, 985.

(15) Worth, G. A.; Cederbaum, L. S. *Annu Rev Phys Chem* **2004**, *55*, 127.

(16) Levine, B. G.; Martinez, T. J. *Annual Review of Physical Chemistry* **2007**, *58*, 613.

(17) Domcke, W.; Yarkony, D. R. *Annual Review of Physical Chemistry, Vol 63* **2012**, *63*, 325.

(18) Domcke, W.; Yarkony, D. R.; Köppel, H. Conical intersections: theory, computation and experiment; World Scientific, 2011; Vol. 17.

(19) Dreuw, A.; Head-Gordon, M. *Chemical Reviews* **2005**, *105*, 4009.

(20) Marques, M. A. L.; Gross, E. K. U. *Annual Review of Physical Chemistry* **2004**, *55*, 427.

(21) Isegawa, M.; Peverati, R.; Truhlar, D. G. *J Chem Phys* **2012**, *137*.

(22) Laurent, A. D.; Jacquemin, D. *Int J Quantum Chem* **2013**, *113*, 2019.

(23) Levine, B. G.; Ko, C.; Quenneville, J.; Martinez, T. J. *Molecular Physics* **2006**, *104*, 1039.

(24) Szabó, A.; Ostlund, N. S. Modern quantum chemistry : introduction to advanced electronic structure theory; Free Press ; Collier Macmillan: New York,London, 1982.

(25) Hirata, S.; Head-Gordon, M. *Chemical Physics Letters* **1999**, *314*, 291.

(26) Headgordon, M.; Rico, R. J.; Oumi, M.; Lee, T. J. *Chem Phys Lett* **1994**, *219*, 21.

(27) Shao, Y. H.; Head-Gordon, M.; Krylov, A. I. *J Chem Phys* **2003**, *118*, 4807.

(28) Krylov, A. I. *Accounts Chem Res* **2006**, *39*, 83.

(29) Szabo, A.; Ostlund, N. S. Modern quantum chemistry: introduction to advanced electronic structure theory; Courier Corporation, 1989.

(30) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. *Chem Phys* **1980**, *48*, 157.

(31) Hunt, W. J.; Goddard, W. A. *Chem Phys Lett* **1969**, *3*, 414.

(32) Huzinaga, S.; Arnau, C. *Phys Rev A* **1970**, *1*, 1285.

(33) Beebe, N. H. F. *Int J Quantum Chem* **1979**, *15*, 589.

(34) Potts, D. M.; Taylor, C. M.; Chaudhuri, R. K.; Freed, K. F. *J Chem Phys* **2001**, *114*, 2592.

(35) Chattopadhyay, S.; Chaudhuri, R. K.; Freed, K. F. *J Phys Chem A* **2011**, *115*, 3665.

(36) Toniolo, A.; Ben-Nun, M.; Martinez, T. J. *J Phys Chem A* **2002**, *106*, 4679.

(37) Shu, Y. N.; Levine, B. G. *Journal of Chemical Physics* **2013**, *139*.

(38) Shu, Y. A.; Hohenstein, E. G.; Levine, B. G. *Journal of Chemical Physics* **2015**, *142*.

(39) Slavicek, P.; Martinez, T. J. *J Chem Phys* **2010**, *132*.

(40) Lu, Z.; Matsika, S. *Abstr Pap Am Chem S* **2012**, *244*.

(41) Krylov, A. I. *Chemical Physics Letters* **2001**, *338*, 375.

(42) Shao, Y.; Head-Gordon, M.; Krylov, A. I. The Journal of chemical physics 2003, 118, 4807.

(43) Li, S. H. L.; Marenich, A. V.; Xu, X. F.; Truhlar, D. G. *J Phys Chem Lett* **2014**, *5*, 322.

(44) Bofill, J. M.; Pulay, P. *J. Chem. Phys.* **1989**, *90*, 3637.

(45) Abrams, M. L.; Sherrill, C. D. *Chem. Phys. Lett.* **2004**, *395*, 227.

(46) Lu, Z.; Matsika, S. *J. Chem. Theory Comput.* **2012**, *8*, 509.

(47) Granucci, G.; Persico, M.; Toniolo, A. *J. Chem. Phys.* **2001**, *114*, 10608.

(48) Slavicek, P.; Martinez, T. J. *J. Chem. Phys.* **2010**, *132*, 234102.

(49) Purvis, G. D.; Bartlett, R. J. *J Chem Phys* **1982**, *76*, 1910.

(50) Scuseria, G. E.; Schaefer, H. F. *Chem Phys Lett* **1987**, *142*, 354.

(51) Krylov, A. I.; Sherrill, C. D.; Byrd, E. F. C.; Head-Gordon, M. J Chem Phys 1998, 109, 10669.

(52) Bozkaya, U.; Schaefer, H. F. *J Chem Phys* **2012**, *136*.

(53) Lochan, R. C.; Head-Gordon, M. *J Chem Phys* **2007**, *126*.

(54) Neese, F.; Schwabe, T.; Kossmann, S.; Schirmer, B.; Grimme, S. J Chem Theory Comput

2009, 5, 3060.

(55) Bozkaya, U.; Sherrill, C. D. *J Chem Phys* **2013**, *138*.

(56) Bozkaya, U. *J Chem Phys* **2011**, *135*.

(57) Bozkaya, U.; Turney, J. M.; Yamaguchi, Y.; Schaefer, H. F.; Sherrill, C. D. J Chem Phys 2011, 135.

(58) Krylov, A. I.; Sherrill, C. D.; Head-Gordon, M. *J Chem Phys* **2000**, *113*, 6509.

(59) Soydas, E.; Bozkaya, U. *J Chem Theory Comput* **2013**, *9*, 1452.

(60) Kurlancheek, W.; Head-Gordon, M. *Mol Phys* **2009**, *107*, 1223.

(61) Minezawa, N.; Gordon, M. S. *J Phys Chem A* **2011**, *115*, 7901.

(62) Casanova, D. *J Chem Phys* **2012**, *137*.

(63) Xu, X. F.; Gozem, S.; Olivucci, M.; Truhlar, D. G. *J Phys Chem Lett* **2013**, *4*, 253.

(64) Zhang, X.; Herbert, J. M. *J Chem Phys* **2014**, *141*.

(65) Huix-Rotllant, M.; Natarajan, B.; Ipatov, A.; Wawire, C. M.; Deutsch, T.; Casida, M. E. Phys Chem Chem Phys 2010, 12, 12811.

(66) Minezawa, N.; Gordon, M. S. *J Phys Chem A* **2009**, *113*, 12749.

(67) Minezawa, N.; Gordon, M. S. *J Chem Phys* **2012**, *137*.

(68) Harabuchi, Y.; Maeda, S.; Taketsugu, T.; Minezawa, N.; Morokuma, K. J Chem Theory Comput 2013, 9, 4116.

(69) Ben-Nun, M.; Martinez, T. J. *Chem Phys* **2000**, *259*, 237.

(70) Serrano-Andres, L.; Merchan, M.; Lindh, R. *J Chem Phys* **2005**, *122*.

(71) Levine, B. G.; Coe, J. D.; Martinez, T. J. *J Phys Chem B* **2008**, *112*, 405.

(72) Levine, B. G.; Coe, J. D.; Virshup, A. M.; Martinez, T. J. *Chem Phys* **2008**, *347*, 3.

(73) Foresman, J. B.; Headgordon, M.; Pople, J. A.; Frisch, M. J. J Phys Chem-Us 1992, 96, 135.

(74) Headgordon, M.; Grana, A. M.; Maurice, D.; White, C. A. J Phys Chem-Us 1995, 99, 14261.

(75) Knowles, P. J.; Werner, H. J. *Chem Phys Lett* **1985**, *115*, 259.

(76) Werner, H. J.; Knowles, P. J. *J Chem Phys* **1985**, *82*, 5053.

(77) Werner, H. J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schutz, M. Wiley Interdisciplinary Reviews-Computational Molecular Science 2012, 2, 242.

(78) Werner, H. J.; Knowles, P. J. *J Chem Phys* **1988**, *89*, 5803.

(79) Langhoff, S. R.; Davidson, E. R. *Int J Quantum Chem* **1974**, *8*, 61.

Chapter 6

Efficient Electronic Structure Theories of Investigating Excited State Dynamics and Modeling Potential Energy Surfaces In The Vicinity of Conical Intersection

"The best of prophets of the future is the past"

-Lord Byron (1788-1824)

Solar energy is a promising clean energy that can be alternatives to fossil fuels. In recent years, there is an increasing demand for a comprehensive understanding of the complex photophysical processes for materials and biological systems upon excitation. This is partially driven by the shortcoming of fossil fuels and associated global warming problems. Theoretical modeling is becoming an important methodology to simulate the photophysical processes of these complex systems which are hard to be illustrated directly from the current experimental techniques. As the microscopic systems obey the quantum mechanics, the direct simulation of the equation of motion of the system through time-dependent Schrödinger equation (TDSE) is the way to obtain the photophysical behaviors. However, as both analytical and numerical exact solution for

TDSE is impossible, various approximations have been applied to obtain approximated results. Among which, non-adiabatic *ab initio* molecular dynamics (AIMD) is a powerful tool to simulate the excited state dynamics with the capability of describing the electron transitions among different electronic states.[1-22] This requires a suitable electronic structure theory to calculate the potential energy surfaces (PESs) which nucleus are propagated on. The PESs in dynamical simulations are often computed on-the-fly. A suitable electronic structure theory should provide the capability of describing the PESs efficiently, smoothly, and at least qualitatively accurately to capture the important features of the PESs. Because of the ability to describe wavefunctions of varying characters and intersections between electronic states, multi-reference electronic structure methods are attractive for this purpose. Specifically, the state averaged complete active space self-consistent field (SA-CASSCF) method[23] is a very popular choice owing to its low computational cost, ability to provide a qualitatively accurate description of complex PESs, and the availability of analytic calculations for gradients and nonadiabatic coupling vectors.[24]

However, SA-CASSCF wavefunctions contain some undesirable properties as well. For example, the energies computed on SA-CASSCF level strongly depend on the active space which is user chosen and its effect on energy is unpredictable. Also, the orbital coefficients and CI vectors are optimized simultaneously in SA-CASSCF algorithm[25] which causes problems for example difficulties in convergence, multiple solutions, PES discontinuities and unphysical symmetry breaking of the wavefunctions. All these

undesirable behaviors prevent the use of SA-CASSCF PES in AIMD.[26-29]

There exist alternatives to SA-CASSCF, which alleviates the above-mentioned problems of SA-CASSCF wavefunctions. Two-step multi-reference methods (we will denote this set of methods as two-step methods in the following text), which perform an orbital optimization step followed by a CI expansion, for example, complete active space configuration interaction (CASCI) calculation or configuration interaction singles (CIS) are a promising alternative. The two-step methods decouple the orbital coefficient and CI vector optimizations but maintains similar accuracy as compared to SA-CASSCF. Approaches based on improved Hartree-Fock virtual orbitals (IVO),[30] unrestricted Hartree-Fock natural orbitals (UNO),[31] natural orbitals of correlated wavefunction calculations,[32] natural orbitals of high-multiplicity states,[33] floating occupation molecular orbitals (FOMO),[34,35] singly excited active space self-consistent field (SEAS) orbitals,[29] CIS natural orbitals (CISNO) and its variant with relaxed density matrix (CISNOr)[36] have shown great promise as algorithmically simpler and computationally less expensive alternatives to SA-CASSCF. Figure 6-1 shows the computational time of CISNO-CASCI and SA-CASSCF performed on one sila-adamantane molecule with the same number of states being calculated. The CISNO-CASCI can be an order of magnitude faster than SA-CASSCF. The calculations are performed with MolPro software package[37,38] and single core Intel Xeon E5624 2.40GHz CPU. This computational cost difference usually holds for most of the two-step approaches as long as orbital optimization step is not computationally too expensive as compared to the CASCI step. Also promising are low-cost approaches

which use spin flip excitations[23-25] or perturbation theory[26] to describe wavefunctions of multi-reference character within a single-reference framework.

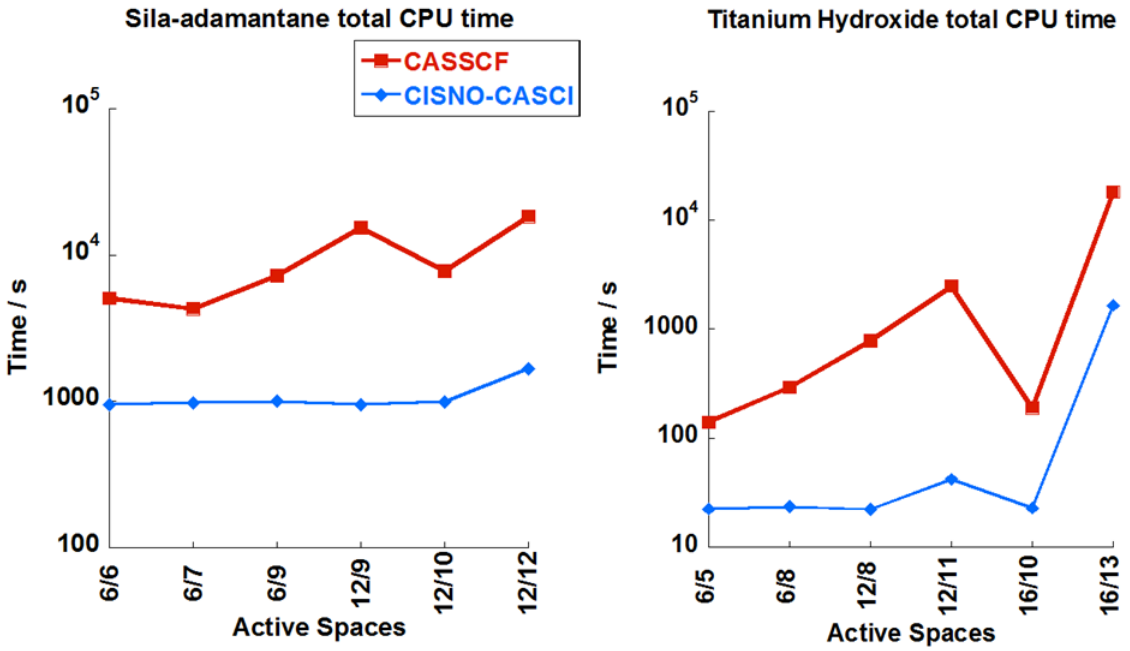


Figure 6-1. Computational cost comparison between SA-CASSCF and CISNO-CASCI for sila-adamantane and titanium hydroxide system for various active spaces.

In this chapter, I will summarize the electronic structure theories we developed in Chapter 3-5 and illustrate some of the SA-CASSCF problems into depth with mathematical and/or numerical demonstrations. Especially, I will emphasize on the arising of unphysical symmetry breaking of the wavefunctions in SA-CASSCF method and the importance of introducing natural orbitals as one of the efficient ways to constitute the

reference determinant which CI Hamiltonian is built on. These have motivated the development of SEAS-CASCI and CISNO-CASCI. Lastly, I will show some benchmark calculations to compare existing two-step methods to illustrate their behaviors on different problems for example size intensivity, wavefunction symmetry breaking, n->π* excitation and the accuracy in the vicinity of conical intersection geometry. This chapter can be treated as a supplement chapter to achieve a deeper understanding of Chapter 3-5.

6.1 The Arising of Unphysical Symmetry Breaking of The Wavefunction in SA-CASSCF Wavefunction Ansatz

6.1.1 Orbital Pseudo-Jahn-Teller Effect (OPJTE)

In Chapter 3, we have seen examples of systems which exhibit unphysical wavefunction symmetry breaking at SA-CASSCF level of theory. In many cases, the broken symmetry of the wavefunction is persistent with respect to the choice of active space. We refer the reader Chapter 3 to acquire a detailed introduction of the symmetry problem which is firstly suggested by Löwdin that the variationally optimized wavefunction does not guarantee the full symmetry of the Hamiltonian.[39] Here we present a mathematical explanation for such symmetry breaking: the orbital pseudo-Jahn-Teller effect (OPJTE).

The OPJTE can be understood by recognizing the formal equivalence of two problems familiar to computational chemists: the MCSCF problem and the problem of optimizing

a geometry on a CI PES. In both problems, the Hamiltonian matrix elements depend on a set of input variables which are varied in order to minimize the lowest eigenvalue of the Hamiltonian. The two problems differ in the identity of those variables; in MCSCF the orbital coefficients are varied, while in geometry optimization it is the atomic coordinates which are varied. It is well known that under certain circumstances the geometry optimization problem can lead to a broken symmetry result via the Jahn-Teller (JT) or pseudo-Jahn-Teller (PJT) effect.[40] As illustrated below, the same mathematics that result in the PJT symmetry breaking of the optimal geometry can lead to symmetry breaking of the optimal orbitals, as well. Unlike the JTE and PJTE, which are real physical effects, the OPJTE is an artifact of the choice of approximate MCSCF trial wavefunction. In the full CI limit, this effect would disappear. After introducing the OPTJE, we will demonstrate that spatial symmetry breaking in the simple model system He_2^+ can be explained in terms of the OPJTE. The OPJTE motivated the development of SEAS-CASCI method described in Chapter 3.

To illustrate the OPJTE, we present the following simple model. Suppose a MCSCF Hamiltonian including two electronic configurations, *A* and *B*, with matrix elements dependent on the coordinate describing the orbitals, C,

$$H(C) = \begin{bmatrix} E_A(C) & V_{AB}(C) \\ V_{AB}(C) & E_B(C) \end{bmatrix} \tag{6-1}$$

The lowest eigenvalue of $H(C)$ can be written

$$E_{-} = \frac{E_A + E_B}{2} - \sqrt{\frac{(E_A - E_B)^2}{4} + V_{AB}{}^2} \tag{6-2}$$

Solving the ground state MCSCF problem involves finding the orbital coefficients which minimize E_{-}. Note that below we develop equations assuming that $E_A > E_B$, but the labeling of states *A* and *B* is arbitrary, so this is of no significance.

We define an orbital coordinate, C, such that all orbitals transform as irreducible representations of the full point group of the molecule at $C = 0$. When $C = 0$, the configurations *A* and *B* transform as different irreducible representations, *a* and *b*. The orbital coordinate, C, is chosen such that it transforms as $(a \times b)$. By symmetry, it can be shown that

$$V_{AB}\Big|_{C=0} = \frac{dE_A}{dC}\bigg|_{C=0} = \frac{dE_B}{dC}\bigg|_{C=0} = \frac{d^2V_{AB}}{dC^2}\bigg|_{C=0} = 0 \tag{6-3}$$

$$E_{-}\Big|_{C=0} = E_B \tag{6-4}$$

and assuming a non-degenerate ground state

$$\frac{dE_{-}}{dC}\bigg|_{C=0} = 0 \tag{6-5}$$

Thus, the second derivative of E_{-} with respect to *C* is the lowest order derivative of E_{-} which is not zero by symmetry at $C = 0$. We develop an expression for this derivative to investigate the possibility of symmetry-breaking in this model MCSCF wavefunction. In the event that d^2E_{-}/dC^2 is less than zero, the eigenvalue, E_{-}, would be reduced by the breaking of orbital symmetry. Thus simultaneous variational optimization of the orbital and CI coefficients would result in a symmetry-broken solution.

Making use of eqs. (6-3), (6-4), and (6-5), it can be shown that

$$\left.\frac{d^2E_-}{dC^2}\right|_{C=0} = \left.\frac{d^2E_B}{dC^2}\right|_{C=0} - \frac{2}{(E_A(0)-E_B(0))}\left(\left.\frac{dV_{AB}}{dC}\right|_{C=0}\right)^2 \qquad (6\text{-}6)$$

Inspecting this equation, we can see that even in the case that the diagonal elements of the MCSCF Hamiltonian are stable with respect to *C* ($d^2E_A/dC^2\big|_{C=0} > 0$, $d^2E_B/dC^2\big|_{C=0} > 0$), it is still possible for $d^2E_-/dC^2\big|_{C=0}$ to be less than zero, thus leading to a variational solution of broken symmetry. This occurs when the second term in the right-hand side of eq. (6-6) is large and negative. Analogous to the pseudo Jahn-Teller effect, this is true when $\left(E_A(0)-E_B(0)\right)$ is small, and dV_{AB}/dC is large in magnitude.

The principle illustrated by this simple two-configuration model problem can be generalized to an arbitrary number of configurations, and to state-averaged MCSCF calculations, where coupling terms between states included in the state average and those outside it could promote symmetry breaking. In fact, in large systems with high densities of states, it may prove difficult or impossible to choose a number of states to average over such that the energy gap between states is large enough to prevent symmetry breaking.

One strategy to avoid unphysical symmetry breaking caused by the OPJTE would be to minimize the average energy of all eigenvalues of the Hamiltonian rather than minimize a single energy or the average energies of a select few configurations. This is equivalent to determining the orbitals that minimize the trace of the Hamiltonian, and this trace does not depend on the off-diagonal matrix elements:

$$E_{ave} = \frac{1}{N_{config}} \sum_{I=1}^{N_{config}} E_I = \frac{1}{N_{config}} Tr(\mathbf{H}) = \frac{1}{N_{config}} \sum_{A=1}^{N_{config}} E_A \quad (6\text{-}7)$$

Here N_{config} is the total number of configurations in the CI Hamiltonian, *I* indexes eigenvalues of the Hamiltonian, and *A* indexes the configurations used to build the Hamiltonian. The second derivative of E_{ave} depends only on the second derivative of the diagonal elements of the Hamiltonian,

$$\frac{d^2 E_{ave}}{dC^2} = \frac{1}{N_{config}} \sum_{A=1}^{N_{config}} \frac{d^2 E_A}{dC^2} \quad (6\text{-}8)$$

As such, there can be no OPJTE between configurations included in the Hamiltonian; if the average energy of the diagonal elements of the Hamiltonian is stable with respect to orbital symmetry breaking, the optimal orbitals will reflect the full molecular symmetry regardless of the magnitude of dV_{AB}/dC or the size of the energy gaps between states.

6.1.2 Symmetry breaking in He_2^+ as an example of OPJTE

The He_2^+ radical cation is a simple model system which exhibits unphysical spatial symmetry breaking when its wavefunction is approximated as a single Slater determinant.[41,42] The symmetry-preservation of the wavefunction (or lack thereof) is determined by a balance between two competing contributions to the energy: *resonance effects* and *orbital size effects*. Resonance effects describe the favorable energy associated with overlap between atomic orbitals on neighboring atoms (covalent bond formation). Orbital size effects reflect the relaxation of the orbital shape with respect to the instantaneous number of electrons localized near a given atom.

At short internuclear separation, the variationally optimal spin-restricted single

determinant wavefunction of He_2^+ is constructed by doubly occupying the σ-bonding orbital and singly occupying the σ-antibonding (σ^*) orbital. In this case, both orbitals reflect the full symmetry of the molecule, and thus so does the many-electron wavefunction. This symmetry-preserving solution is favored by strong resonance effects. As the internuclear separation is increased, resonance effects are weakened, and the variationally optimal single determinant wavefunction becomes one of broken symmetry where the s orbital on one He atom is doubly-occupied, while the s orbital on the other He atom is singly-occupied. This solution is favored by the relaxation energy of the s orbitals; the doubly-occupied s orbital becomes more diffuse relative to the symmetric solution, while the singly-occupied s orbital shrinks towards the nucleus. These correspond to the above-mentioned orbital size effects.

Though this behavior can easily be explained within this single determinant wavefunction picture, here we point out that the spatial symmetry breaking in He_2^+ can also be viewed as an example of the OPTJE if we change the representation of the wavefunction. Here we model the He_2^+ wavefunction at the SA-CASSCF level of theory with three active electrons in two active orbitals (CAS(3/2)). The CAS(3/2) Hamiltonian includes two Slater determinants, any arbitrary mixture of which can itself be represented as a single Slater determinant. Thus, the optimal state-specific CAS(3/2) wavefunction is identical to that of the optimal restricted open-shell HF wavefunction described above. Because it is constructed in a basis of two determinants, the CAS(3/2) Hamiltonian takes exactly the form in eq. (6-1). These calculations were performed using the double-zeta

6-31G basis because it is the smallest basis which allows relaxation of the s orbital sizes (orbital size effects). An internuclear separation of 2.452 Angstrom was used.

This SA-CASSCF problem was solved both with and without symmetry constraints on the wavefunction leading to two different sets of orbitals. Keeping with eq. (6-1) we define a coordinate, *C*, connecting these two sets of orbitals. The symmetry constrained and fully optimized orbitals correspond to $C = 0$ and $C = \delta C$, respectively. The symmetry-constrained solution ($C = 0$, Figure 6-2, bottom left) is composed of symmetry-preserving (σ and σ*) orbitals. The fully optimized solution is not symmetry-preserving, as evidenced by a non-zero dipole moment. In order to understand this calculation in terms of the OPJTE, we applied an orbital rotation to the fully optimized active orbitals to maximize their overlap with the symmetry-preserving σ and σ* orbitals determined in the symmetry-constrained calculation (orbital diabatization; the diabatized orbitals are pictured in Figure 6-2, bottom right). The distortion of the symmetry-broken orbitals relative to the symmetry-preserving ones is very small; the overlap matrix elements $\left\langle \phi_{\sigma,C=0} \middle| \phi_{\sigma,C=\delta C} \right\rangle$ and $\left\langle \phi_{\sigma^*,C=0} \middle| \phi_{\sigma^*,C=\delta C} \right\rangle$ are both greater than 0.995.

The eigenvalues and diagonal matrix elements of the 2-by-2 Hamiltonian as a function of *C* are plotted in Figure 6-2. The eigenvalue and diagonal matrix element curves can be thought of as PESs of adiabatic and diabatic states, respectively, in orbital space. The breaking of orbital symmetry (displacing the orbitals from $C = 0$ to $C = \delta C$) increases the diagonal Hamiltonian matrix elements. Taking the identities in eq. (6-3) into account and assuming that this orbital displacement is small enough that

the matrix elements depend approximately quadratically on *C*, this positive change in the matrix elements indicates that they are stable with respect to orbital symmetry breaking ($d^2E_A / dC^2\big|_{C=0} > 0$, $d^2E_B / dC^2\big|_{C=0} > 0$). Though breaking the orbital symmetry increases the values of the diagonal Hamiltonian matrix elements, it introduces a splitting that decreases the value of the lowest eigenvalue of the Hamiltonian significantly. Thus dV_{AB} / dC is large enough to introduce the OPJTE, and therefore spatial symmetry breaking of the wavefunction. The resulting eigenvalue curves look identical to the typical PJTE PESs, with nuclear coordinates replaced by orbital displacements. As predicted above, performing a state-averaged CASSCF calculations (averaging over both states) results in a symmetry-preserving wavefunction.

As can be seen in the bottom panel of Figure 6-2, the OPJTE is a reflection of the same orbital size effects which are invoked when discussing symmetry breaking in the orbital frame. It is important to note that orbitals size effects are a real physical effect that can be accurately described by a symmetry-preserving correlated wavefunction (such as in the nonorthogonal-CI approach[42]). Enforcing symmetry on the wavefunction by constraint or by state averaging involves a trade-off between correctly describing these effects and symmetry preservation. It can be argued that in many cases symmetry-broken solutions are preferable to the symmetry-preserving one.[43-47] However, in other cases (such as in AIMD calculations—where a smooth PES is highly desirable—and when the initial wavefunction is to be used as a reference for a subsequent treatment of electron correlation) the symmetry-preserving solution is more desirable, and it is with these cases

in mind that we have developed SEAS-CASCI to decrease the propensity for unphysical spatial symmetry breaking in compact multi-reference wavefunctions.

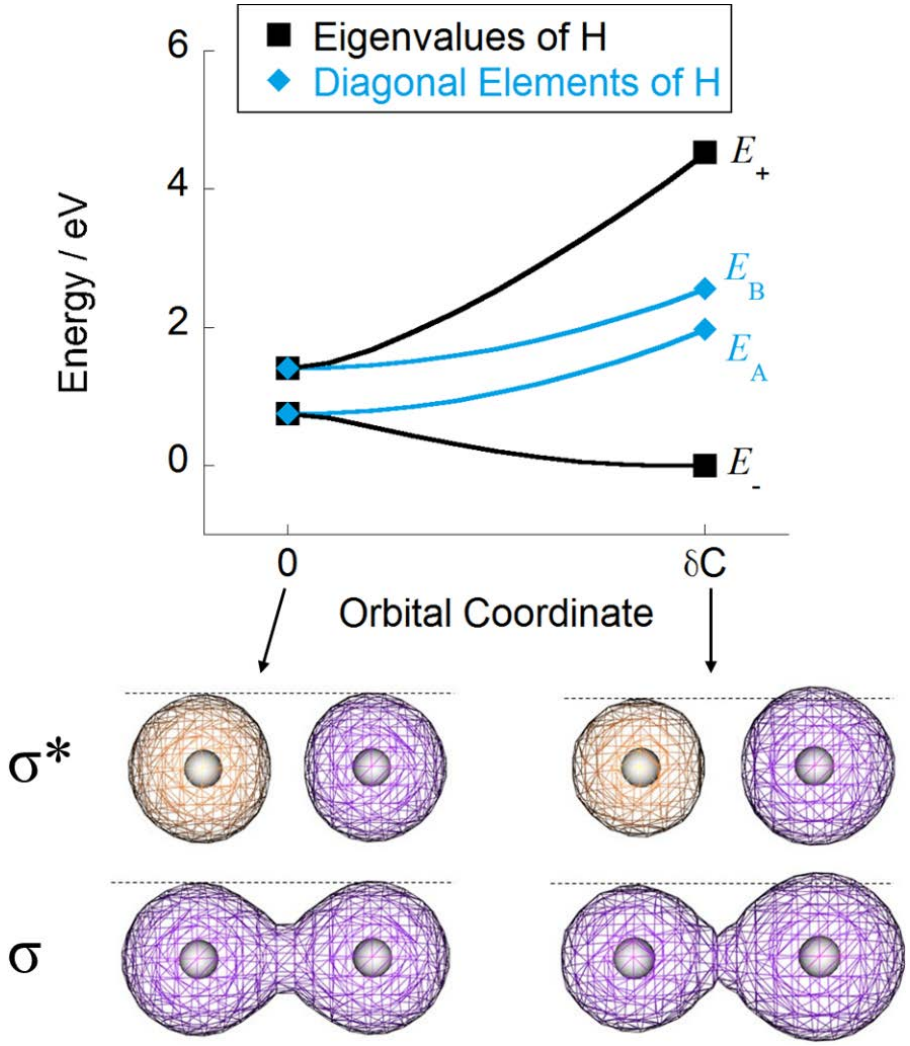


Figure 6-2. Upper panel: Black shows the eigenvalues (E_+ and E_-) of the CAS(3/2)/6-31G Hamiltonian describing He2+ as a function of a distortion of the orbitals from variationally optimal symmetry constrained orbitals ($C=0$) to fully optimized orbitals (no symmetry constraint, $C=\delta C$). The diagonal matrix elements of the Hamiltonian (E_A and E_B) are presented in blue. Despite the fact that the diagonal elements of the

Figure 6-2 (cont'd). Hamiltonian are increased by symmetry breaking, the variationally optimal solution ($E_{-}(\delta C)$) is symmetry-broken because symmetry breaking introduces strong coupling between configurations A and B. The square markers indicate data that was calculated, while the lines connecting them are interpolated assuming quadratic behavior of the Hamiltonian matrix elements. Lower panel: The orbital basis from which the CAS(3/2) Hamiltonian is constructed. The dashed lines guide the eye to see that the orbitals are symmetric when $C=0$ and symmetry-broken when $C=\delta C$. Configuration A corresponds to double occupation of the σ orbital and single occupation of the σ* orbital; the occupations are reversed in configuration B.

6.1.3 Singly-excited active space self-consistent field method

SEAS-CASCI is a two-step approach to multi-reference quantum chemistry designed to decrease the incidence of wavefunction symmetry breaking in the calculation of low-lying singlet excited states (see Chapter 3 for a rigorous definition of the method). In the first step, a SEAS SCF calculation determines a set of orbitals. The SEAS SCF step minimizes the averaged energy of all eigenvalues of the Hamiltonian, and hence the symmetry breaking problem is alleviated. The initial orbital determination step is followed by the construction and diagonalization of a CASCI Hamiltonian. Though developed below for singlet states, the OPJTE could trivially be used to develop similar methods to describe states of other spin symmetries.

SEAS-CASCI is suitable for calculating the PESs near conical intersection as the CASCI

Hamiltonian contains doubly excited determinants which couple to the closed-shell Hartree-Fock (HF) determinant.[29,48] In addition, the SEAS orbital couples the occupied and virtual orbitals in HF reference determinant and hence semi-quantitative accurate PES can be obtained in the vicinity of ground and excited states conical intersection geometry (see Chapter 5 for detailed discussion of the importance of the virtual occupied coupling in describing conical intersection region PES and the definition of semi-quantitative accuracy). What is more, the SEAS energy can be expressed in a generalized HF expression, and hence the SEAS wavefunction obeys the Hellmann-Feynman theorem for orbital rotation parameters. Thus, the energy gradient of the SEAS-CASCI approach may be easy to compute.

6.2 Effective Selection of The Determinants in Configuration Interaction Hamiltonian By Utilizing Natural Orbitals

6.2.1 SEAS and HF orbitals are not always most efficient orbitals in achieving small active spaces

The SEAS-CASCI approach provides us symmetry preserved wavefunctions, however in some cases, a large active space cannot be avoided. For example, $Ti(OH)_4$ requires 16/10 active space to achieve accurate excitation energy and symmetry preserved wavefunction. This is because when small active spaces are employed, the orbitals outside the active space has strong contributions to the energy of the state we are interested in by looking at the EOM-CCSD amplitudes and HF orbital characters. That is to say, SEAS orbitals do

not always pertain an orbital order that those most important orbitals of describing low-lying excited states are ranked closest to highest occupied molecular orbitals (HOMOs) and lowest unoccupied molecular orbitals (LUMOs). The SEAS orbitals are ordered according to the diagonal elements of the effective Fock operator. This is not surprising as SEAS orbitals are optimized by minimizing the averaged energy of all singly excited determinants in the active space, and some of the singly excited determinants are not main excitation characters of the low-lying excited states. To investigate this assumption we discussed above, we did restricted active space configuration interaction (RASCI) calculations upon SEAS orbitals. In SEAS-RASCI approach, we restrict the orbital occupation numbers in the CASCI expansion, that is to say, certain orbitals are forbidden to form excited determinants in configuration interaction expansion. We will denote these orbitals that have occupation numbers always being zero or two in the CASCI expansion as excited forbidden orbitals (EFOs) in the following text. Then we compare the SEAS-RASCI low-lying excited states energies to trustworthy EOM-CCSD results. If the difference is large, that means the orbitals forbidden being excited are important orbitals for that specific excited state. The results are shown in Figure 6-3. Figure 6-3(a) shows the $Ti(OH)_4$ closed shell determinant constituted by SEAS orbitals with indexed orbital orders. The orbital degeneracy is represented graphically by the vertical alignment. When SEAS-RASCI approach is employed, the EFOs we choose do not break the degeneracy of the orbitals. Figure 6-3(b) shows the absolute energy difference of SEAS-RASCI and EOM-CCSD for $Ti(OH)_4$ molecule. It is noteworthy that SEAS-RASCI with

EFO 22 and 23 (EFO(22,23)) has larger error compared with EFO(24,25,26). This indicates that determinants formed by excitations from orbital 22 and 23 have larger contributions to the energy of S_1 state as compared to orbital 24, 25 and 26. EFO(27,28,29) has a large difference as they are degenerate HOMOs, which is not surprising to be important in describing S_1. This observation explains the fact that 16 electrons in the active space is required in SEAS-CASCI to obtain accurate and symmetry preserved S_1 state as orbital 22 and 23 should be included in the active space. The last column of Figure 6-3(b) shows the absolute energy difference between the SEAS-CASCI with 16/13 active space and EOM-CCSD. As can be seen, the difference is largely reduced when we include all the determinants in the 16/13 active space. $Ti(OH)_4$ is not a unique molecule that has such results for SEAS orbitals. Figure 6-3(c) and (d) show the similar behavior for HF orbitals of $Ni(CO)_4$ molecule. We can see that orbitals 35,36 and 37 has a much larger contribution to S_1 state than orbital 38 and 39. This will also happen for virtual orbitals. As shown in Figure 6-3(d) EFO(49,50,51) has much larger difference compared with EOM-CCSD result than EFO(43,44,45) or EFO(46,47,48). Both results indicate that an active space 16/17 will be possibly required to achieve accurate S_1 excitation energy with preserved symmetry for $Ni(CO)_4$ molecule. Unfortunately the HF-CASCI calculation with 16/17 active space for $Ni(CO)_4$ molecule is beyond our current computational power with the existing codes.

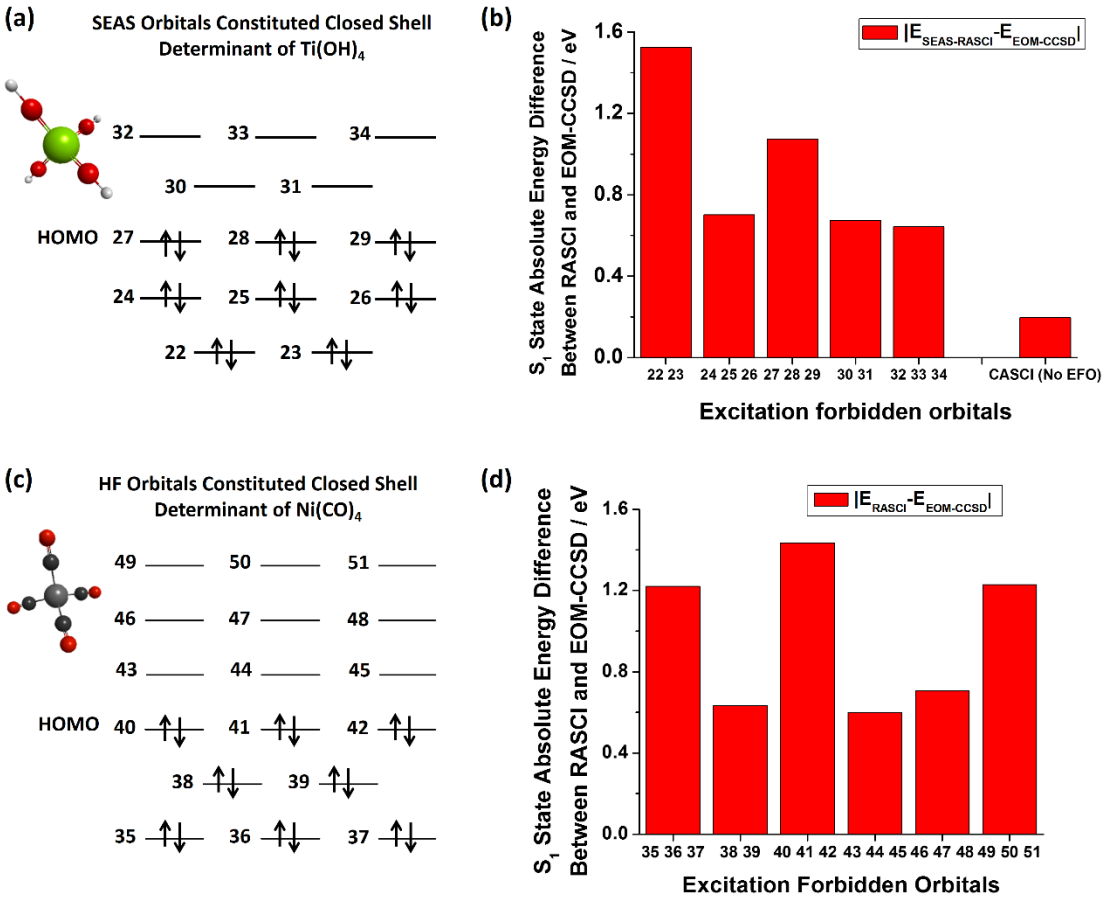


Figure 6-3. (a) The SEAS orbital of $Ti(OH)_4$ and associated closed shell determinant constituted by SEAS orbitals, the number indexes the orbitals with the vertical alignment of the orbitals indicate the degeneracy. The inset shows the $Ti(OH)_4$ molecule. (b) The absolute difference between SEAS-RASCI and EOM-CCSD of $Ti(OH)_4$ in the unit of eV. RASCI is defined in the text. (c) The SEAS orbital of $Ni(CO)_4$ and associated closed shell determinant constituted by SEAS orbitals, the number indexes the orbitals with the vertical alignment of the orbitals indicate the degeneracy. The inset shows the $Ni(CO)_4$ molecule. (b) The absolute difference between SEAS-RASCI and EOM-CCSD of $Ni(CO)_4$ in the unit of eV. RASCI is defined in the text.

However, the above phenomena do not mean that all the molecules will have such SEAS or HF orbitals, the purpose is to show that it is possible to happen. What is more important is that for low-lying excited states, these important orbitals are not always being rotated with closest orders to HOMO and LUMO and result in the necessity of employing large active space to improve the accuracy of SEAS-CASCI approach. This conclusion motivated the development of CISNO-CASCI approach which has been detailed discussed in Chapter 4.

6.2.2 Configuration interaction singles natural orbitals form an efficient orbital space for configuration interaction expansion

The CISNO is defined by diagonalizing the state-averaged (SA-) CIS and HF first order reduced density matrices,

$$\gamma_{\mu\upsilon}^{\text{SA-CIS}} = \frac{1}{N+1}\left(\gamma_{\mu\upsilon}^{\text{HF}} + \sum_{I=1}^{N}\gamma_{\mu\upsilon}^{\text{CIS}(I)}\right) \tag{6-9}$$

where $\gamma_{\mu\upsilon}^{\text{HF}}$ and $\gamma_{\mu\upsilon}^{\text{CIS}(I)}$ are elements of the first order reduced density matrices of the HF ground state and the *I*th CIS singlet excited state, respectively, and lowercase Greek letters index basis functions. *N* represents the number of low-lying singlet excited states employed in SA-CIS. The CIS first order reduced density matrix is defined

$$\gamma_{\mu\upsilon}^{CIS(I)} = \left\langle \Psi_{CIS(I)} \middle| a_{\mu}^{+} a_{\upsilon} \middle| \Psi_{CIS(I)} \right\rangle \tag{6-10}$$

where $\left|\Psi_{CIS(I)}\right\rangle$ is the wavefunction of the *I*th CIS excited state. Because $\boldsymbol{\gamma}^{\text{SA-CIS}}$ represents an approximate average first order reduced density matrix for the low-lying excited states of interest, diagonalizing it provides an efficient natural orbital basis in

which to expand those states.

To see whether the CISNO or CISNO with relaxed density matrix (CISNOr) has the capability of rotating the orbital orders such that those important orbitals are ordered closest to HOMO and LUMO, we computed CISNOr-RASCI energies and compare them with EOM-CCSD results. Figure 6-4 is a similar figure as Figure 6-3 with the same molecules and basis sets applied in the calculations, instead of employing SEAS or HF orbitals, we employed CISNOr. It can be seen from Figure 6-4(b) that the first excited state excitation energy absolute differences between CISNOr-RASCI and EOM-CCSD for $Ti(OH)_4$ molecule are largest when we forbid excitations from degenerate HOMOs (orbital 27, 28, 29) or degenerate LUMOs (30,31). In addition, these energy differences are smaller if the EFOs are farther from HOMO or LUMO in terms of orbital orders. For example, the absolute energy difference with EFO(22,23,24), EFO(25,26), EFO(27,28,29) are 0.295, 0.736 and 1.709 eV respectively. This behavior also holds for the $Ni(CO)_4$ molecule, as can be seen from Figure 6-4(d). For example, the S_1 excitation energy absolute difference between CISNOr-RASCI with EFO(43,44,45), EFO(46,47,48),EFO(49,50,51) and EOM-CCSD are 1.134, 0.232, 0.118 eV respectively. We have also computed CISNOr-CASCI energy for $Ti(OH)_4$ molecule by employing 16/13 active space for comparison. This is shown in the last column of Figure 6-4(b). The absolute excitation energy difference between CISNOr-CASCI(16/13) and EOM-CCSD is 0.292 eV which is larger than CISNOr-RASCI with EFO(32,33,34). This indicates that there exist some error cancellations in the CISNOr-RASCI or CISNOr-CASCI approach. Otherwise

CISNOr-RASCI should not be more accurate than CISNOr-CASCI(16/13). However, this small error should not hinder the attempts of employing CISNOr as a set of efficient orbitals in describing the low-lying excited states. The above calculations suggest that when CISNOr is employed, we may only need 10/7 active spaces for $Ti(OH)_4$ molecule and 6/6 active space for $Ni(CO)_4$ molecule to obtain accurate excitation energies at CISNOr-CASCI approach.

Numerical results for $Ti(OH)_4$ molecule shows that SEAS-CASCI with 16/13 active space predicts 6.076 eV S_1 excitation energy as compared with 6.271 eV predicted at EOM-CCSD theoretical level. This is 0.195 eV discrepancy. However, CISNOr-CASCI with 10/7 active space predicts 6.165 eV S_1 excitation energy indicates 0.106 eV error. The 10/7 active space is employed because the order of CISNOr is different from that of SEAS orbitals. SEAS predicts orbitals as (27,28,29), (24,25,26), (22,23) with the degenerated orbitals listed in the same parenthesis, while CISNOr changes the order to (27,28,29), (25,26), (22,23,23) as can be seen from Figure 6-3(a) and Figure 6-4(a). And hence a relatively small active space 10/7 can be employed to achieve high accuracy. This is also supported by Figure 6-4(b) which shows a large error of CISNOr-RASCI approach compared with the EOM-CCSD result when EFO(25,26), EFO(27,28,29) and EFO(30,31) are employed. The reduction of active space size with CISNOr-CASCI approach becomes more significant for $Ni(CO)_4$ molecule. The CISNOr-CASCI approach employs 6/6 active spaces predicts 4.818 eV S_1 excitation energy as compared with 4.910 eV predicted at EOM-CCSD theoretical level. This indicates 0.092 eV error. However, SEAS-CASCI with active space

as large as 10/11 and 16/14 predicts 5.144 and 5.199 eV excitation energy of S_1 state which show a 0.234 and 0.289 eV error. HF-CASCI predicts 5.956 and 5.919 eV S_1 excitation energy which indicates an even larger error as HF orbitals are optimal orbitals for ground state optimized with single determinant. As suggested from Figure 6-3(d), a bigger active space for SEAS-CASCI and HF-CASCI is required for an accurate excitation energy. However, this active space is already beyond our current computational power with the current hardware and codes. The S_1 excitation energies as a function of various active spaces at CISNOr-CASCI, SEAS-CASCI, and HF-CASCI theoretical levels are summarized in Figure 6-5 (a) and (c) with EOM-CCSD results for comparison. Figure 6-5 (b) and (d) lists the different active spaces employed, and these active spaces are chosen without breaking the degeneracies of the orbitals.

We will not discuss CISNO here into detail as it behaves very similar with CISNOr. That is to say, CISNO also ranks the order of orbitals such that those most important orbitals for low-lying excited states are closest to HOMO and LUMO as can be seen from Figure 6-6. Figure 6-6 is similar with Figure 6-3 and Figure 6-4 but by employing CISNO.

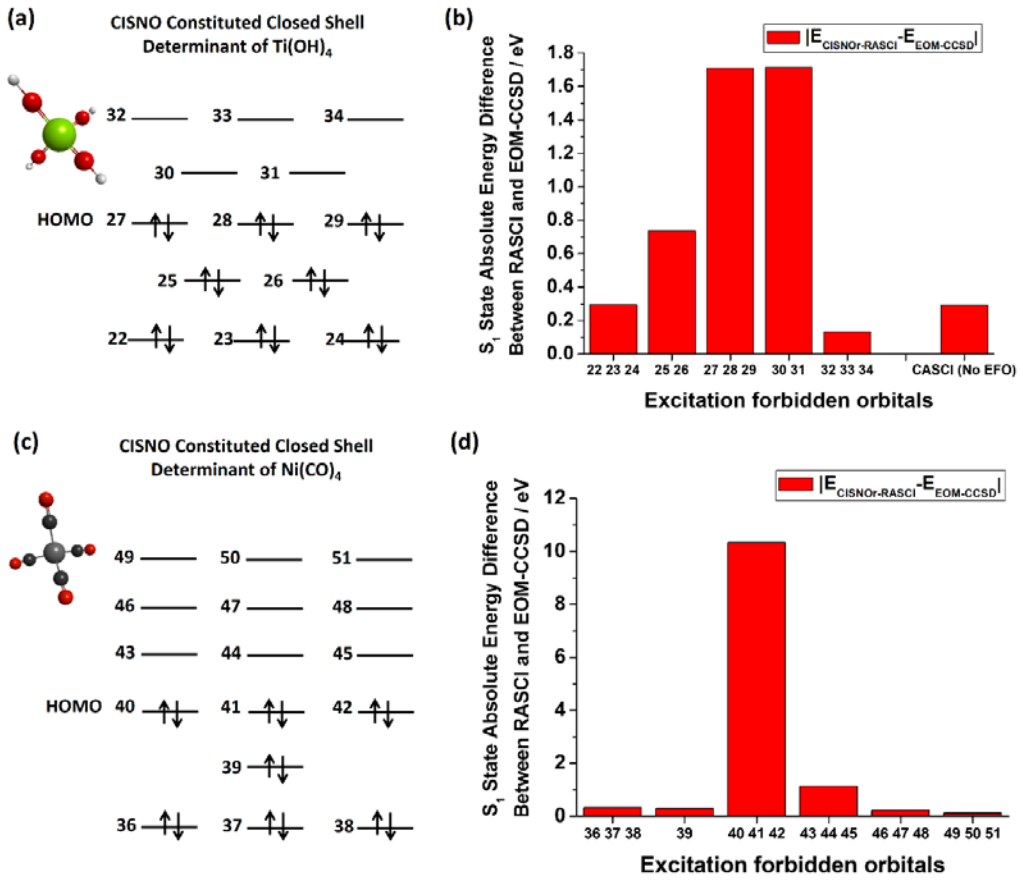
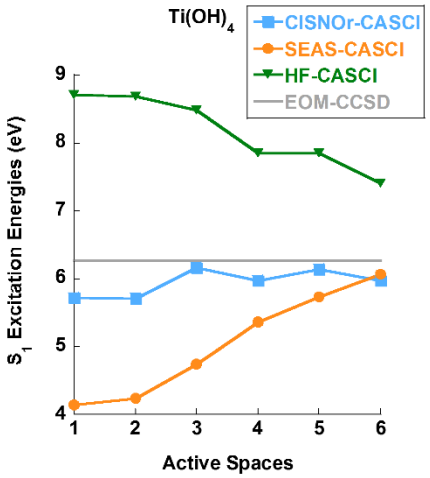


Figure 6-4. (a) The CISNOr of $Ti(OH)_4$ and associated closed shell determinant constituted by CISNOr, the number indexes the orbitals with the vertical alignment of the orbitals indicate the degeneracy of the occupation numbers of CISNOr. The inset shows the $Ti(OH)_4$ molecule. (b) The absolute difference between CISNOr-RASCI and EOM-CCSD of $Ti(OH)_4$ in the unit of eV. RASCI is defined in the text. (c) The CISNOr of $Ni(CO)_4$ and associated closed shell determinant constituted by CISNOr, the number indexes the orbitals with the vertical alignment of the orbitals indicate the degeneracy of the occupation numbers of CISNOr. The inset shows the $Ni(CO)_4$ molecule. (b) The absolute difference between CISNOr-RASCI and EOM-CCSD of $Ni(CO)_4$ in the unit of eV. RASCI is defined in the text.

Both numerical results for $Ti(OH)_4$ and $Ni(CO)_4$ have demonstrated the efficiency of CISNO and CISNOr in CASCI expansion with small active space to achieve accurate low-lying excited state excitation energies. Constructed by an ensemble-averaging procedure without iteratively re-optimizing the orbitals in CISNO in addition with the fact that $\boldsymbol{\gamma}^{\text{SA-CIS}}$ and $\boldsymbol{\gamma}^{HF}$ commute as proven in Chapter 5, symmetry breaking is not expected for CISNO-CASCI. Figure 6-7 has shown the PESs generated by asymmetric stretching coordinates for $Be(OH)_2$, linear $Ti(OH)_4$ and bent $Ti(OH)_4$ molecules at CISNO-CASCI, SA-CASSCF, and EOM-CCSD theoretical levels. The smooth CISNO-CASCI PES indicates the symmetry preserved wavefunction is achieved in CISNO-CASCI calculations.

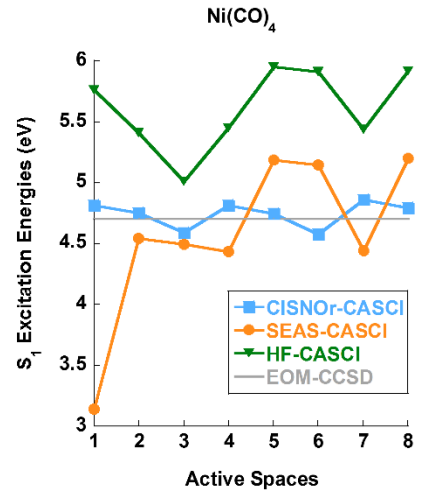
(a)



(b)

Active Spaces	CISNOr-CASCI	SEAS-CASCI	HF-CASCI
1	6/5	6/5	6/5
2	6/8	6/8	6/8
3	10/7	12/8	12/8
4	10/10	12/11	12/11
5	16/10	16/10	16/10
6	16/13	16/13	16/13

(c)



(d)

Active Spaces	CISNOr-CASCI	SEAS-CASCI	HF-CASCI
1	6/6	6/6	6/6
2	6/9	6/9	6/9
3	6/12	6/12	6/12
4	8/7	10/8	10/8
5	8/10	10/11	10/11
6	8/13	10/14	10/14
7	14/10	16/11	16/11
8	14/13	16/14	16/14

Figure 6-5. The S_1 excitation energy as a function of active spaces for (a) $Ti(OH)_4$ and (c) $Ni(CO)_4$ molecules calculated at CISNOr-CASCI, SEAS-CASCI, HF-CASCI and EOM-CCSD theoretical levels. (b) and (d) lists the active spaces employed in CISNOr-CASCI, SEAS-CASCI and HF-CASCI.

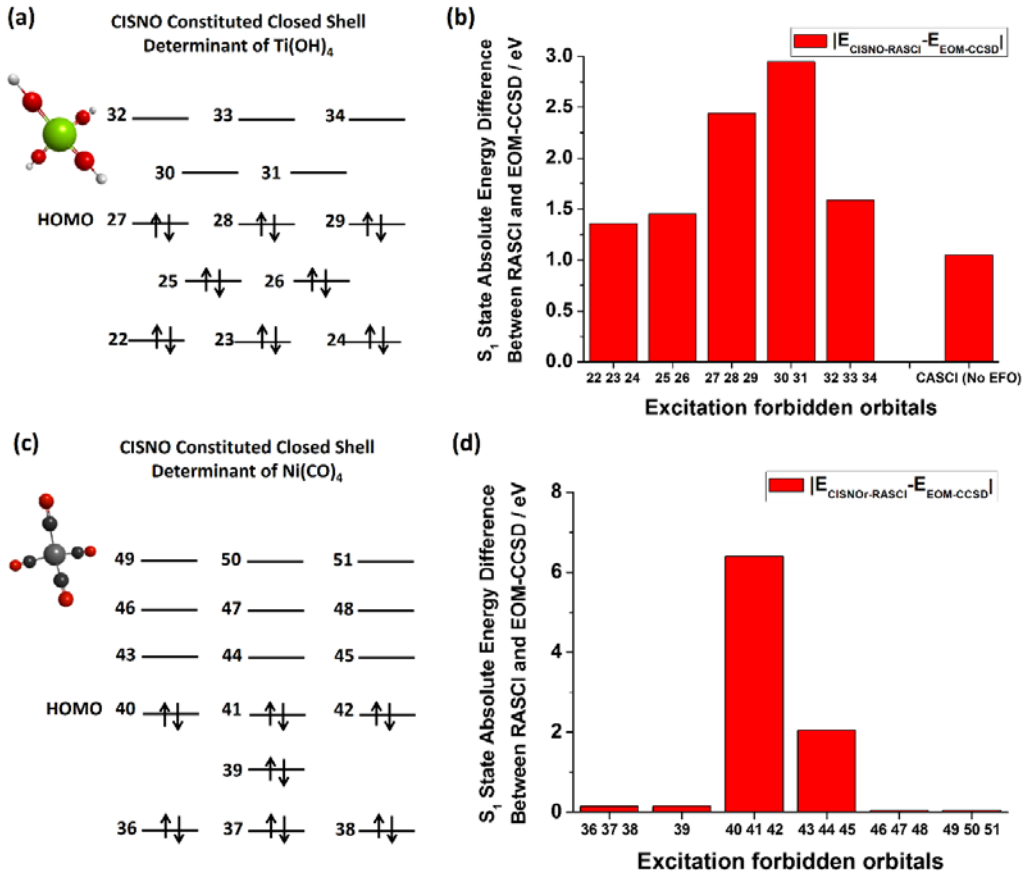


Figure 6-6. (a) The CISNO of $Ti(OH)_4$ and associated closed shell determinant constituted by CISNO, the number indexes the orbitals with the vertical alignment of the orbitals indicate the degeneracy of the occupation numbers of CISNO. The inset shows the $Ti(OH)_4$ molecule. (b) The absolute difference between CISNO-RASCI and EOM-CCSD of $Ti(OH)_4$ in the unit of eV. RASCI is defined in the text. (c) The CISNO of $Ni(CO)_4$ and associated closed shell determinant constituted by CISNO, the number indexes the orbitals with the vertical alignment of the orbitals indicate the degeneracy of the occupation numbers of CISNO. The inset shows the $Ni(CO)_4$ molecule. (b) The absolute difference between CISNO-RASCI and EOM-CCSD of $Ni(CO)_4$ in the unit of eV. RASCI is defined in the text.

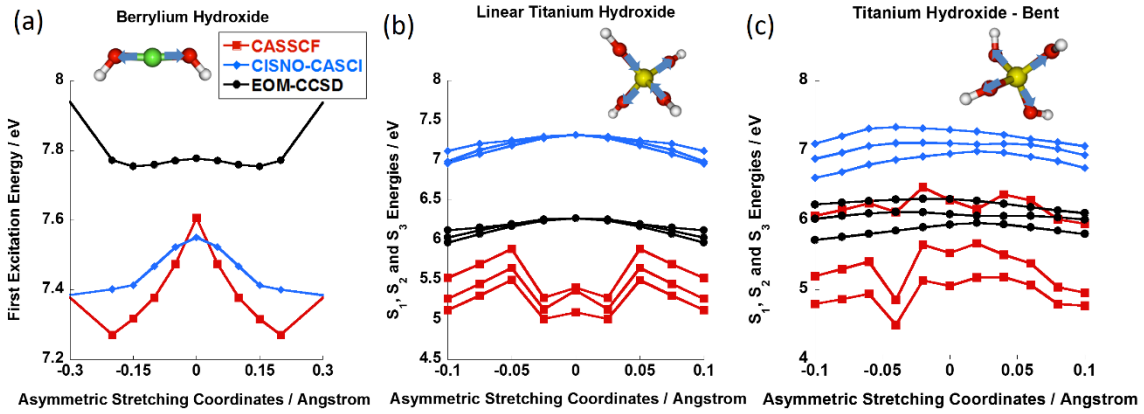


Figure 6-7. (a) The PES of the first excited state of beryllium hydroxide as a function of a rigid asymmetric stretching (AS) coordinate (as illustrated in the inset) calculates at three levels of theory. CASSCF predicts a large cusp in the PES at the C_{2v} symmetric (zero AS) geometry, while the CISNO-CASCI surface is much smoother with respect to AS. The shape of the CISNO-CASCI curve is in much better agreement with the symmetry-preserving EOM-CCSD reference. (b) and (c) The PESs for the linear $Ti(OH)_4$ and bent $TI(OH)_4$ molecules as a function of an asymmetric stretching coordinate (illustrated in the insets). The CISNO-CASCI PESs (blue) are smoothly varying and in qualitative agreement with EOM-CCSD (black), while the CASSCF surfaces (red) have unphysical cusps at the T_d symmetric geometry for linear $Ti(OH)_4$ and is in poor agreement with EOM-CCSD.

6.3 Benchmark Calculations for Two-step Methods

6.3.1 Size intensivity

With the increasing interests toward nanoscale systems in recent years, investing the scaling of the energies with system size is necessary. For photochemistry, we are specifically interested in excited state properties, for example, the size dependence of the excitation energy in systems composed of multiple excitable subsystems, like conjugated polymers, organic semiconductors, or defective nanoparticles. As has been noted in Chapter 4, the scaling of energies with system size depends not only on the CI or cluster expansion employed but also on the choice of orbitals. Here, we will use the same terminology as we introduced in Chapter 4.

We summarize the size intensivity behavior for the various two-step methods here and the results are shown in Figure 6-8. Figure 6-8 shows the first excited state excitation energy as a function of the number of non-interacting hydrogen molecules. All the two-step methods are represented as *orbital optimization method-configuration interaction expansion*, same as we have introduced in Chapter 5.

Among various methods, SA-CASSCF, SEAS-CACI, IVO-CASCI are not size intensive, CISNO-CASCI, FOMO-CASCI, and HF-CASCI are size intensive as shown in Figure 6-8. The arising of the size intensivity behaviors for these six methods have been discussed in Chapter 4 and we will not discuss it again here.

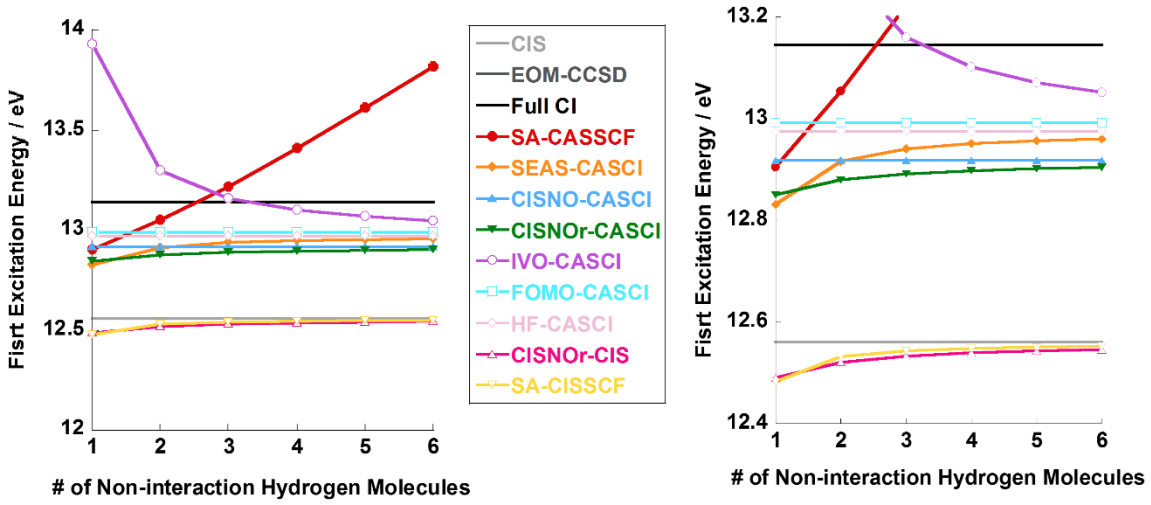


Figure 6-8. (a) The excitation energies of the monomer-localized singly excited states of systems of non-interacting H_2 molecules computed at the CIS, EOM-CCSD, Full CI, SA-CASSCF, SEAS-CASCI, CISNO-CASCI, CISNOr-CASCI, IVO-CASCI, FOMO-CASCI, HF-CASCI, CISNOr-CIS and SA-CISSCF theoretical levels as a function of the number of non-interacting subsystems are shown in grey, dark grey, black, red, orange, blue, green, purple, cyan, light red, magenta and yellow respectively. (b) Same as (a), but with a narrower energy scale for clarity.

As we have noted in Chapter 4, in various two-step methods, the non-size intensive method (except those employs CISNOr which will be discussed separately in the following text) will asymptotically get close to the corresponding *ground state orbital optimization method-configuration interaction expansion* result as the number of non-interaction subsystems increases. The *ground state orbital optimization method* here specifically means the optimal orbitals for the ground state. That is to say, the orbitals are not

optimized in a state averaged procedure, instead, optimized in a state specific procedure. For example, the corresponding *ground state orbital optimization method* for SA-CASSCF is the ground state state specific (SS-) CASSCF; for SA-CISSCF is the ground state SS-CISSCF. The ground state SS-CASSCF orbitals are CASSCF orbitals. The ground state SS-CISSCF orbitals are HF orbitals. Let us analyze the ground state SS-CISSCF approach here into detail to understand why the ground state SS-CISSCF orbitals are HF orbitals. At the first iteration of ground state SS-CISSCF approach, we build a CIS Hamiltonian based on HF orbitals. Due to the zero coupling between HF closed shell determinant and singly excited determinants, the ground state energy obtained by diagonalizing the Hamiltonian is just HF energy. Now we want to optimize the orbitals to minimize this ground state energy which is HF energy, and we know that HF orbitals are already the orbitals such that HF energy is minimized if the wavefunction is expanded with closed shell and singly excited determinants only. And hence, in ground state SS-CISSCF procedure, the orbitals will not be re-optimized. It is worth noting here that the above discussion is only validated when the CIS state has higher energy compared with HF state. It can sometimes happen that CIS state energy can be lower than HF state when the geometry is close to the conical intersection. Why can SA-CISSCF be employed to optimize the orbitals that are different from HF orbitals? Consider a SA2-CISSCF approach, at the first iteration, we build a CIS Hamiltonian based on HF orbitals. However, the HF orbitals are not the optimal orbitals such that the averaged HF and the first CIS state energies are minimized. And hence, the orbitals will be re-optimized at SA2-CISSCF theoretical level.

It is hence can be understood that the corresponding ground state state specific version of SEAS orbitals are HF orbitals as well.

The asymptotical behavior is due to the fact that the percentage of ground state optimal orbitals in the optimized orbital for the whole system increases as the number of non-interacting subsystems increases as has been discussed in Chapter 4. This conclusion can be demonstrated numerically from the Figure 6-8 that for various orbital optimization-CASCI methods, for example, SEAS-CASCI asymptotically get close to HF-CASCI, SA-CISSCF asymptotically get close to SS-CISSCF-CIS, which is just HF-CIS.

It is important to note here that CISNOr-*configuration interaction expansion* approach asymptotically get close to CISNO- *configuration interaction expansion* result. CIS relaxed first order reduced density matrix has non-zero occupied virtual (OV) block which is the only difference compared with CIS first order reduced density matrix. Figure 6-9 shows first order reduced density matrices being averaged for the whole system with one, two and three subsystems. If we only focus on one of the subsystems, it can be seen that as the non-interacting subsystems increase, the OV contribution to the SA-CIS density matrix reduces (the OV contribution equals OV/(N+1), where N is the number of non-interacting subsystems). And hence when the number of non-interaction subsystems increases to infinity, this OV contribution can be negligible. Thus, CISNOr-CASCI or CISNOr-CIS is asymptotically get close to CISNO-CASCI or CISNO-CIS (which is just HF-CIS) result.

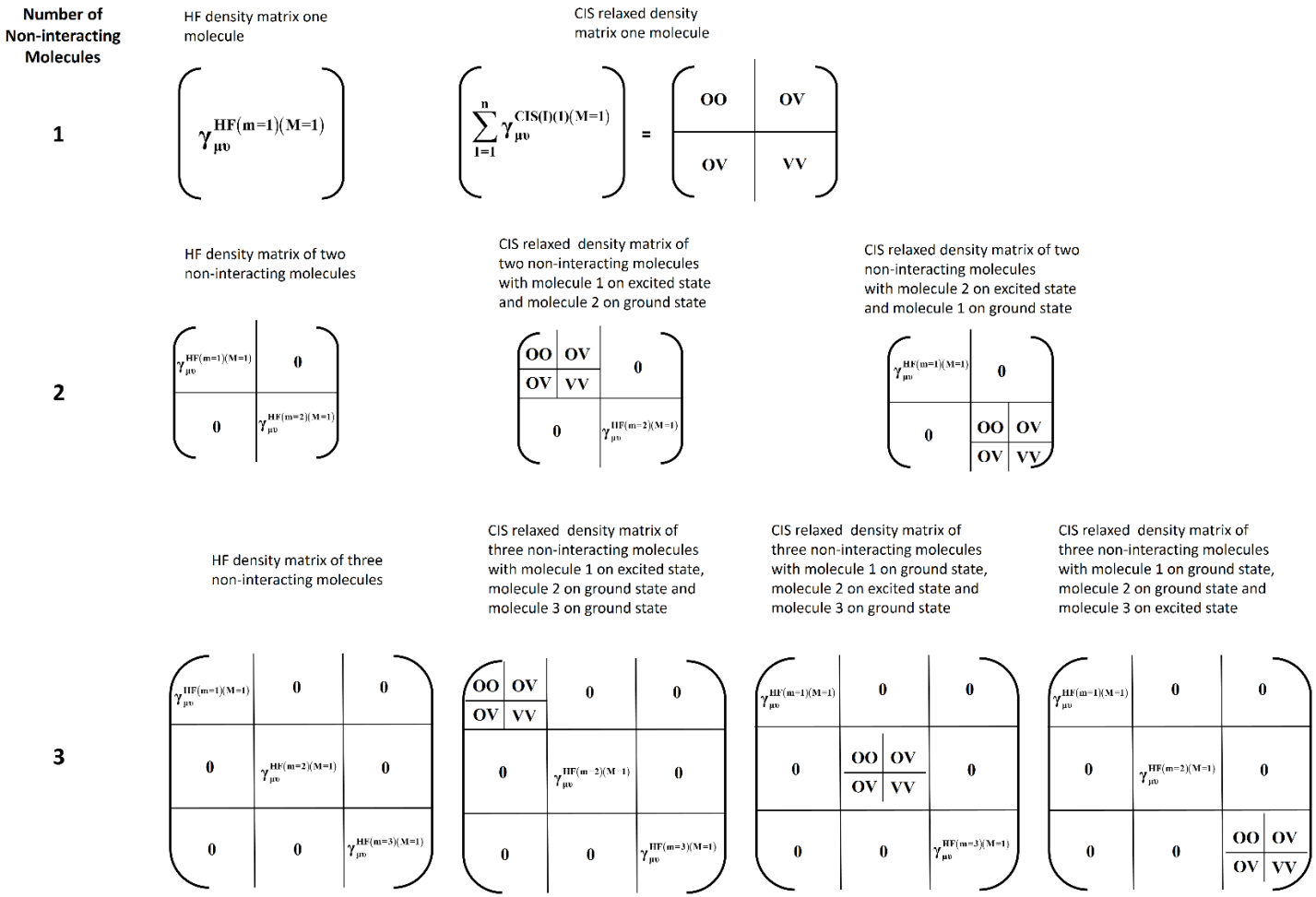


Figure 6-9. The HF ground state and CIS excited states relaxed first order reduced density matrices for one, two and three non-interacting subsystems.

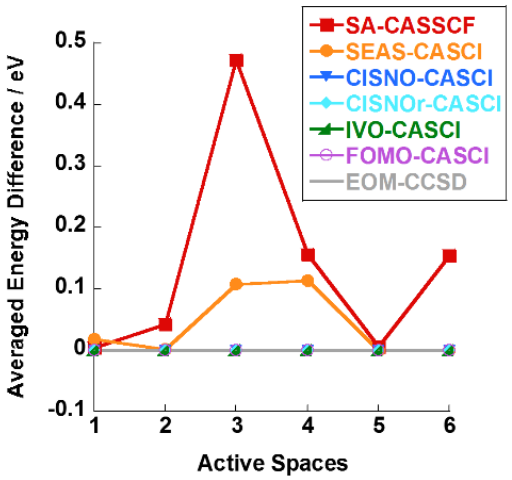
In Chapter 5, we have proposed the orbital optimization method-CIS approach to provide the method with capability in describing the PES in the vicinity of conical intersection geometry without explicitly introducing the doubly excited determinants into the wavefunction. We have also demonstrated the importance of employing virtual occupied coupling orbitals (VOCOs). And hence proposed the VOCO-CIS method to provide accurate description of PES near the conical intersection region (see Chapter 5 for detailed discussion). However, it is noteworthy here that the asymptotically size intensivity behavior for the two VOCO-CIS approaches we investigated (CISNOr-CIS and SA-CISSCF) as shown in Figure 6-8 suggests its failure when the number of non-interacting subsystems increases. The increased number of non-interaction subsystems dilutes the CISNOr or SA-CISSCF introduced coupling between closed shell reference determinant and singly excited determinants. This again, is caused by the fact that HF orbital percentage is increased when the number of non-interaction subsystems increases. Hence it would be interesting in the future to investigate the behavior of FOMO-CIS approach. FOMO-CASCI is size intensive and it is not surprising to expect that FOMO-CIS does not have the asymptotical behavior as SA-CISSCF or CISNOr-CIS. This suggests FOMO is possibly a promising VOCO can be employed in the VOCO-CIS in studying photophysical process in systems composed of multiple excitable subsystems.

6.3.2 Symmetry breaking

As has been extensively discussed in the section 6.1 as well as Chapter 3, symmetry breaking of the wavefunction causes unsmooth PES and is an undesirable property of

multi-reference electronic structure theory especially for AIMD simulations. Here we summarize this issue by employing linear titanium hydroxide cluster as an example. Figure 6-10(a) shows the averaged energy difference between degenerate states as a function of active spaces at SA-CASSCF, SEAS-CASCI, CISNO-CASCI, CISNOr-CASCI, IVO-CASCI and FOMO-CASCI theoretical levels with EOM-CCSD result shown for comparison. The averaged energy difference between degenerate states is defined as the average of the energy difference between S_1 and S_2, S_2 and S_3. Due to the degeneracy of orbitals are different for various optimization method, the active spaces employed for each multi-reference method are different as shown in Figure 6-10(b). The zero averaged energy difference for CISNO-CASCI, CISNOr-CASCI, FOMO-CASCI and IVO-CASCI for all active spaces indicate that these methods are capable to provide symmetry preserved wavefunction. SEAS-CASCI alleviates the symmetry breaking problem in SA-CASSCF.

(a)



(b)

Theoretical Level	Active Spaces					
	1	2	3	4	5	6
SA-CASSCF	6/5	6/8	12/8	12/11	16/10	16/13
SEAS-CASCI	6/5	6/8	12/8	12/11	16/10	16/13
CISNO-CASCI	6/5	6/8	10/7	10/10	16/10	16/13
CISNOr-CASCI	6/5	6/8	10/7	10/10	16/10	16/13
IVO-CASCI	6/9	6/12	12/12	12/15	16/14	
FOMO-CASCI	6/7	6/9	12/10	12/12	16/12	16/14

Figure 6-10. (a) The averaged energy difference between degenerate states as a function of active spaces for the linear titanium hydroxide cluster calculated at SA-CASSCF, SEAS-CASCI, CISNO-CASCI, CISNOr-CASCI, IVO-CASCI, FOMO-CASCI and EOM-CCSD theoretical levels. (b) The active spaces employed at each level of multi-reference theories.

6.3.3 Two-step methods in describing n->π* excitation

The n->π* excitation represents one type of important excitations in organic photochemistry, especially for molecules have, for example, carbonyl group, N=N double bond. And hence, an accurate description of the n->π* excitation is of great importance in investigating photophysical processes for molecules such as nucleobases, azo compound.[49-54] Similar groups or types of bonds exist in materials as well.[55-59] Due to the hole and particle orbitals are specifically described by one particle picture with n and π* characters, a balanced treatment of occupied and virtual orbitals are required to achieve accurate description. A balanced treatment means both occupied and virtual orbitals should be optimized beyond HF orbitals. Hence it is expected that orbitals are important in describing the n->π* excitation. Unbalanced treatment of occupied and virtual orbitals tend to loose quantitative accuracy, for example, in various two-step methods, orbitals that are re-optimized toward ground state with correlated wavefunction will end up with a better description for n orbital only; IVO re-optimizes the virtual orbitals while leaving the occupied orbitals in HF form and hence obtains a better description for π* orbitals only. Have seen the importance of orbitals in various two-step methods (with the configuration interaction expansion employed are CIS or CASCI) in current and previous chapters (Chapter 3-5), it remains to investigate the importance of the orbitals in describing the excited state with n->π* character as well.

The n->π* excited state excitation energies as a function of reasonable choices of active spaces for cytosine, formaldehyde, pyradizine, pyrazine, thymine and uracil

molecules as shown in Figure 6-11 (a) to (f) respectively. The results are computed at SA-CASSCF, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, FOMO-CASCI, COSNOr-CASCI and SA-active space (AS) CISSCF theoretical levels with EOM-CCSD results for comparison. The SA-ASCISSCF is a simplified version of SA-CISSCF with frozen core and frozen virtual approximations for CIS. This means we only consider the singly excited determinants in active space while leaving the orbitals outside the active space doubly occupied. We have employed SA3, SA2, SA4, SA4, SA3 and SA3 for cytosine, formaldehyde, pyradizine, pyrazine, thymine and uracil molecules at SA-CASSCF and SA-ASCISSCF theoretical levels. Table 6-1 shows the mean absolute error and mean signed error of S_1 excitation energy between two-step approaches and EOM-CCSD method for various two-step methods. The mean absolute error is computed by averaging the absolute error of S_1 excitation energy between the two-step approach with different active spaces and EOM-CCSD method for all the molecules in Figure 6-11. Among these two-step methods, except for IVO-CASCI whose occupied orbitals are left in HF orbital form, SA-CASSCF, CISNO-CASCI, SEAS-CASCI, FOMO-CASCI, COSNOr-CASCI and SA-ASCISSCF all re-optimize both the occupied and virtual orbitals. And hence it is not surprising from the Figure 6-11 that IVO-CASCI has the relative bad accuracy (with most of the computed excitation energies have more than 1 eV error compared with EOM-CCSD result) compared with other two-step methods for n->π* excited state if we trust EOM-CCSD results. Generally speaking, SEAS-CASCI, CISNO-CASCI, CISNOr-CASCI, SA-ASCISSCF and SA-CASSCF have relative better accuracy than FOMO-CASCI for these molecules. CISNOr-CASCI has better accuracy than

CISNO-CASCI which is expected as both occupied and virtual orbitals are re-optimized by employing the relaxed density matrices. By comparing the mean absolute error and mean signed error, IVO-CASCI, FOMO-CASCI and CISNO-CASCI are tend to overestimate the excitation energy of the n->π* excited state because all these methods show same mean absolute error and mean signed error.

It is also noteworthy here "periodic" behavior of the excitation energies for IVO-CASCO, FOMO-CASCI, CISNO-CASCI and CISNOr-CASCI as a function of active space. This is due to the fact that the orbitals are independent of the active spaces employed in CI expansion. This behavior is contrast to SA-CASSCF at which level the effect of the active space is unpredictable. Figure 6-12 have listed the corresponding n and π* orbitals associated with the states we calculated as shown in Figure 6-11 for these molecules.

The above conclusions almost holds for all difference active spaces. Thus it is reasonable to attribute such behavior to the orbitals. Hence again, we have shown the importance of the orbitals in the two-step methods.

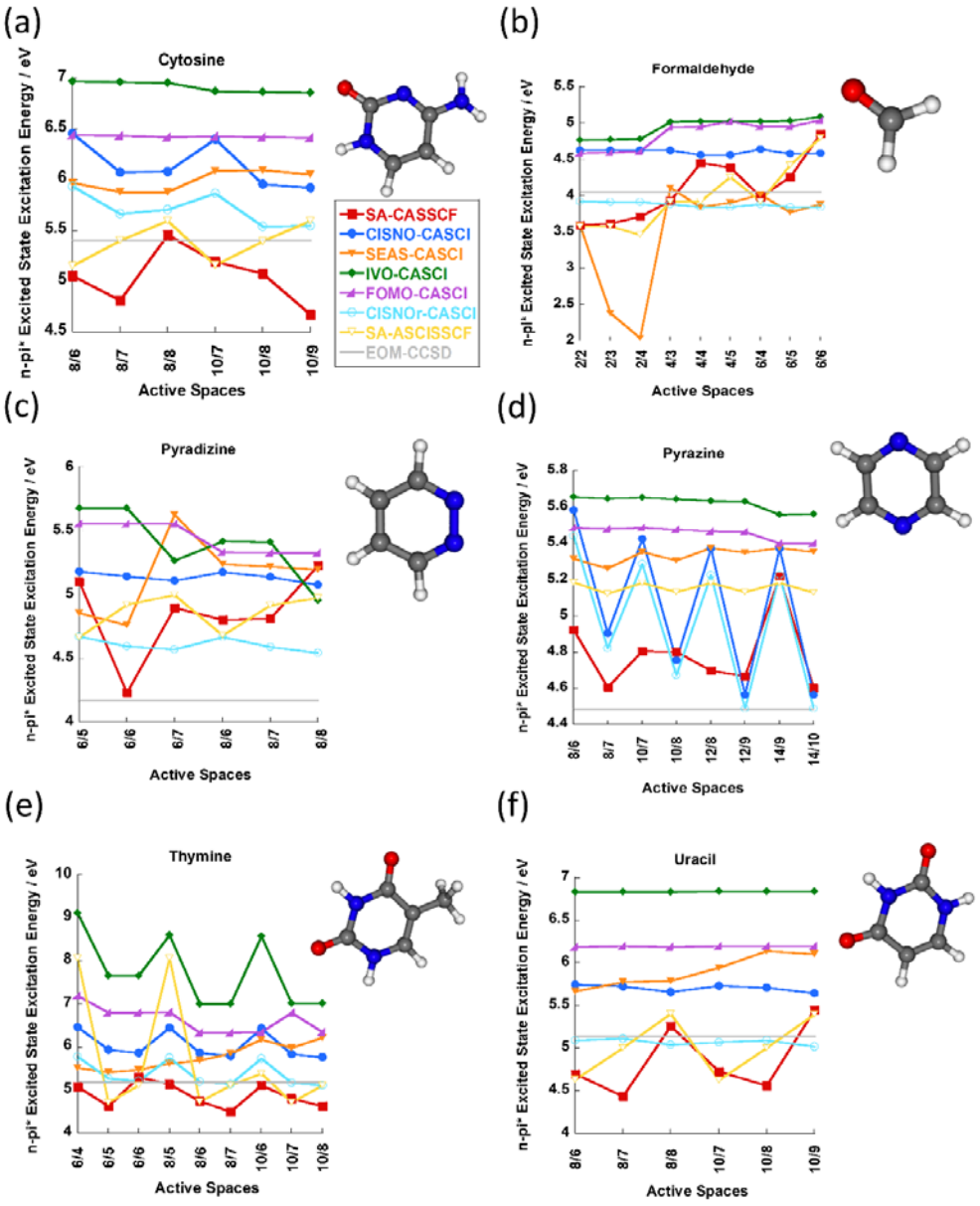


Figure 6-11. The excitation energies of n->π* state as a function of active spaces employed in SA-CASSCF and various two-step methods. The SA-CASSCF, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, FOMO-CASCI, CISNO-CASCI, SA-ASCISSCF and EOM-CCSD are shown as red, blue, orange, green, purple, cyan, yellow and grey respectively for (a) cytosine, (b) formaldehyde, (c) pyradizine, (d) pyrazine, (e) thymine and (f) uracil molecules.

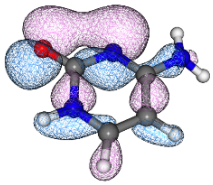
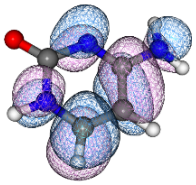
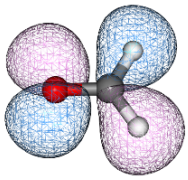
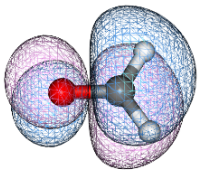
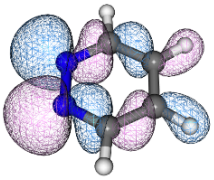
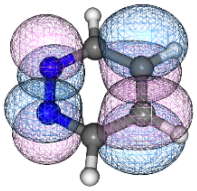
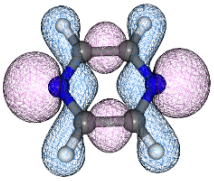
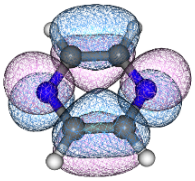
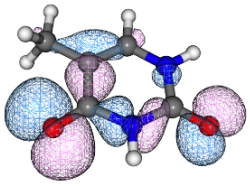
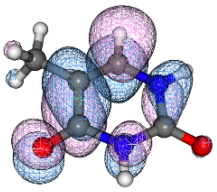
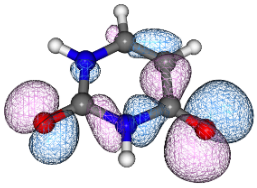
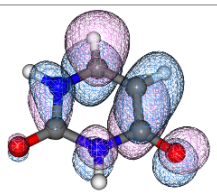
Molecules	n orbital	π* orbital
Cytosine		
Formaldehyde		
Pyradizine		
Pyrazine		
Thymine		
Uracil		

Figure 6-12. The n->π* excited state characters for (a) cytosine, (b) formaldehyde, (c) pyradizine, (d) pyrazine, (e) thymine and (f) uracil molecules. The n and π* orbitals shown here are HF orbitals.

6.3.4 Quantitative accuracy in description of the conical intersection

The various two-step method we have discussed in this chapter are all suitable in investigating PESs in the vicinity of conical intersection geometries between ground and excited states in a qualitative sense, as either we employ CASCI expansion after the orbital optimization, or we employ CIS expansion while utilize VOCOs. And hence, both approaches are capable of lifting the degeneracy between ground and excited states once the nuclear geometry distorts from the conical intersection geometry. However, as we have seen from Chapter 5 that orbitals without virtual occupied coupling are tend to fail in quantitative perspective and hence we introduced the VOCO idea. I refer the readers Chapter 5 for a detailed discussion of the arising of the problem. The quantitative accuracy of the two-step method in the vicinity of conical intersection are assessed by computing the S_1-S_0 energy gaps at trustworthy conical intersection geometries. The trustworthy conical intersection geometry is defined such that high level of theory for example, MRCI(Q) predicts small S_1-S_0 energy gap. And hence, if a two-step method predicts a relative large S_1-S_0 energy gap for example, larger than 1 eV, at the trustworthy conical intersection geometry, we believe that this two-step method fails quantitatively in computing the PES near conical intersection.

Table 6-1. The mean absolute error and mean signed error of S_1 excitation energy between two-step approaches and EOM-CCSD method. The S_1 state has n->π* excitation character.

Method	Mean Absolute Error / eV	Mean Signed Error / eV
SA-CASSCF	0.397	0.004
CISNO-CASCI	0.704	0.704
SEAS-CASCI	0.708	0.480
IVO-CASCI	1.518	1.518
FOMO-CASCI	1.094	1.094
CISNOr-CASCI	0.279	0.180
SA-ASCISSCF	0.524	0.279

Here, I will just show some calculations based on FOMO and IVO to further demonstrate the importance of VOCO. It should be expected that FOMO is accurate in that sense, while IVO may have severe quantitative disagreement with higher level of theory for example, MRCI(Q). At FOMO-CASCI optimized pyramidalized ethylene conical intersection geometry, VOCO-configuration interaction methods, for example, SA-CASSCF, CISNOr-CASCI, SEAS-CASCI predict 0.121, 0.392 and 0.118 eV energy gap between S_1 and S_0 states. While those two-step methods employs orbitals without virtual occupied coupling, for example, CISNO-CASCI, HF-CASCI, IVO-CASCI predict 1.741, 1.852 and 1.791 eV S_1 and S_0 energy gaps. Higher level theory with dynamical correlations like MRCI(Q)

predicts 0.175 eV energy gap between S_1 and S_0 states. This further validates the hypothesis that virtual occupied coupling is important in describing conical intersection regions of PESs with quantitative accuracy. And hence, we can conclude here that a suitable two-step method in investigating non-radiative decay processes, should employ the VOCO-CI expansion approach.

6.4 Conclusions and Outlook

The two-step methods employ the computational scheme as orbital optimization followed by configuration interaction expansion is a very promising set of methods to be widely used as alternatives to SA-CASSCF method. As the two-step methods decouple the optimization for orbital coefficients and CI vectors, some undesirable properties of SA-CASSCF can be removed, for example, wavefunction symmetry breaking, size intensivity, discontinuous of the PES, unpredictable effect of the active space. In addition, we have shown the importance of the orbitals employed in the two-step methods despite the fact that final wavefunction is expanded by multiple configurations. With that, it is very important to recognize the fact that we can take advantage of the different properties of different orbitals and utilize different orbitals in studying various chemical problems. For example, VOCOs are suitable for investigating non-radiative recombination processes; CISNO and CISNOr can be made black box by choosing the active space according to the occupation numbers of natural orbitals; IVO and FOMO requires the computational cost as close as HF orbitals and can be made very efficient; FOMO and CISNO are size intensive

and hence will be suitable for investigating systems composed of multiple excitable subsystems.

With the advances in both theoretical methods and computer hardware, we are now able to simulate very big systems (for example, nano-sized cluster, small proteins) at multi-reference theoretical levels. And we believe with the gradients available for various two-step methods, we are able to directly simulate the excited and ground states dynamics for nano particles, proteins, etc. and to investigate more interesting large-scale problems that cannot be simulated directly in the past.

APPENDIX

APPENDIX

Table 6-S1. The absolute energies of S_0 and S_1 states of $Ti(OH)_4$ FC geometry at SEAS-RASCI, CISNOr-RASCI and CISNO-RASCI theoretical levels with various EFOs, corresponding to Figure 6-3(a), Figure 6-4(a) and Figure 6-6(a).

SEAS-RASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
EFO 22 23	-1150.272535	-1150.098104
EFO 24 25 26	-1150.301335	-1150.096659
EFO 27 28 29	-1150.324261	-1150.054309
EFO 30 31	-1150.268445	-1150.062777
EFO 32 33 34	-1150.302357	-1150.095487
CASCI	-1150.334383	-1150.111099
CISNOr-RASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
EFO 22 23 24	-1150.324156	-1150.104537
EFO 25 26	-1150.306679	-1150.103264
EFO 27 28 29	-1150.326213	-1150.032947
EFO 30 31	-1150.296818	-1150.003361
EFO 32 33 34	-1150.313888	-1150.088394
CASCI	-1150.352712	-1150.132994
CISNO-RASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
EFO 22 23 24	-1150.337049	-1150.056624
EFO 25 26	-1150.337454	-1150.053617
EFO 27 28 29	-1150.344104	-1150.023972
EFO 30 31	-1150.325909	-1149.987162
EFO 32 33 34	-1150.327238	-1150.038442
CASCI	-1150.361508	-1150.092498

Table 6-S2. The absolute energies of S_0 and S_1 states of $Ni(CO)_4$ FC geometry at SEAS-RASCI, CISNOr-RASCI and CISNO-RASCI theoretical levels with various EFOs, corresponding to Figure 6-3(b), Figure 6-4(b) and Figure 6-6(b).

HF-RASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
EFO 35 36 37	-1957.567536	-1957.350275
EFO 38 39	-1957.517133	-1957.320873
EFO 40 41 42	-1957.485564	-1957.259883
EFO 43 44 45	-1957.484585	-1957.289521
EFO 46 47 48	-1957.511297	-1957.312295
EFO 49 50 51	-1957.568749	-1957.351229
CISNOr-RASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
EFO 36 37 38	-1957.579934	-1957.411618
EFO 39	-1957.591232	-1957.421746
EFO 40 41 42	-1957.417475	-1956.856816
EFO 43 44 45	-1957.430958	-1957.208789
EFO 46 47 48	-1957.586930	-1957.414982
EFO 49 50 51	-1957.570704	-1957.39457
CISNO-RASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
EFO 36 37 38	-1957.551417	-1957.3763180
EFO 39	-1957.581470	-1957.406600
EFO 40 41 42	-1957.450314	-1957.034910
EFO 43 44 45	-1957.44816	-1957.192155
EFO 46 47 48	-1957.575529	-1957.396870
EFO 49 50 51	-1957.561588	-1957.379101

Table 6-S3. The absolute energies of S_0 and S_1 states of $Ti(OH)_4$ FC geometry at CISNOr-CASCI, SEAS-CASCI, HF-CASCI theoretical levels with various active spaces, corresponding to Figure 6-5(a).

CISNOr-CASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
6/5	-1150.298440	-1150.143875
6/8	-1150.350862	-1150.176923
10/7	-1150.308613	-1150.152104
10/10	-1150.366600	-1150.198190
16/10	-1150.316579	-1150.155559
16/13	-1150.444635	-1150.253098
SEAS-CASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
6/5	-1150.265357	-1150.113239
6/8	-1150.274650	-1150.119019
12/8	-1150.271273	-1150.097100
12/11	-1150.308413	-1150.111115
16/10	-1150.305740	-1150.095031
16/13	-1150.334383	-1150.111099
HF-CASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
6/5	-1150.316475	-1149.996325
6/8	-1150.320821	-1150.001601
12/8	-1150.319013	-1150.007031
12/11	-1150.329758	-1150.04102
16/10	-1150.326518	-1150.038124
16/13	-1150.348185	-1150.075672

Table 6-S4. The absolute energies of S_0 and S_1 states of $Ni(CO)_4$ FC geometry at CISNOr-CASCI, SEAS-CASCI, HF-CASCI theoretical levels with various active spaces, corresponding to Figure 6-5(b).

CISNOr-CASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
6/6	-1957.539593	-1957.355356
6/9	-1957.541126	-1957.358207
6/12	-1957.550226	-1957.374391
8/7	-1957.539862	-1957.356017
8/10	-1957.541914	-1957.359571
8/13	-1957.551417	-1957.376318
14/10	-1957.554473	-1957.369365
14/13	-1957.561588	-1957.379101
SEAS-CASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
6/6	-1957.455700	-1957.340213
6/9	-1957.548837	-1957.381824
6/12	-1957.552390	-1957.387106
10/8	-1957.547891	-1957.384779
10/11	-1957.619672	-1957.619672
10/14	-1957.625326	-1957.436275
16/11	-1957.542706	-1957.379469
16/14	-1957.629305	-1957.438239

Table 6-S4(cont'd).

HF-CASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
6/6	-1957.475895	-1957.264021
6/9	-1957.496068	-1957.297165
6/12	-1957.505724	-1957.321648
10/8	-1957.494101	-1957.293858
10/11	-1957.563166	-1957.344295
10/14	-1957.567536	-1957.350275
16/11	-1957.495253	-1957.295499
16/14	-1957.568749	-1957.351229

Table 6-S5. The absolute energies of S_0 and S_1 states of $Ti(OH)_4$ and $Ni(CO)_4$ FC geometries at EOM-CCSD theoretical level, corresponding to Figure 6-5.

EOM-CCSD	S_0 Energy / a.u.	S_1 Energy / a.u.
$Ti(OH)_4$	-1151.189298	-1151.189298
$Ni(CO)_4$	-1959.029496	-1958.856468

Table 6-S6. The absolute energies of S_0 and S_1 states of pyrazine FC geometries at SA-CASSCF, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, FOMO-CASCI, CISNOr-CASCI, SA-ASCISSCF with various active spaces and EOM-CCSD theoretical levels, corresponding to Figure 6-11.

Methods		S_0 Energy / a.u.	S_1 Energy / a.u.
SA-CASSCF	8/6	-262.7209091	-262.5400132
	8/7	-262.7356877	-262.5664416
	10/7	-262.7334334	-262.5568366
	10/8	-262.7685735	-262.5920637
	12/8	-262.7337632	-262.5610887
	12/9	-262.7692072	-262.5977478
	14/9	-262.7324069	-262.5406699
	14/10	-262.7689082	-262.5997330
CISNO-CASCI	8/6	-262.7282940	-262.5230312
	8/7	-262.7316129	-262.5514319
	10/7	-262.7283945	-262.5289539
	10/8	-262.7317711	-262.5570789
	12/8	-262.7284626	-262.5309721
	12/9	-262.7321854	-262.5643726
	14/9	-262.7285317	-262.5310347
	14/10	-262.7322914	-262.5644780
SEAS-CASCI	8/6	-262.7228971	-262.5275781
	8/7	-262.7208774	-262.5275138
	10/7	-262.7243481	-262.5275470
	10/8	-262.7227901	-262.5277697
	12/8	-262.7263009	-262.5287719
	12/9	-262.728966	-262.5323811
	14/9	-262.7297626	-262.5322113
	14/10	-262.7302645	-262.5335781

Table 6-S6 (cont'd).

Methods		S_0 Energy / a.u.	S_1 Energy / a.u.
IVO-CASCI	8/6	-262.7282729	-262.5203338
	8/7	-262.7293864	-262.5217692
	10/7	-262.7283329	-262.5204813
	10/8	-262.7295502	-262.5220727
	12/8	-262.7308740	-262.5238114
	12/9	-262.7325919	-262.5255747
	10/8	-262.7256564	-262.5243522
	12/8	-262.7280203	-262.5271370
	12/9	-262.7292558	-262.5284369
	14/9	-262.7280464	-262.5295692
	14/10	-262.7295870	-262.5311265
CISNOr-CASCI	8/6	-262.7276026	-262.5275477
	8/7	-262.7309831	-262.5537685
	10/7	-262.7277061	-262.5333558
	10/8	-262.7311432	-262.5594258
	12/8	-262.7277733	-262.5356980
	12/9	-262.7315529	-262.5666071
	14/9	-262.7278444	-262.5357648
	14/10	-262.7316605	-262.5667141
ASCISSCF	8/6	-262.6944009	-262.5038247
	8/7	-262.6849529	-262.4967398
	10/7	-262.6944044	-262.5039292
	10/8	-262.6852898	-262.4967407
	12/8	-262.6944019	-262.5039181
	12/9	-262.6853117	-262.4967291
	14/9	-262.6944418	-262.5039479
	14/10	-262.6853097	-262.4968576
E0M-CCSD		-263.5497345	-263.3849097

Table 6-S7. The absolute energies of S_0 and S_1 states of thymine FC geometries at SA-CASSCF, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, FOMO-CASCI, CISNOr-CASCI, SA-ASCISSCF with various active spaces and EOM-CCSD theoretical levels, corresponding to Figure 6-11.

Methods		S_0 Energy / a.u.	S_1 Energy / a.u.
SA-CASSCF	6/4	-451.5215611	-451.2839487
	6/5	-451.530531	-451.3124912
	6/6	-451.5345491	-451.3189209
	8/5	-451.5216404	-451.2849189
	8/6	-451.5306307	-451.315164
	8/7	-451.5346633	-451.3218278
	10/6	-451.5216787	-451.2851515
	10/7	-451.5307238	-451.3162722
	10/8	-451.534787	-451.3230835
CISNO-CASCI	6/4	-451.5264652	-451.3396022
	6/5	-451.552176	-451.3821683
	6/6	-451.5795386	-451.3849354
	8/5	-451.527933	-451.339283
	8/6	-451.5562364	-451.3816471
	8/7	-451.5758671	-451.4107987
	10/6	-451.5281649	-451.3406274
	10/7	-451.557579	-451.3811382
	10/8	-451.5809969	-451.4111028
SEAS-CASCI	6/4	-451.5111191	-451.3084006
	6/5	-451.523535	-451.3245938
	6/6	-451.5214613	-451.3203662
	8/5	-451.493249	-451.2866156
	8/6	-451.5173038	-451.308073
	8/7	-451.5175396	-451.3026339
	10/6	-451.5030628	-451.2761715
	10/7	-451.5278573	-451.3079573

Table 6-S7 (cont'd).

Methods	S0 Energy / a.u.	S1 Energy / a.u.	Methods
IVO-CASCI	6/4	-451.5265055	-451.2317762
	6/5	-451.5308964	-451.2496468
	6/6	-451.531066	-451.2497871
	8/5	-451.5265134	-451.2317912
	8/6	-451.5309252	-451.2737446
	8/7	-451.5311662	-451.2739395
	10/6	-451.5302941	-451.2385149
	10/7	-451.537169	-451.2792897
	10/8	-451.5374329	-451.2795003
FOMO-CASCI	6/4	-451.5202696	-451.2560791
	6/5	-451.5252683	-451.2756912
	6/6	-451.5254123	-451.2757637
	8/5	-451.5202824	-451.2701668
	8/6	-451.5253126	-451.2925064
	8/7	-451.525506	-451.2926232
	10/6	-451.5311277	-451.2977894
	10/7	-451.5221634	-451.2724373
	10/8	-451.5313578	-451.2979338
CISNOr-CASCI	6/4	-451.5163735	-451.304072
	6/5	-451.5255445	-451.3317088
	6/6	-451.529985	-451.3384182
	8/5	-451.5164493	-451.3051669
	8/6	-451.5256411	-451.3345239
	8/7	-451.5301015	-451.3414091
	10/6	-451.5164802	-451.3055047
	10/7	-451.5257418	-451.3357642
	10/8	-451.5302476	-451.3427705

Table 6-S7 (cont'd).

Methods	S_0 Energy / a.u.	S_1 Energy / a.u.	Methods
	6/4	-451.5163735	-451.304072
	6/5	-451.5255445	-451.3317088
	6/6	-451.529985	-451.3384182
	8/5	-451.5164493	-451.3051669
ASCISSCF	8/6	-451.5256411	-451.3345239
	8/7	-451.5301015	-451.3414091
	10/6	-451.5164802	-451.3055047
	10/7	-451.5257418	-451.3357642
	10/8	-451.5302476	-451.3427705
EOM-CCSD		-452.8775950	-452.6871081

Table 6-S8. The absolute energies of S_0 and S_1 states of formaldehyde FC geometries at SA-CASSCF, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, FOMO-CASCI, CISNOr-CASCI, SA-ASCISSCF with various active spaces and EOM-CCSD theoretical levels, corresponding to Figure 6-11.

Methods		S_0 Energy / a.u.	S_1 Energy / a.u.
SA-CASSCF	2/2	-113.8670439	-113.6971621
	2/3	-113.8671264	-113.6971649
	2/4	-113.8671407	-113.6971649
	4/3	-113.8671874	-113.697221
	4/4	-113.8674604	-113.6997846
	4/5	-113.8685113	-113.70074
	6/4	-113.86772	-113.697221
	6/5	-113.8682035	-113.7000008
	6/6	-113.87058	-113.7019616
CISNO-CASCI	2/2	-113.8559184	-113.7239353
	2/3	-113.8653663	-113.732407
	2/4	-113.8687064	-113.7323799
	4/3	-113.8969344	-113.7524507
	4/4	-113.919	-113.7553065
	4/5	-113.9205384	-113.7592202
	6/4	-113.8988042	-113.7524341
	6/5	-113.9329692	-113.776543
	6/6	-113.9643381	-113.7859371
SEAS-CASCI	2/2	-113.8558361	-113.7240166
	2/3	-113.8256424	-113.7379098
	2/4	-113.8149907	-113.7399989
	4/3	-113.8923118	-113.7413167
	4/4	-113.885986	-113.7447326
	4/5	-113.8879742	-113.7444445
	6/4	-113.8937111	-113.7461644
	6/5	-113.8858865	-113.7470781
	6/6	-113.8886104	-113.7459974

Table 6-S8 (cont'd).

Methods		S_0 Energy / a.u.	S_1 Energy / a.u.
IVO-CASCI	2/2	-113.8670742	-113.6918442
	2/3	-113.8674249	-113.6918442
	2/4	-113.8675705	-113.6918442
	4/3	-113.8965806	-113.7121844
	4/4	-113.8969578	-113.7122657
	4/5	-113.8973812	-113.7126451
	6/4	-113.8969638	-113.7121844
	6/5	-113.897689	-113.7127687
	6/6	-113.9020862	-113.7151487
FOMO-CASCI	2/2	-113.864169	-113.6955111
	2/3	-113.8644587	-113.6955111
	2/4	-113.8655233	-113.6959925
	4/3	-113.8912475	-113.7095498
	4/4	-113.8915422	-113.7095943
	4/5	-113.8948011	-113.7100556
	6/4	-113.8915751	-113.7095498
	6/5	-113.8920188	-113.7100085
	6/6	-113.8958312	-113.7107292
CISNOr-CASCI	2/2	-113.8606065	-113.7164309
	2/3	-113.8606537	-113.7167159
	2/4	-113.860668	-113.7167159
	4/3	-113.8607583	-113.7178955
	4/4	-113.8609519	-113.7197519
	4/5	-113.8619907	-113.7207338
	6/4	-113.8607633	-113.7178955
	6/5	-113.8609632	-113.7197581
	6/6	-113.8620242	-113.7207544

Table 6-S8 (cont'd).

Methods		S_0 Energy / a.u.	S_1 Energy / a.u.
ASCISSCF	2/2	-113.8555451	-113.7240166
	2/3	-113.8555451	-113.7240166
	2/4	-113.8596193	-113.7324377
	4/3	-113.8955247	-113.7514851
	4/4	-113.8955247	-113.7514851
	4/5	-113.9089917	-113.7523549
	6/4	-113.8955247	-113.7514851
	6/5	-113.9130432	-113.7505697
	6/6	-113.9270363	-113.7508137
EOM-CCSD		-114.1935923	-114.0445844

Table 6-S9. The absolute energies of S_0 and S_1 states of pyradizine FC geometries at SA-CASSCF, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, FOMO-CASCI, CISNOr-CASCI, SA-ASCISSCF with various active spaces and EOM-CCSD theoretical levels, corresponding to Figure 6-11.

Methods		S_0 Energy / a.u.	S_1 Energy / a.u.
SA-CASSCF	6/5	-262.6530373	-262.4626575
	6/6	-262.6536243	-262.4646775
	6/7	-262.65383	-262.4661087
	8/6	-262.6532277	-262.4630421
	8/7	-262.6538903	-262.4650753
	8/8	-262.6542286	-262.4676117
CISNO-CASCI	6/5	-262.6938711	-262.506474
	6/6	-262.686831	-262.5314565
	6/7	-262.7124122	-262.5324652
	8/6	-262.6990436	-262.522592
	8/7	-262.7111658	-262.5343698
	8/8	-262.7363279	-262.5442182
SEAS-CASCI	6/5	-262.6617543	-262.483384
	6/6	-262.6575174	-262.4826602
	6/7	-262.6873179	-262.4805432
	8/6	-262.6911656	-262.4987358
	8/7	-262.6890224	-262.497285
	8/8	-262.6875959	-262.4967674

Table 6-S9(cont'd).

Methods		S_0 Energy / a.u.	S_1 Energy / a.u.
IVO-CASCI	6/5	-262.6917466	-262.4830758
	6/6	-262.6918617	-262.483131
	6/7	-262.6965408	-262.5030992
	8/6	-262.6920001	-262.4928643
	8/7	-262.692238	-262.4932435
	8/8	-262.697258	-262.5152229
FOMO-CASCI	6/5	-262.6874972	-262.4834326
	6/6	-262.6875658	-262.4834701
	6/7	-262.687631	-262.4835244
	8/6	-262.6877296	-262.4918642
	8/7	-262.6879218	-262.4921783
	8/8	-262.6881292	-262.4923598
CISNOr-CASCI	6/5	-262.6457184	-262.4741465
	6/6	-262.6462148	-262.4774236
	6/7	-262.6462728	-262.4784376
	8/6	-262.645848	-262.4744693
	8/7	-262.6463927	-262.4777551
	8/8	-262.6466302	-262.4797128
ASCISSCF	6/5	-262.660426	-262.4890159
	6/6	-262.6708355	-262.489994
	6/7	-262.674484	-262.4909455
	8/6	-262.6610604	-262.4891869
	8/7	-262.6710517	-262.4903855
	8/8	-262.6741901	-262.4914878
EOM-CCSD		-263.5197005	-263.3665334

Table 6-S10. The absolute energies of S_0 and S_1 states of cytosine FC geometries at SA-CASSCF, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, FOMO-CASCI, CISNOr-CASCI, SA-ASCISSCF with various active spaces and EOM-CCSD theoretical levels, corresponding to Figure 6-11.

Methods		S_0 Energy / a.u.	S_1 Energy / a.u.
SA-CASSCF	8/6	-392.6423654	-392.4049004
	8/7	-392.6507473	-392.4275432
	8/8	-392.6565001	-392.4329341
	10/7	-392.6424503	-392.4070773
	10/8	-392.6509145	-392.4318938
	10/9	-392.6574757	-392.4396457
CISNO-CASCI	8/6	-392.6528814	-392.4670421
	8/7	-392.6705067	-392.4935933
	8/8	-392.6974048	-392.5000248
	10/7	-392.6494881	-392.4585069
	10/8	-392.6795237	-392.5017429
	10/9	-392.6992297	-392.5274112
SEAS-CASCI	8/6	-392.6249305	-392.4055178
	8/7	-392.6231017	-392.4069153
	8/8	-392.6243889	-392.4082913
	10/7	-392.6392835	-392.4269001
	10/8	-392.639291	-392.4272229
	10/9	-392.6378641	-392.4237122
IVO-CASCI	8/6	-392.6374235	-392.3879142
	8/7	-392.6312065	-392.4234318
	8/8	-392.6381187	-392.38945
	10/7	-392.6462286	-392.4163686
	10/8	-392.6464013	-392.4166412
	10/9	-392.6471489	-392.4180256

Table 6-S10 (cont'd).

Methods		S_0 Energy / a.u.	S_1 Energy / a.u.
FOMO-CASCI	8/6	-392.6321711	-392.3953669
	8/7	-392.6323333	-392.3957257
	8/8	-392.632847	-392.3966454
	10/7	-392.640249	-392.4154732
	10/8	-392.640522	-392.4160498
	10/9	-392.6412151	-392.4171797
CISNOr-CASCI	8/6	-392.6397332	-392.4214132
	8/7	-392.648458	-392.4400901
	8/8	-392.6544075	-392.4446567
	10/7	-392.6398166	-392.4242624
	10/8	-392.6486231	-392.4449565
	10/9	-392.6552368	-392.4514268
ASCISSCF	8/6	-392.6245605	-392.4351208
	8/7	-392.648458	-392.4387928
	8/8	-392.646016	-392.4403894
	10/7	-392.6247448	-392.435182
	10/8	-392.6373733	-392.4391345
	10/9	-392.6462236	-392.4404841
EOM-CCSD		-393.827024	-393.6421338

Table 6-S11. The absolute energies of S_0 and S_1 states of uracil FC geometries at SA-CASSCF, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, FOMO-CASCI, CISNOr-CASCI, SA-ASCISSCF with various active spaces and EOM-CCSD theoretical levels, corresponding to Figure 6-11.

Methods		S_0 Energy / a.u.	S_1 Energy / a.u.
SA-CASSCF	8/6	-412.4881997	-412.2768936
	8/7	-412.4959065	-412.2854088
	8/8	-412.502471	-412.2944362
	10/7	-412.4883746	-412.2776514
	10/8	-412.4972389	-412.2873157
	10/9	-412.5044712	-412.297078
CISNO-CASCI	8/6	-412.5140084	-412.3412678
	8/7	-412.5335837	-412.3704284
	8/8	-412.5607781	-412.3674963
	10/7	-412.5152197	-412.3413252
	10/8	-412.5386258	-412.3709668
	10/9	-412.5661239	-412.3658852
SEAS-CASCI	8/6	-412.4756963	-412.2673633
	8/7	-412.4756874	-412.263352
	8/8	-412.4770218	-412.2641101
	10/7	-412.4867103	-412.2681529
	10/8	-412.4884066	-412.2625882
	10/9	-412.4879428	-412.2635967
FOMO-CASCI	8/6	-412.4888498	-412.2377047
	8/7	-412.4890443	-412.2378657
	8/8	-412.4894321	-412.2383052
	10/7	-412.4971207	-412.2457243
	10/8	-412.4973332	-412.2458972
	10/9	-412.4979051	-412.2465166

Table 6-S11 (cont'd). The absolute energies of S_0 and S_1 states of uracil FC geometries at SA-CASSCF, CISNO-CASCI, SEAS-CASCI, IVO-CASCI, FOMO-CASCI, CISNOr-CASCI, SA-ASCISSCF with various active spaces and EOM-CCSD theoretical levels, corresponding to Figure 6-11.

Methods		S_0 Energy / a.u.	S_1 Energy / a.u.
IVO-CASCI	8/6	-412.4837685	-412.2561153
	8/7	-412.4839486	-412.2562365
	8/8	-412.4842981	-412.256605
	10/7	-412.4906426	-412.2627959
	10/8	-412.4908532	-412.2629395
	10/9	-412.491251	-412.2633489
CISNOr-CASCI	8/6	-412.4831505	-412.2960963
	8/7	-412.4926792	-412.3047568
	8/8	-412.4979092	-412.3126552
	10/7	-412.4833412	-412.2969419
	10/8	-412.4942981	-412.3070214
	10/9	-412.4997229	-412.3151807
ASCISSCF	8/6	-412.4769792	-412.3065781
	8/7	-412.4939685	-412.3099937
	8/8	-412.5101495	-412.311422
	10/7	-412.4770585	-412.3069055
	10/8	-412.4941933	-412.3102834
	10/9	-412.5097224	-412.3113527
EOM-CCSD		-413.6786402	-413.6786402

Optimized Geometries (In Angstrom) corresponding to Figure 6-3 to Figure 6-6

Optimized $Ti(OH)_4$ FC geometry (MP2/6-31G)**

Ti	0.000028029	-4.9761E-060	-3.11318E-05
O	0.579850493	1.642100915	0.343054576
O	0.932322302	-1.168589833	0.956868520
O	0.206103850	-0.351365992	-1.727614069
O	-1.718271976	-0.122164398	0.427645664
H	0.890857688	2.524711763	0.528550454
H	1.432192981	-1.797178175	1.471787659
H	0.317344068	-0.540165405	-2.65611480

Optimized $Ti(OH)_4$ FC geometry (KS/B3LYP/6-31G)**

Ni	0.001942512	0.004982954	0.000814315
C	0.000793323	0.004146248	1.814466774
C	1.712407995	0.006663590	-0.602585037
C	-0.852111545	-1.476760186	-0.602560243
C	-0.854400697	1.484797317	-0.604441955
O	-1.393015187	-2.413734100	-0.981924589
O	2.793934055	0.007582901	-0.982991287
O	4.97172E-05	0.002687739	2.960913291

Optimized Geometries (In Angstrom) corresponding to Figure 6-11

Optimized pyrazine FC geometry (MP2/6-31G)**

C	0.043212	-0.111863	0.030960
C	1.178081	-0.158657	1.998825
C	2.321112	-0.600233	1.329034
C	1.186051	-0.553557	-0.638847
N	0.021661	0.092331	1.359971
N	2.342440	-0.804539	0.000080
H	1.187513	-0.001787	3.071559
H	-0.877047	0.083135	-0.508064
H	1.176552	-0.710913	-1.711343
H	3.241428	-0.795320	1.867838

Optimized thymine FC geometry (MP2/6-31G)**

C	0.036270331	-0.001873237	-0.012052387
C	-0.002086435	-0.000929483	1.450786869
N	1.261011984	0.000237037	2.061903043
C	2.511912768	0.000685839	1.463308462
C	1.244489942	-0.001481990	-0.625817911
N	2.430800781	-0.000225806	0.079675216
O	3.570956188	0.001751233	2.079927968
O	-1.029199485	-0.001210169	2.128974622
C	-1.269925059	-0.003111507	-0.741773855
H	-1.114104897	-0.007852373	-1.819856236
H	-1.859272110	-0.877618944	-0.467517768
H	-1.857153955	0.875172479	-0.475013868
H	1.347597555	-0.002251441	-1.703316807
H	3.318772926	0.000005622	-0.398053759
H	1.264764465	0.000783740	3.074251409

Optimized formaldehyde FC geometry (MP2/6-31G)**

C	0.000000000	0.000000000	0.065813147
O	0.000000000	0.000000000	1.287044067
H	0.931369973	0.000000000	-0.520928607
H	-0.931369973	0.000000000	-0.520928607

Optimized pyradizine FC geometry (MP2/6-31G)**

N	0.076995414	0.058042985	0.034910149
N	0.083642799	0.04748119	1.382637956
C	1.262768479	0.022541485	2.028697160
C	1.249631623	0.042925519	-0.623055230
C	2.506041666	0.006225818	1.389897335
C	2.499168281	0.016128170	0.003105043
H	1.183251341	0.015079099	3.109177215
H	3.424779825	-0.014369227	1.961548212
H	3.412213238	0.005448038	-0.577864018
H	1.159562066	0.052972660	-1.702701352

Optimized cytosine FC geometry (MP2/6-31G)**

C	0.066422346	0.067298336	0.008288115
N	0.118124851	0.043932684	1.365403531
C	1.322481420	-0.003935816	2.113722006
N	2.491300447	-0.015778955	1.376294620
C	2.437826298	0.003083693	0.058991639
C	1.222265783	0.044982649	-0.707483998
O	1.257651773	-0.022950209	3.338973977
N	3.630986405	-0.078004342	-0.608558437
H	1.218752984	0.047461010	-1.786689002
H	-0.916292301	0.099571496	-0.443219120
H	-0.724249098	0.057471105	1.922121041
H	4.444858098	0.071765566	-0.032212642
H	3.679747993	0.270518783	-1.550896730

Optimized uracil FC geometry (MP2/6-31G)**

N	0.084109415	0.110677239	0.075938359
C	-0.023656707	0.095386081	1.461726570
N	1.211589383	0.058097642	2.087937735
C	2.495229818	0.047296664	1.508841107
C	2.467576982	0.064450258	0.051402091
C	1.283071108	0.096180001	-0.600620648
H	3.405024046	0.055789751	-0.480564299
O	3.499403825	0.025541777	2.215684892
O	-1.097259583	0.112269041	2.050826475
H	1.207865692	0.112105226	-1.679236026
H	-0.794434403	0.134672052	-0.418307743
H	1.183134425	0.053264268	3.099859488

REFERENCES

REFERENCES

(1) Tully, J. C.; Preston, R. K. *J Chem Phys* **1971**, *55*, 562.

(2) Tully, J. C. *J Chem Phys* **1990**, *93*, 1061.

(3) Mclachlan, A. D. *Mol Phys* **1964**, *8*, 39.

(4) Micha, D. A. *J Chem Phys* **1983**, *78*, 7138.

(5) Kirson, Z.; Gerber, R. B.; Nitzan, A.; Ratner, M. A. *Surf Sci* **1984**, *137*, 527.

(6) Sawada, S. I.; Nitzan, A.; Metiu, H. *Phys Rev B* **1985**, *32*, 851.

(7) Gerber, R. B.; Buch, V.; Ratner, M. A. *J Chem Phys* **1982**, *77*, 3022.

(8) Beck, M. H.; Jackle, A.; Worth, G. A.; Meyer, H. D. *Phys Rep* **2000**, *324*, 1.

(9) Wang, H. B.; Thoss, M. *J Chem Phys* **2003**, *119*, 1289.

(10) Meyer, H. D.; Worth, G. A. *Theor Chem Acc* **2003**, *109*, 251.

(11) Ben-Nun, M.; Martinez, T. J. *J Chem Phys* **1998**, *108*, 7244.

(12) Levine, B. G.; Coe, J. D.; Virshup, A. M.; Martinez, T. J. *Chem Phys* **2008**, *347*, 3.

(13) Makhov, D. V.; Glover, W. J.; Martinez, T. J.; Shalashilin, D. V. *J Chem Phys* **2014**, *141*.

(14) Topaler, M.; Makri, N. *J Phys Chem-Us* **1996**, *100*, 4430.

(15) Shim, S.; Aspuru-Guzik, A. *J Chem Phys* **2012**, *137*.

(16) Miller, W. H. *Chem Phys Lett* **1970**, *7*, 431.

(17) Herman, M. F. *J Chem Phys* **1984**, *81*, 754.

(18) Sun, X.; Wang, H. B.; Miller, W. H. *J Chem Phys* **1998**, *109*, 7064.

(19) Thoss, M.; Stock, G. *Phys Rev A* **1999**, *59*, 64.

(20) Bonella, S.; Coker, D. F. *J Chem Phys* **2001**, *114*, 7778.

(21) Virshup, A. M.; Punwong, C.; Pogorelov, T. V.; Lindquist, B. A.; Ko, C.; Martinez, T. J. *J Phys Chem B* **2009**, *113*, 3280.

(22) Zhu, C. Y.; Teranishi, Y.; Nakamura, H. Advances in Chemical Physics, Vol. 117 2001, 117, 127.

(23) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. *Chem Phys* **1980**, *48*, 157.

(24) Lengsfield, B. H.; Saxe, P.; Yarkony, D. R. *J Chem Phys* **1984**, *81*, 4549.

(25) Tao, H. L.; Levine, B. G.; Martinez, T. J. *J Phys Chem A* **2009**, *113*, 13656.

(26) Bauschlicher Jr, C. W.; Langhoff, S. R. *The Journal of chemical physics* **1988**, *89*, 4246.

(27) Demeras, A. S.; Lepetit, M. B.; Malrieu, J. P. *Chem Phys Lett* **1990**, *172*, 163.

(28) Zaitsevskii, A.; Malrieu, J. P. *Chem Phys Lett* **1994**, *228*, 458.

(29) Shu, Y. N.; Levine, B. G. *J Chem Phys* **2013**, *139*.

(30) Potts, D. M.; Taylor, C. M.; Chaudhuri, R. K.; Freed, K. F. *J. Chem. Phys.* **2001**, *114*, 2592.

(31) Bofill, J. M.; Pulay, P. *J. Chem. Phys.* **1989**, *90*, 3637.

(32) Abrams, M. L.; Sherrill, C. D. *Chem. Phys. Lett.* **2004**, *395*, 227.

(33) Lu, Z.; Matsika, S. *J. Chem. Theory Comput.* **2012**, *8*, 509.

(34) Granucci, G.; Persico, M.; Toniolo, A. *J. Chem. Phys.* **2001**, *114*, 10608.

(35) Slavicek, P.; Martinez, T. J. *J. Chem. Phys.* **2010**, *132*, 234102.

(36) Shu, Y. A.; Hohenstein, E. G.; Levine, B. G. *J Chem Phys* **2015**, *142*.

(37) Werner, H. J.; Knowles, P. J. *The Journal of chemical physics* **1985**, *82*, 5053.

(38) Knowles, P. J.; Werner, H.-J. *Chem Phys Lett* **1985**, *115*, 259.

(39) Lykos, P.; Pratt, G. *Rev Mod Phys* **1963**, *35*, 496.

(40) Jahn, H. A.; Teller, E. *Proc R Soc Lon Ser-A* **1937**, *161*, 220.

(41) Mclean, A. D.; Lengsfield, B. H.; Pacansky, J.; Ellinger, Y. *J Chem Phys* **1985**, *83*, 3567.

(42) Ayala, P. Y.; Schlegel, H. B. *J Chem Phys* **1998**, *108*, 7560.

(43) McGrady, J. E.; Lovell, T.; Stranger, R. *Inorg Chem* **1997**, *36*, 3242.

(44) Li, X. Z.; Paldus, J. *J Chem Phys* **2009**, *131*.

(45) Sorkin, A.; Iron, M. A.; Truhlar, D. G. *J Chem Theory Comput* **2008**, *4*, 307.

(46) Edgecombe, K. E.; Becke, A. D. *Chem Phys Lett* **1995**, *244*, 427.

(47) Jimenez-Hoyos, C. A.; Henderson, T. M.; Scuseria, G. E. J Chem Theory Comput 2011, *7*, 2667.

(48) Levine, B. G.; Ko, C.; Quenneville, J.; Martinez, T. J. *Mol Phys* **2006**, *104*, 1039.

(49) Alexandrova, A. N.; Tully, J. C.; Granucci, G. *J Phys Chem B* **2010**, *114*, 12116.

(50)Valeur, B.; Berberan-Santos, M. N. *Molecular fluorescence: principles and applications*; John Wiley & Sons, 2012.

(51) Hurley, R.; Testa, A. *J Am Chem Soc* **1966**, *88*, 4330.

(52) Siewertsen, R.; Schönborn, J. B.; Hartke, B.; Renth, F.; Temps, F. Phys Chem Chem Phys 2011, 13, 1054.

(53) Lu, Y. C.; Chang, C. W.; Diau, E. W. G. *J Chin Chem Soc-Taip* **2002**, *49*, 693.

(54) Yuan, S.; Wu, W. F.; Dou, Y.; Zhao, J. S. *Chinese Chem Lett* **2008**, *19*, 1379.

(55) Shu, Y. A.; Levine, B. G. *J Phys Chem C* **2014**, *118*, 7669.

(56) Meek, G. A.; Baczewski, A. D.; Little, D. J.; Levine, B. G. *J Phys Chem C* **2014**, *118*, 4023.

(57) Wang, X.; Maeda, K.; Thomas, A.; Takanabe, K.; Xin, G.; Carlsson, J. M.; Domen, K.; Antonietti, M. Nature materials 2009, 8, 76.

(58) Xu, X.; Liu, G.; Randorn, C.; Irvine, J. T. *Int J Hydrogen Energ* **2011**, *36*, 13501.

(59) Wang, X.; Blechert, S.; Antonietti, M. *Acs Catalysis* **2012**, *2*, 1596.

Chapter 7

Non-Radiative Recombination via Conical Intersection at Silicon Epoxide Defect

"There are two possible outcomes: if the result confirms the hypothesis, then you've made a measurement. If the result is contrary to the hypothesis, then you've made a discovery."

-Enrico Fermi (1901-1954)

This chapter is reproduced from paper: Non-radiative recombination via conical intersection at a semiconductor defect, Y. Shu, B. G. Levine, J. Chem. Phys., 139, 081102 (2013).

The ability to tune the optical properties of silicon nanocrystals (Si-NCs) by surface modification makes them appealing materials for light emission, sensing, and other applications where incorporation with existing silicon materials would be valuable. Characterizing the role of surface defects in trapping charge carriers and facilitating radiative and non-radiative decay would allow engineers to rationally design materials with desirable properties, but these processes involve a complex series of coupled electronic and nuclear motions which are difficult both to measure experimentally and to model theoretically. Though the photoluminescence (PL) of silicon nanostructures[1] was

first reported more than two decades ago, a complete microscopic understanding of the effect of surface oxidation on Si-NC photochemistry remains elusive due to the complex interplay of structure, electronic structure, and dynamics involved in Si-NC photochemistry.

In this chapter, the direct simulation of fast and efficient non-radiative (NR) decay via a conical intersection resulting from deformation of a known oxygen-containing defect at the silicon surface is reported. Conical intersections,[2] molecular structures that bring abobut true degeneracy between adiabatic electronic states, have long been known to facilitate NR decay in molecules. Our simulations demonstrate that Shockley-Read-Hall (SRH) non-radiative recombination,[3,4] which follows the localization of excitations to molecule-sized semiconductor defects, can occur via conical intersection as well. In addition, the identification and analysis of a conical intersection that arises at a known oxygen-containing defect provides insights into the experimentally observed PL of oxidized Si-NCs.

Modeling non-radiative recombination at defects requires a careful treatment of the nonadiabatic coupling between nuclear and electronic motions, a quantum mechanical effect with no classical analogue. Thus a thorough theoretical treatment requires consideration of the quantal nature of nuclear motion. To this end, the *ab initio* multiple spawning method[5] (AIMS) is applied to study the photodynamics of a cluster model of an oxygen-containing defect on the silicon surface. Application of similar nonadiabatic molecular dynamics methods to excited semiconductor materials have

provided important insights into the mechanisms of relaxation of hot electrons, charge and energy transfer, and multiple exciton generation.[6-9] The AIMS method is a hierarchy of approximations which allows the solution of the full time-dependent Schrodinger equation of a chemical system, including both nuclear and electronic degrees of freedom. Though described in full mathematical detail elsewhere,[5,10] here the features of the AIMS ansatz most relevant to the present work are briefly summarized. The AIMS nuclear wavefunction is expanded in a basis of frozen Gaussian functions,[11] which are associated with a particular adiabatic electronic state and whose average positions and momenta evolve according to classical equations of motion. In this sense, each individual basis function evolves like an *ab initio* molecular dynamics trajectory. The time-dependent expansion coefficients of the nuclear basis functions are determined by integration of the time-dependent Schrodinger equation. To accurately model nonadiabatic events, additional basis functions are added to the basis set ("spawned") when one of the frozen Gaussian functions enters a region of high nonadiabatic coupling to another state. The electronic wavefunction is determined by CASSCF[12] calculations performed on-the-fly.

In this work AIMS is applied to model the photodynamics of a silicon cluster containing a silicon epoxide unit, a three-membered ring composed of two silicon atoms and an oxygen atom. The presence of silicon epoxide defects upon oxidation of silicon surfaces has been observed by infrared spectroscopy,[13] but little is known about the effect these defects have on Si-NC PL or photochemistry. The $Si_8H_{12}O$ cluster studied in this work (pictured in the inset of Figure 7-1) has all its dangling bonds capped by hydrogen

atoms. The initial wavepacket is prepared assuming an instantaneous excitation; 20 initial nuclear basis functions with their average positions and momenta drawn from the Wigner distribution of the vibrational ground state of the ground electronic state are placed on the first adiabatic excited state of the cluster. Adiabatic state labels (ground state, S_0; first excited state S_1) are used throughout this work and should not be interpreted to indicate anything other than the ordering of the states. The FMS-MolPro software package[10] (based on the MolPro electronic structure package[14]) was used to perform the dynamical calculations.[15]

Within 50 fs after excitation, the silicon epoxide unit undergoes geometric distortions which result in a reduction of the S_1-S_0 energy gap from its initial value (5 eV on average) to nearly zero, and this reduction in energy gap corresponds with the onset of NR decay to S_0, as shown in Figures 7-1 and 7-2a. The NR decay is extremely efficient, with over half of the population reaching S_0 within 100 fs after excitation, and essentially all of the population quenched to S_0 by 400 fs after excitation. As shown in Figure 7-2b, this nonadiabatic decay is accompanied by an increase in the Si-O-Si bending angle from its initial value of ~80º to a maximum approaching 180º at 80 fs after excitation, corresponding to the breaking of the Si-Si bond of the epoxide ring. All 20 initial basis functions exhibit identical ring opening behavior.

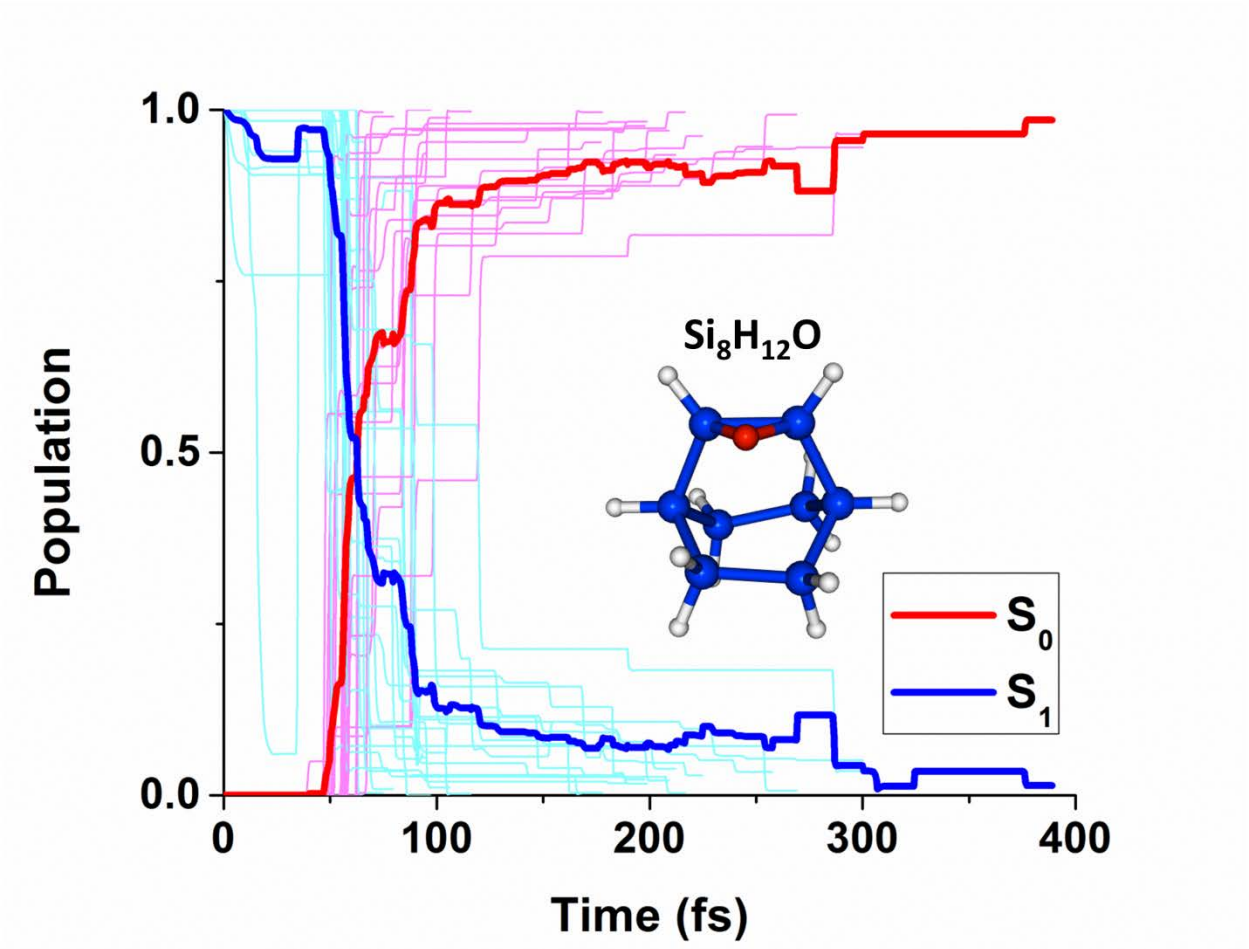


Figure 7-1. The population of adiabatic states S_0 and S_1 as a function of time after excitation. The average over all 20 initial nuclear basis functions is shown in bold red and blue for S_0 and S_1, respectively, while the values associated with individual initial basis functions are represented by thin lines. The $Si_8H_{12}O$ cluster is shown in the inset, with silicon, oxygen, and hydrogen atoms represented by blue, red, and white, respectively.

The efficient and time-local nature of the NR decay suggests the existence of a low-energy conical intersection in this system. To demonstrate the existence of and characterize this intersection, the CIOpt software package[16] was used to identify the

minimal energy point on the conical intersection seam (MECI) of the $Si_8H_{12}O$ cluster at the same CASSCF level of theory employed in the AIMS dynamics simulations. The structure and energy of this S_0/S_1 MECI are presented in Figure 7-3 along with optimized ground and excited state minima for comparison (denoted Franck-Condon (FC) point and S_1 min, respectively). The S_1 min and the S_0/S_1 MECI are similar in both energy and structure. Both are ring-opened structures similar to those explored by the dynamical simulations (with Si-O-Si angles of 175º and 164º for the S_0/S_1 MECI and S_1 min, respectively). The S_0/S_1 MECI is predicted to be 2.78 eV above the S_0 energy at the FC point and is only 0.04 eV above the S_1 minimum. This barrier is on the order of *kT* at room temperature, and thus not a significant hindrance to NR decay of a thermalized electronically excited population.

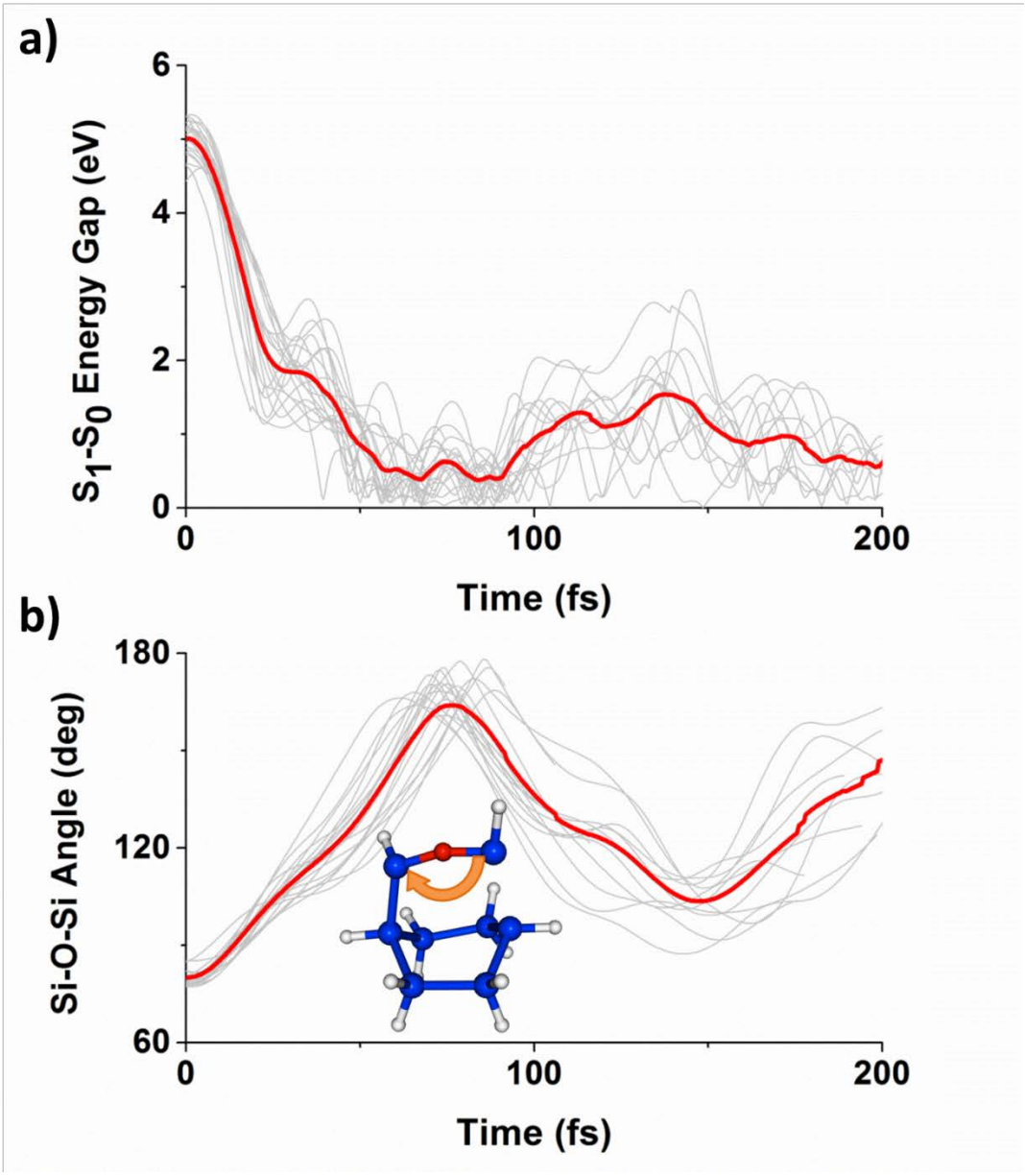


Figure 7-2. The a) S_1-S_0 energy gap and b) the Si-O-Si angle of the epoxide ring (defined in the inset) as a function of time after excitation. The average over the 20 initial basis functions on S_1 is shown in bold red lines, while the values of the individual basis functions are presented in thin gray lines. (Population on S_0 is excluded from this average.)

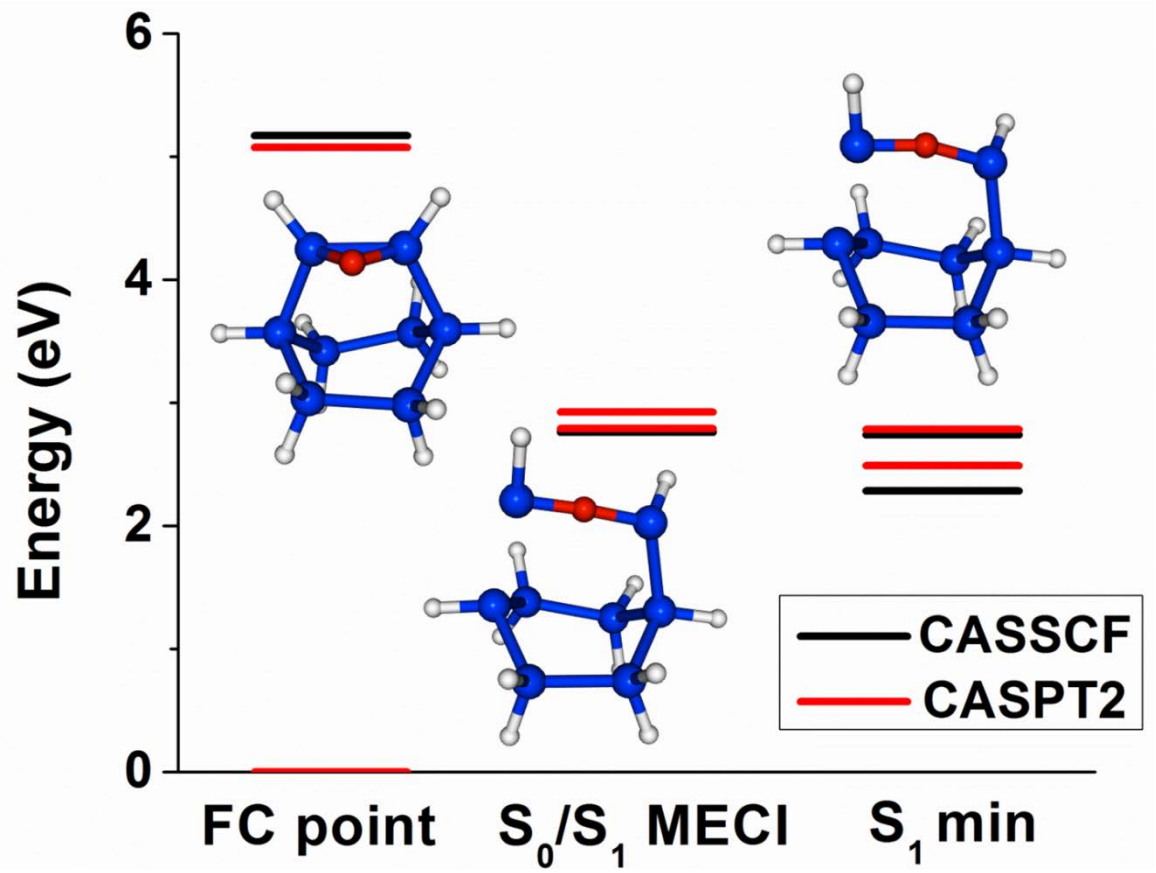


Figure 7-3. The S_0 and S_1 energies of three important points on the PES of the silicon epoxide cluster: the Franck-Condon (FC) point, the S_0/S_1 MECI, and the S_1 min. Black and red lines correspond to the energies of the CASSCF optimized geometries computed at the CASSCF and CASPT2 levels of theory, respectively. Energies are shifted so that the S_0 energy of the FC point is zero for both levels of theory. Only one black line is visible for the S_0/S_1 MECI because S_0 and S_1 are degenerate. Optimized geometries are shown as insets.

The electronic excitation is localized on or near the epoxide unit at both the FC point and the MECI geometry, as seen in the partially occupied natural orbitals (occupation numbers, $0.1 < n < 1.9$) presented in Figure 7-4. The FC point natural orbitals were determined by diagonalization of the S_1 reduced first order density matrix, while the MECI natural orbitals were determined by diagonalization of the state-averaged S_0/S_1 reduced first order density matrix to determine a set of orbitals which are invariant to rotations between these degenerate states. At the FC point, the lowest lying excitation is localized entirely on the epoxide unit. At the MECI geometry, the transition between states involves the movement of an electron from a non-bonding orbital on one of the epoxide silicon atoms to an antibonding orbital between one of the epoxide silicon atoms and an adjacent silicon atom. In neither structure does the excitation extend beyond the atoms adjacent to the epoxide ring.

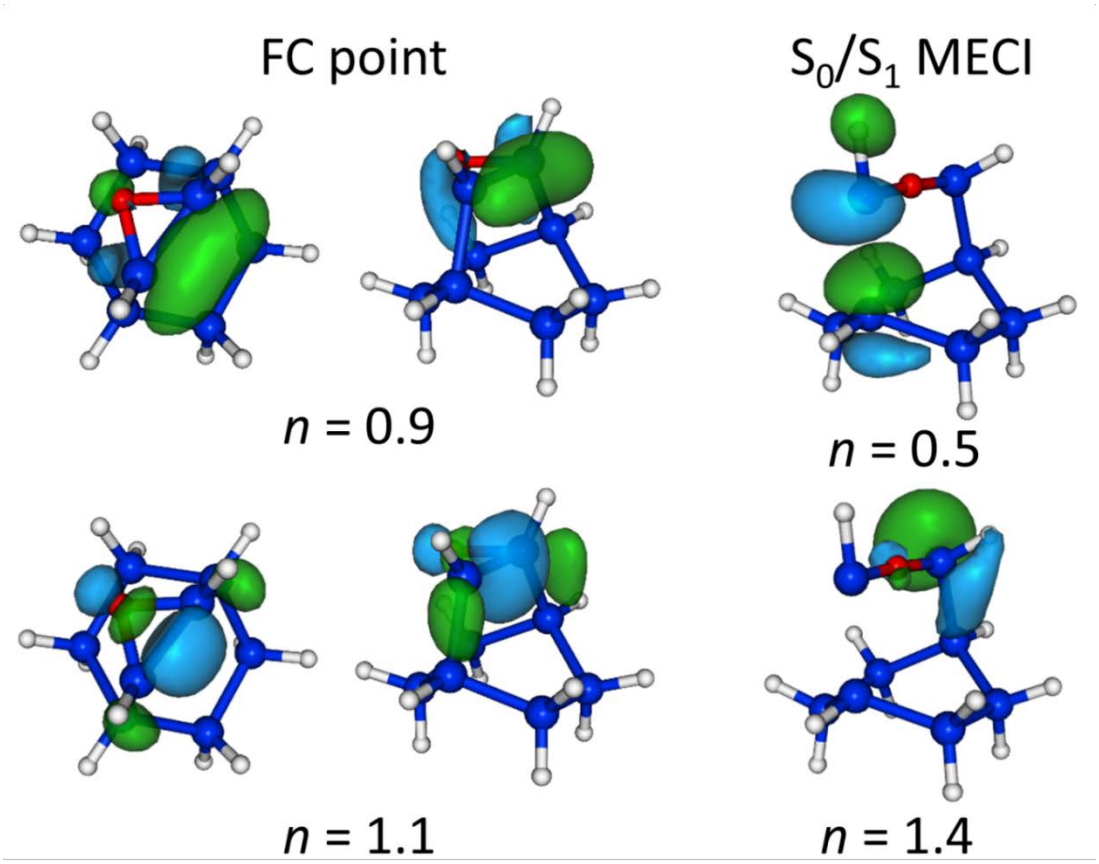


Figure 7-4. Partially occupied natural orbitals of the CASSCF wavefunctions at the FC point (left) and the optimized S_0/S_1 MECI structure (right). Two different views of each orbital are presented for the FC point. Occupation numbers, *n*, are presented for each orbital. Orbitals were determined as described in the text.

In the context of a larger material, the predicted conical intersection can be thought of as a deep-level trap brought about by nuclear distortion of the oxygen defect in response to electronic excitation. Obviously the existence of a fast and efficient mechanism for NR decay at an oxygen-containing surface defect provides a possible explanation for the observed decrease in PL yield upon oxidation, but the 2.78 eV energy

of the MECI relative to the ground state energy of the system suggests a possible role in the observed PL spectrum as well. The localization of the excitation to the silicon epoxide unit suggests that the energy of the conical intersection relative to the S_0 minimum would not be strongly affected by the extension of the present cluster to the nanometer scale. If a similar intersection exists at 2.8 eV above the ground state of an oxidized Si-NC, its role in the photochemistry will depend on the band gap of that particular Si-NC. Assuming trapping occurs on timescales shorter than the μs radiative lifetime of excitons in Si-NCs,[17] silicon epoxide defects will act as centers for NR decay in Si-NCs with a band gap above ~2.8 eV (those where the energy of an excited electron-hole pair is sufficient to reach the MECI region of the PES), but would remain inaccessible in larger Si-NCs with smaller band gaps. Thus, larger gapped Si-NCs are effectively removed from the ensemble-averaged PL, leaving only smaller gapped particles to luminesce. Such selective quenching would result in the experimentally observed red-shifted, size-insensitive PL of ensembles of oxidized Si-NCs even in the absence of emissive defects. Silicon epoxide defects also provide a potential explanation for the observation that only a small percentage of oxidized Si-NCs in a given sample are emissive[18] and that ensembles of oxidized Si-NCs often contain no emitters with PL energies above 2.5 eV.[19,20] The observation of short (nanosecond) PL lifetimes and oriented transition dipoles in single particle experiments, however, supports the existence of emissive defects in addition to the non-radiative ones reported here,[19] and the ensemble-averaged PL is likely determined by the existence of both radiative and non-radiative defects on the oxidized

surface.

The CASSCF wavefunction employed in our dynamical study was chosen because CASSCF is capable of describing the static electron correlation present when a small number of electronic configurations are mixed, as is the case near conical intersections. However, the inability of CASSCF to accurately describe dynamic electron correlation is a potential source of error in our dynamics simulations. To demonstrate the accuracy of the CASSCF surface, additional calculations at the multi-state complete active space second order perturbation theory (CASPT2) level,[21] which accurately describes both static and dynamical correlation effects, were performed. Single point CASPT2 energy calculations at the CASSCF optimized geometries show near quantitative agreement between the two levels of theory. CASPT2 predicts a small energy gap of 0.13 eV at the S_0/S_1 MECI structure, and CASPT2 energies of 2.92 and 2.78 eV computed at the S_0/S_1 MECI and S_1 min, respectively, agree well with the CASSCF values of 2.78 and 2.74 eV. Reoptimization of these two structures at the CASPT2 level of theory results in energies (2.76 and 2.75 eV, respectively) in even closer agreement with CASSCF and does not significantly affect the structures.[22] Single point CASPT2 calculations performed at the spawning geometries from the dynamics simulations (those geometries where new nuclear basis functions are created due to a large non-adiabatic coupling) also show excellent agreement between the CASSCF PES and that of CASPT2.[23]

In the present study, we identified a pathway for fast and efficient NR decay via a conical intersection resulting from the excitation and subsequent nuclear deformation of

a known defect on the oxidized silicon surface. The energy of the conical intersection responsible for this decay suggests that this pathway might be responsible for the observed decrease in PL yield upon oxidation. In addition, the preferential quenching of small diameter, large-gap Si-NCs by this intersection may contribute to the observed size-insensitivity of the ensemble PL observed upon oxidation. Future studies will investigate potential mechanisms of NR decay in other oxygen containing silicon defects. In addition, work is underway to develop tools to allow us to scale up our multireference nonadiabatic molecular dynamics simulations to the nanoscale.

The authors thank Michigan State University (MSU) for start-up funding, the MSU Institute for Cyber-Enabled Research for computer time, and Paul Reed for technical assistance.

REFERENCES

REFERENCES

(1) Canham, L. T. *Appl Phys Lett* **1990**, *57*, 1046.

(2) Domcke, W.; Yarkony, D. R. *Annu. Rev. Phys. Chem.* **2012**, *63*, 325.

(3) Shockley, W.; Read, W. T. *Phys. Rev.* **1952**, *87*, 835.

(4) Hall, R. N. *Phys. Rev.* **1952**, *87*, 387.

(5) Ben-Nun, M.; Quenneville, J.; Martinez, T. J. *J. Phys. Chem. A* **2000**, *104*, 5161.

(6) Abuabara, S. G.; Rego, L. G. C.; Batista, V. S. *J. Am. Chem. Soc.* **2005**, *127*, 18234.

(7) Craig, C. F.; Duncan, W. R.; Prezhdo, O. V. *Phys Rev Lett* **2005**, *95*.

(8) Nelson, T.; Fernandez-Alberti, S.; Chernyak, V.; Roitberg, A. E.; Tretiak, S. *J. Phys. Chem. B* **2011**, *115*, 5402.

(9) Jailaubekov, A. E.; Willard, A. P.; Tritsch, J. R.; Chan, W. L.; Sai, N.; Gearba, R.; Kaake, L. G.; Williams, K. J.; Leung, K.; Rossky, P. J.; Zhu, X. Y. *Nat. Mater.* **2013**, *12*, 66.

(10) Levine, B. G.; Coe, J. D.; Virshup, A. M.; Martinez, T. J. *Chem. Phys.* **2008**, *347*, 3.

(11) Heller, E. J. *J Chem Phys* **1981**, *75*, 2923.

(12) Roos, B. O.; Taylor, P. R. *Chem. Phys.* **1980**, *48*, 157.

(13) Stefanov, B. B.; Gurevich, A. B.; Weldon, M. K.; Raghavachari, K.; Chabal, Y. J. Phys. Rev. Lett. 1998, 81, 3908.

(14) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schutz, M.; Celani, P.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Koppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; O'Neill, D. P.; Palmieri, P.; Pfluger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stol, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M.; Wolf, A. MOLPRO, version 2010.1, a package of ab initio programs, 2010.

(15) See Supplementary Materials for additional simulation details.

(16) Levine, B. G.; Coe, J. D.; Martinez, T. J. *J Phys Chem B* **2008**, *112*, 405.

(17) Huisken, F.; Ledoux, G.; Guillois, O.; Reynaud, C. *Adv. Mater.* **2002**, *14*, 1861.

(18) Valenta, J.; Juhasz, R.; Linnros, J. *Appl. Phys. Lett.* **2002**, *80*, 1070.

(19) Schmidt, T.; Chizhik, A. I.; Chizhik, A. M.; Potrick, K.; Meixner, A. J.; Huisken, F. *Phys. Rev. B* **2012**, *86*.

(20) Martin, J.; Cichos, F.; Huisken, F.; von Borczyskowski, C. *Nano Lett.* **2008**, *8*, 656.

(21) Finley, J.; Malmqvist, P. A.; Roos, B. O.; Serrano-Andres, L. Chem. Phys. Lett. 1998, 288, 299.

(22) See Supplementary Materials for details of CASPT2 optimized conical intersections and excited state minimum geometries.

(23) See Supplementary Materials for CASPT2 energies calculated at CASSCF spawning geometries.

Chapter 8

Non-Radiative Recombination via Conical Intersections Arising at Defects on The Oxidized Silicon Surface

"In science one tries to tell people, in such a way as to be understood by everyone, something that no one ever knew before."

-Paul A.M. Dirac (1902-1984)

This chapter is reproduced from paper: Nonradiative recombination via conical intersections arising at defects on the oxidized silicon surface, Y. Shu and B. G. Levine, J. Phys. Chem. C, 119, 1737 (2015).

In our previous chapter of epoxide defects we focused on defect structures containing a single oxygen atom, but these defects often occur as part of an oxide layer, and infrared spectroscopy suggests that one or both of the epoxide silicon atoms are generally bound to additional oxygen atoms.[1] Therefore, in this chapter we consider defects containing multiple oxygen atoms. Using nonadiabatic molecular dynamics and multireference quantum chemical approaches, we model the dynamics of cluster models of two such defects following excitation. In doing so, we address the question of how further oxidation of the epoxide defect affects its photochemistry, and what

consequences this might have for our understanding of the PL of oxidized SiNCs.

8.1 Computational Methods

To model the dynamics of electronically excited defects, we apply ab initio multiple spawning (AIMS),[2,3] a hierarchy of approximations which allows us to model the full, time-dependent wavefunction of a molecule after excitation. By working in a relatively small time-dependent basis we can describe all nuclear and electronic degrees of freedom explicitly. The adiabatic electronic wavefunction is calculated on-the-fly at the state averaged complete active space self-consistent field (SA-CASSCF) level of theory.[4] The chosen active space and number of states to be averaged for each cluster are listed in Table 8-1. The LANL2DZdp basis and effective core potentials (ECPs) are used in all cases.[5] These parameters were chosen such that the relative energies and characters of the low-lying excited states are in good agreement with those computed at the equation-of-motion coupled cluster with single and double excitations[6] (EOM-CCSD) level of theory using the 6-31G** basis and that important points on the excited state PES can be validated at other highly correlated levels of theory, as will be discussed below.

Table 8-1. The size of the active space (abbreviated: number of active electrons/number of active orbitals) and number of states to be averaged in our SA-CASSCF calculations for each cluster. Also given are the numbers of occupied orbitals to be correlated in the MS-CASPT2 and MRCI(Q) calculations. Note that more orbitals are correlated in the MS-CASPT2 calculations at the spawning points from our AIMS simulations (reported in Figures 8-7 and 13) than in the remainder of the MS-CASPT2 calculations reported in this work.

Cluster	Active Space for SA-CASSCF	# of States in SA-CASSCF	# of Correlated Orbitals	# of Correlated Orbitals in Calculations of Spawning Points
Three-oxygen Epoxide (3O)	4/3	3	8	12
Five-oxygen Epoxide (5O)	2/4	2	7	11

All dynamical simulations were initiated on the S_1 state. Twenty individual simulations—each starting from a single initial nuclear basis function—were run for each cluster, with the initial positions and momenta of each initial basis function randomly sampled from the ground state vibrational Wigner distribution computed in the harmonic approximation at the MP2/6-31G** level of theory. The reported average data is incoherently averaged over all twenty simulations. The nuclear wavefunction is propagated using an adaptive multiple time step algorithm[7] with a 0.5 fs time step. The widths of the frozen Gaussian basis functions were chosen to be 25.0, 6.0, and 30.0 Bohr for silicon, hydrogen, and oxygen atoms, respectively. These widths were chosen based on the results of ref[8], noting that the rate of decay predicted by AIMS simulations depends only very weakly on this choice. Simulations are stopped when essentially all population has reached the ground state or after 1 ps, whichever is shorter. Thus, in figures showing data from the twenty individual simulations, not all curves are the same length. All dynamic simulations were conducted with the FMSMolPro software package.[3]

The PES was further validated by comparison of CASSCF-optimized minima and minimal energy conical intersections (MECIs) to those optimized using the highly accurate multistate complete active space second order perturbation theory (MS-CASPT2) method.[9] The numbers of orbitals correlated in the MS-CASPT2 calculations are listed in Table 8-1. Additional single point calculations at the Davidson-corrected multireference configuration interaction level of theory with single and double excitations[10] (MRCI(Q)) were performed on all optimized structures to provide an additional point of comparison.

The same active spaces and basis sets were used in the correlated calculations as described above for the CASSCF calculations. All electronic structure calculations were performed with the MolPro software package.[11-18] while geometry and MECI optimizations utilized CIOpt[19] in conjunction with MolPro.

In our previous work, we discussed what we will call the one-oxygen (1O) epoxide defect: a three-membered ring composed of two silicon and one oxygen atoms in which the silicon atoms are each bound to two silicon atoms external to the ring. In this work, we focus on two different defects, which we term the three-oxygen and five-oxygen epoxide defects. In the three-oxygen defect, one of the epoxide silicon atoms is bound to two external silicon atoms, while the other epoxide silicon atom is bound to two external oxygen atoms. In the five-oxygen defect, each epoxide silicon atom is bound to two external oxygen atoms. Both of these defects have been observed on the oxidized surface of silicon in vibrational spectroscopic experiments.[1] We will apply our computational methodology to cluster models of these two defects, pictured in Figure 8-1; the clusters we term the three-oxygen (3O) and five-oxygen (5O) clusters are represented in the left and right panels, respectively. All dangling bonds are terminated by hydrogen atoms in all clusters.

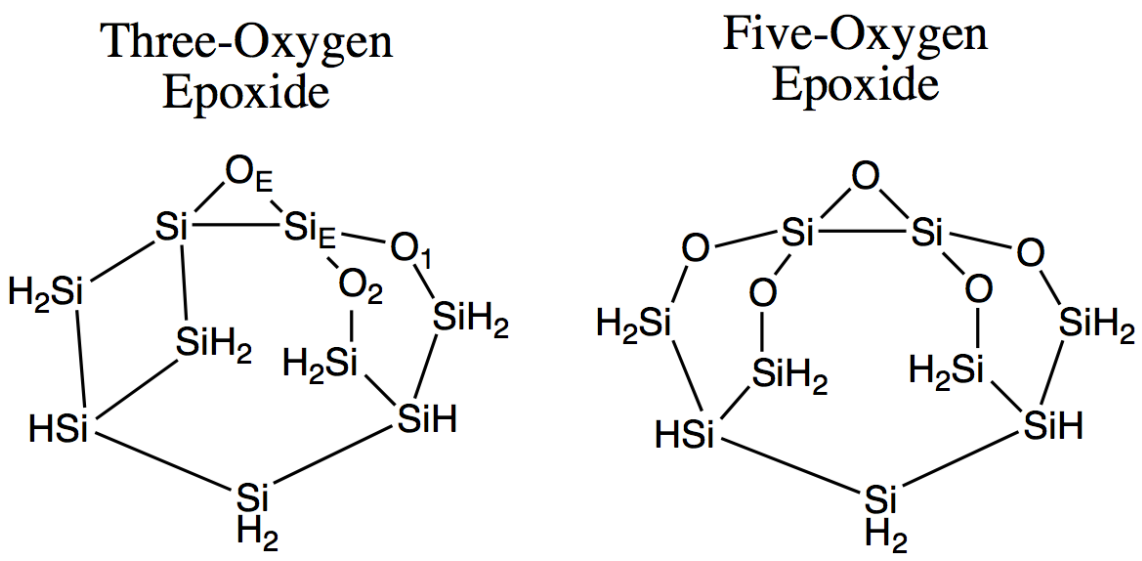


Figure 8-1. The 3O (left) and 5O (right) clusters.

8.2 Results

8.2.1 Dynamical simulation of the 3O cluster

First we discuss the dynamics of the 3O cluster following electronic excitation, as modeled by the AIMS method. The populations of the ground and first excited adiabatic electronic states (S_0 and S_1, respectively) as a function of time after excitation are presented in Figure 8-2. As can be seen, the dynamical simulations predict that 85% of the excited population decays to the ground state in the first picosecond after excitation. This corresponds to a lifetime that is much shorter than the microsecond radiative lifetime of the S-band PL of SiNCs.

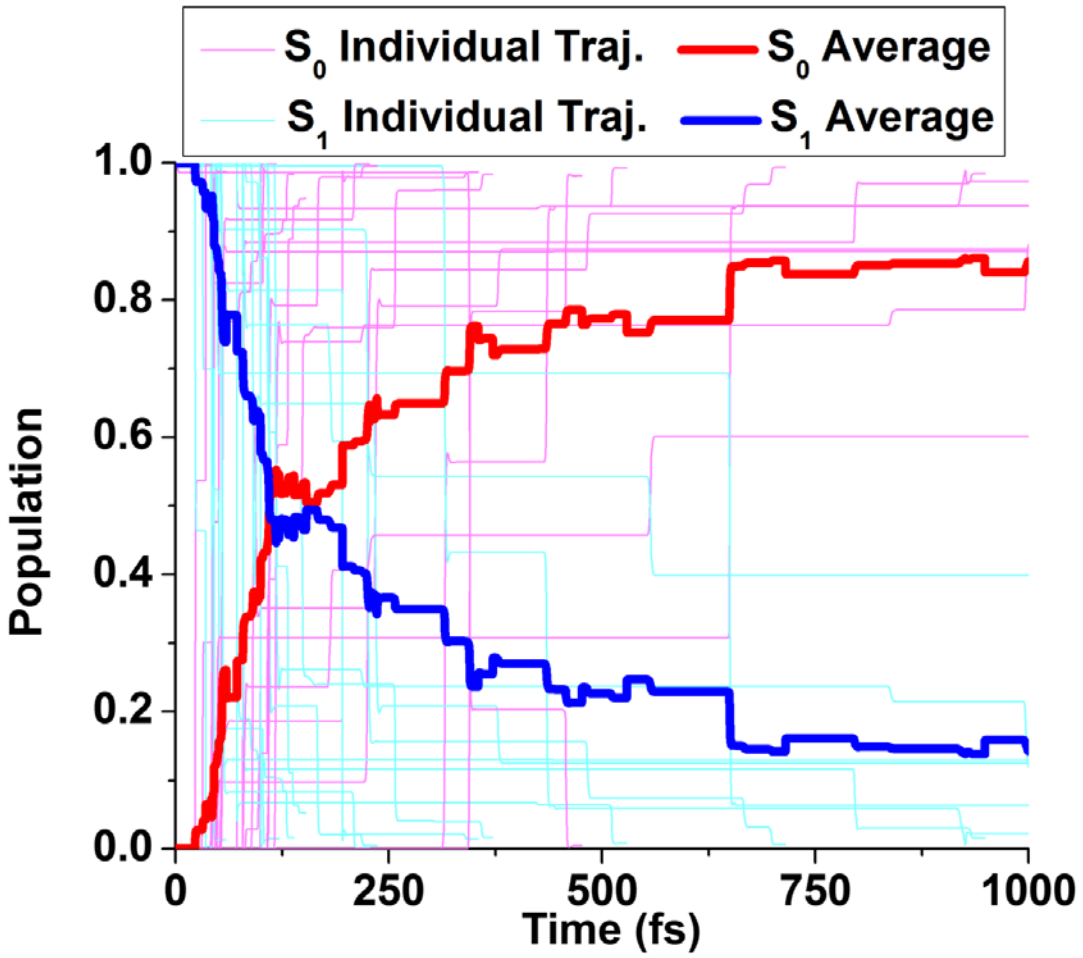


Figure 8-2. The population of S_1 (bold blue) and S_0 (bold red) of the 3O cluster as a function of time after excitation averaged over all simulations. Thin blue and pink lines report the populations corresponding to the twenty individual simulations.

The S_1-S_0 population decay is accompanied by a decrease in the S_1-S_0 energy gap, as shown in Figure 8-3. At the ground state minimum geometry, the CASSCF level of theory predicts a vertical excitation energy of 5.15 eV (as comparable to that computed at the EOM-CCSD/6-31G** level of theory: 4.41 eV), but the average S_1-S_0 energy gap falls below 1 eV in the first 50 fs after excitation and oscillates around 0.8 eV for the rest of the

simulation. The individual trajectories regularly explore regions of near-zero energy gap during these oscillation, and the initial exploration of these regions is coincident with the onset of extremely fast and efficient NR decay, suggesting that conical intersections may be accessed.

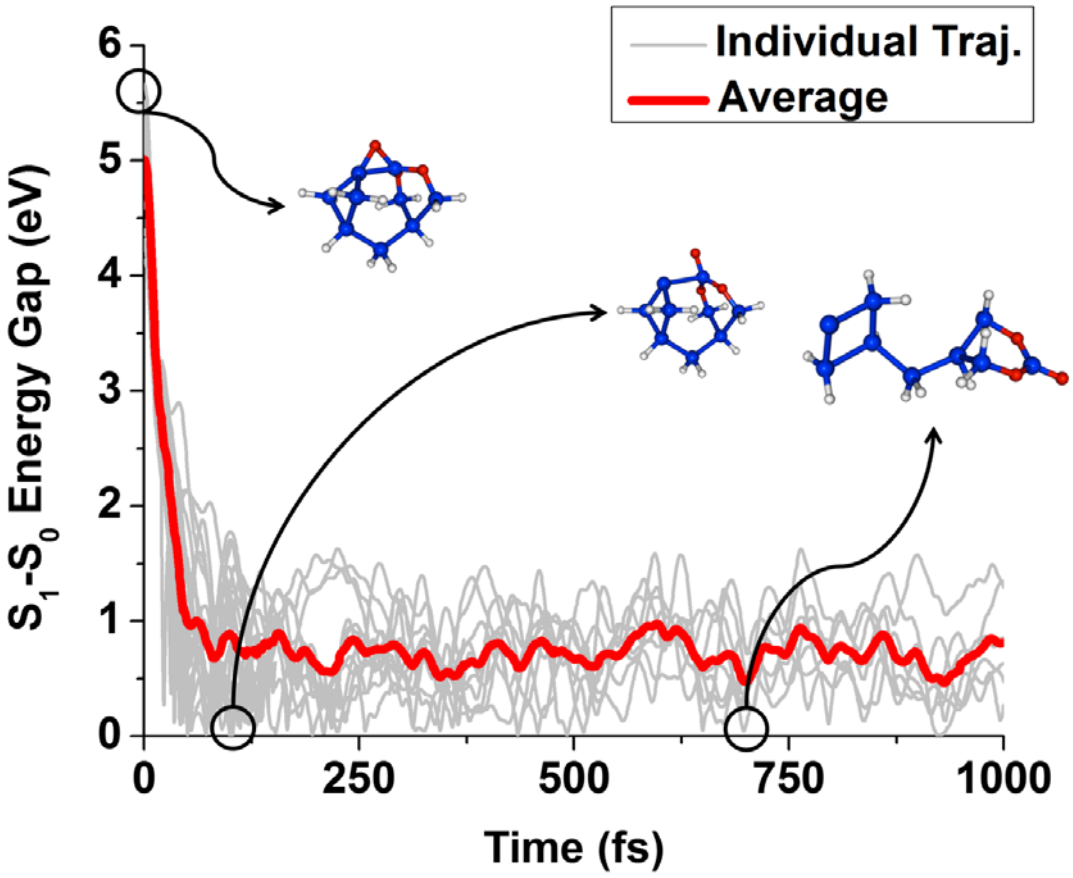


Figure 8-3. Average S_1-S_0 energy gap of the 3O cluster as a function of time (bold red), with the energy gaps of the twenty individual initial nuclear basis functions shown in gray. Snapshots from representative trajectories are shown in the inset, with silicon, oxygen, and hydrogen atoms colored blue, red, and white, respectively.

Significant geometric distortions precede the observed NR decay. In the first 200 fs after excitation, changes in the structure around the more oxidized epoxide silicon atom are observed. We define the O-SiOO pyramidalization angle

$$\theta = \pm \arccos\left(\frac{\left(\vec{R}_{Si_E O_E} \times \vec{R}_{O_1 O_2}\right) \bullet \left(\vec{R}_{Si_E O_1} \times \vec{R}_{Si_E O_2}\right)}{\left\|\vec{R}_{Si_E O_E} \times \vec{R}_{O_1 O_2}\right\| \left\|\vec{R}_{Si_E O_1} \times \vec{R}_{Si_E O_2}\right\|} \right)$$

where $\vec{R}_{AB}$ is the vector from atom *A* to atom *B*, Si_E is the more oxidized epoxide silicon atom, O_E is the oxygen in the epoxide ring, and O_1 and O_2 are the two oxygen atoms bound to Si_E but external to the ring (see Figure 8-1). The sign of this angle indicates the direction of the pyramidalization; negative values correspond to the case where Si_E is puckered in toward the center of the cluster, as is necessary to form the epoxide ring, while positive values correspond to the puckering of Si_E outward. The O-SiOO pyramidalization angle increases from about -10 degrees to zero in the first 50 fs after excitation, and oscillates around zero for the remainder of the simulation. The ballistic change in angle upon excitation suggests that there may be a strong signature in resonance Raman spectrum of this defect, and thus it is worth noting that the ground state normal mode (computed at the MP2/6-31G** level) corresponding to this pyramidalization motion has a vibrational frequency of 345.9 cm^{-1}.

Simultaneous with these pyramidalization motions, two bonds in the epoxide ring break, leaving a double bond between the more oxidized Si atom and the epoxide oxygen atom. Without these bonds, the cluster opens up, with the average distance between the two epoxide silicon atoms (Si-Si distance) increasing from 2.20 to 10 Å as shown in

Figure 8-4(b). A small minority of simulations remain closed. The large amplitude motion required to open the cluster is likely not possible in the constrained environment of a real SiNC surface. However, efficient NR decay begins before the Si-Si distance is so dramatically lengthened, and those simulations which do not result in complete opening of the cluster exhibit very fast decay, suggesting that conical intersections may be accessible before the cluster breaks open. This possibility will be considered in the next section. One might expect that the propensity for ring opening would depend on the initial energy of our simulations, which covers a range of 0.1 eV due to the fact that the initial conditions are sampled from the vibrational Wigner distribution. However, no such correlation is observed (see Figure S1 of supporting information http://pubs.acs.org/doi/suppl/10.1021/jp5114202 for details).

8.2.2 Potential energy surface analysis of the 3O cluster

The fact that the 3O cluster accesses regions of near-zero S_1-S_0 energy gap in our AIMS simulations indicates that NR decay may be facilitated by S_1-S_0 conical intersections. To further investigate this possibility, we applied the CASSCF and MS-CASPT2 theoretical levels to optimize both the excited state minima (S_1 min) and S_1-S_0 MECIs on the PES of the 3O cluster. Via these optimizations we identified and characterized two different regions of the PES: the closed region, in which the backbone of the cluster is not dramatically distorted, and the open region, in which two bonds of the epoxide ring are broken and the entire cluster is very distorted. Distinct S_1 min and S_1-S_0 MECI structures are found in each region and are labeled as closed or open, respectively. At both the

CASSCF- and CASPT2-optimized geometries, single point energy calculations are performed at the CASSCF, MS-CASPT2, and MRCI(Q) theoretical levels. The resulting energies are presented in Table 8-2 and Figure 8-5. The optimized geometries are presented as insets in Figure 8-5.

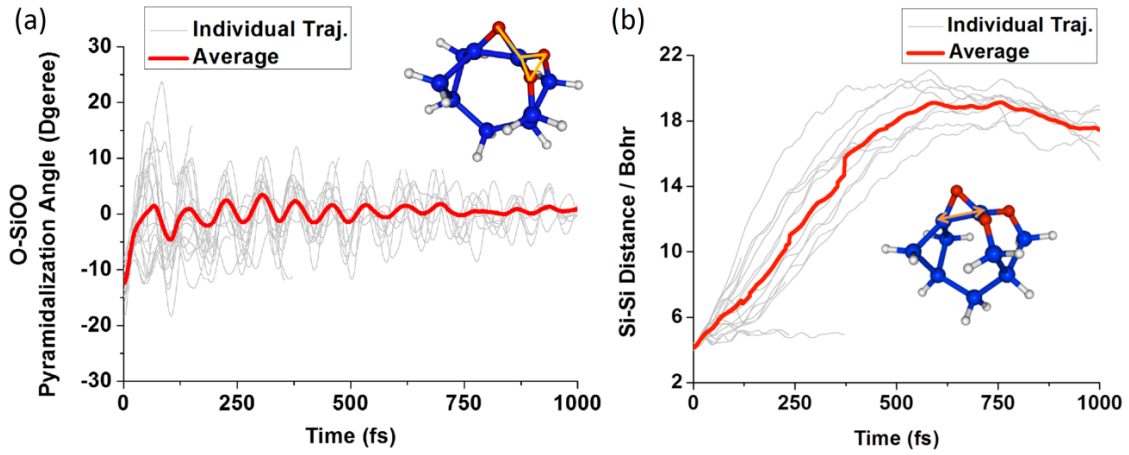


Figure 8-4. (a). Average O-SiOO pyramidalization angle of the 3O cluster as a function of time (bold red) with pyramidalization angles of the twenty individual initial nuclear basis functions (gray). The inset indicates the atoms that define the pyramidalization angle. (b). The distance between epoxide silicon atoms as a function of time averaged over twenty simulations (bold red) and the values corresponding to the twenty individual simulations themselves (gray). The inset indicates the two silicon atoms between which the distance was calculated.

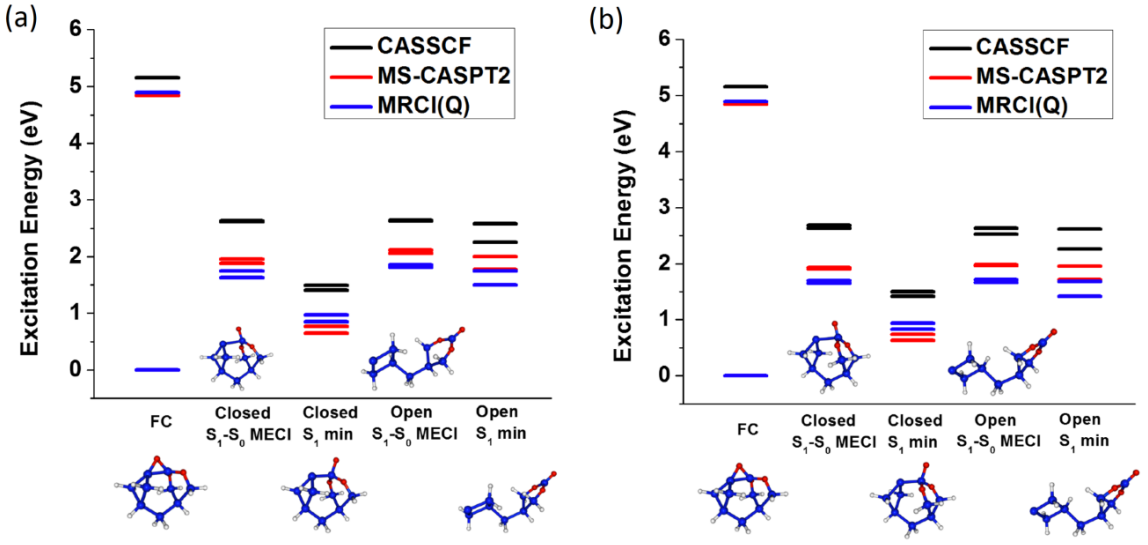


Figure 8-5. (a). S_0 and S_1 energies of the Franck-Condon (FC) point, the CASSCF-optimized S_1 minima, and the CASSCF-optimized S_1-S_0 minimal energy conical intersections of the 3O cluster as calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory (black, red, and blue bars, respectively). Geometries are shown as insets. (b). Same as (a) except that the excited state minima and MECIs have been optimized at the MS-CASPT2 level of theory.

Table 8-2. The S_0 and S_1 energies (in eV) at five important points on the PES of the 3O cluster calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory. S_1 minima and S_1-S_0 MECI geometries have been optimized at both the CASSCF and MS-CASPT2 levels of theory.

Geometry	Optimization Method	CASSCF S_0 Energy	CASSCF S_1 Energy	MS-CASPT2 S_0 Energy	MS-CASPT2 S_1 Energy	MRCI(Q) S_0 Energy	MRCI(Q) S_1 Energy
FC point	MP2	0.00	5.15	0.00	4.84	0.00	4.89
Closed S_1-S_0 CI	CASSCF	2.62	2.63	1.88	1.96	1.63	1.74
	MS-CASPT2	2.64	2.68	1.91	1.93	1.65	1.70
Closed S_1 min	CASSCF	1.41	1.49	0.65	0.76	0.85	0.97
	MS-CASPT2	1.42	1.50	0.64	0.74	0.84	0.94
Open S_1-S_0 CI	CASSCF	2.63	2.64	2.06	2.12	1.82	1.86
	MS-CASPT2	2.53	2.64	1.96	1.98	1.67	1.71
Open S_1 min	CASSCF	2.25	2.58	1.78	2.00	1.50	1.75
	MS-CASPT2	2.26	2.62	1.72	1.96	1.42	1.68

The CASSCF-optimized closed MECI geometry is found to be 2.63, 1.96, and 1.74 eV above the ground state minimum energy at the CASSCF, MS-CASPT2, and MRCI(Q) theoretical levels, respectively. The computed energies of this MECI are well within the visible range at all levels of theory, and thus would be accessible after excitation of a SiNC by visible light. CASSCF overestimates the energy of this MECI by 0.67 and 0.89 eV relative to the highly accurate MS-CASPT2 and MRCI(Q) theoretical levels. Such errors often arise at the CASSCF level of theory due to the absence of dynamic electron correlation. Still, CASPT2 and MRCI(Q) single point calculations at the CASSCF-optimized MECI points predict small S_1-S_0 energy gaps (0.08eV and 0.11eV), suggesting that CASSCF is predicting the conical intersection at roughly the correct nuclear configuration.

The CASSCF optimized closed S_1 min geometry is found to be 1.49, 0.76, and 0.97 eV above the ground state minimum energy at the CASSCF, CASPT2, and MRCI(Q) levels of theory, respectively. Thus, the energetic barrier to reach the conical intersection from the S_1 minimum is estimated to be 1.14, 1.20, and 0.77 eV at these respective levels. This relatively large barrier seems inconsistent with the fast and efficient NR decay observed in the dynamical simulations. An explanation for this discrepancy will be provided below. The optimized S_1 min and MECI geometries have similar structures, with the Si-O bonds between the epoxide oxygen atom and the less oxidized epoxide silicon atom broken and the other two bonds of the epoxide ring intact. The two geometries differ in the structure about the more oxidized epoxide silicon. For the S_1 min, the structure surrounding the oxidized Si atom takes on a roughly tetrahedral sp^3 configuration (with angles that vary from 80 to 125º). In contrast, the structure about this silicon atom in

the MECI is approximately trigonal pyramidal, with the three oxygen atoms occupying the equatorial positions. The distance between epoxide silicon atoms in both structures is 2.65 Å, which is significantly longer than the 2.20 Å Si-Si distance at the FC point, as well as the 2.35 Å Si-Si bond length in crystalline silicon, suggesting a considerable weakening of this bond in the excited state.

The partially occupied natural orbitals (PONOs) of the FC point and closed MECI computed at the CASSCF level of theory are presented in Figure 8-6 to elucidate the electronic structure at these geometries. The natural orbitals are generated by diagonalizing the first order reduced S_1 density matrix at the FC point and the S_1-S_0 state-averaged density matrix at the MECI. At the FC point, the transition can be described as the excitation of a single electron from an orbital of mixed Si-Si and Si-O σ bonding character (top left) to an antibonding orbital involving p orbitals of the epoxide O atom and the less oxidized epoxide silicon (bottom left). The removal of an electron from the σ bonding orbital explains the weakening of the Si-Si bond of the epoxide ring in the excited state. At the MECI geometry, the non-radiative transition is characterized by the transfer of a single electron between a nonbonding (or weakly antibonding) orbital localized on the less oxidized Si atom (top right) and a σ bonding orbital localized on the epoxide ring (bottom right). Mulliken charges at the FC and MECI geometries are available in Figure S2.

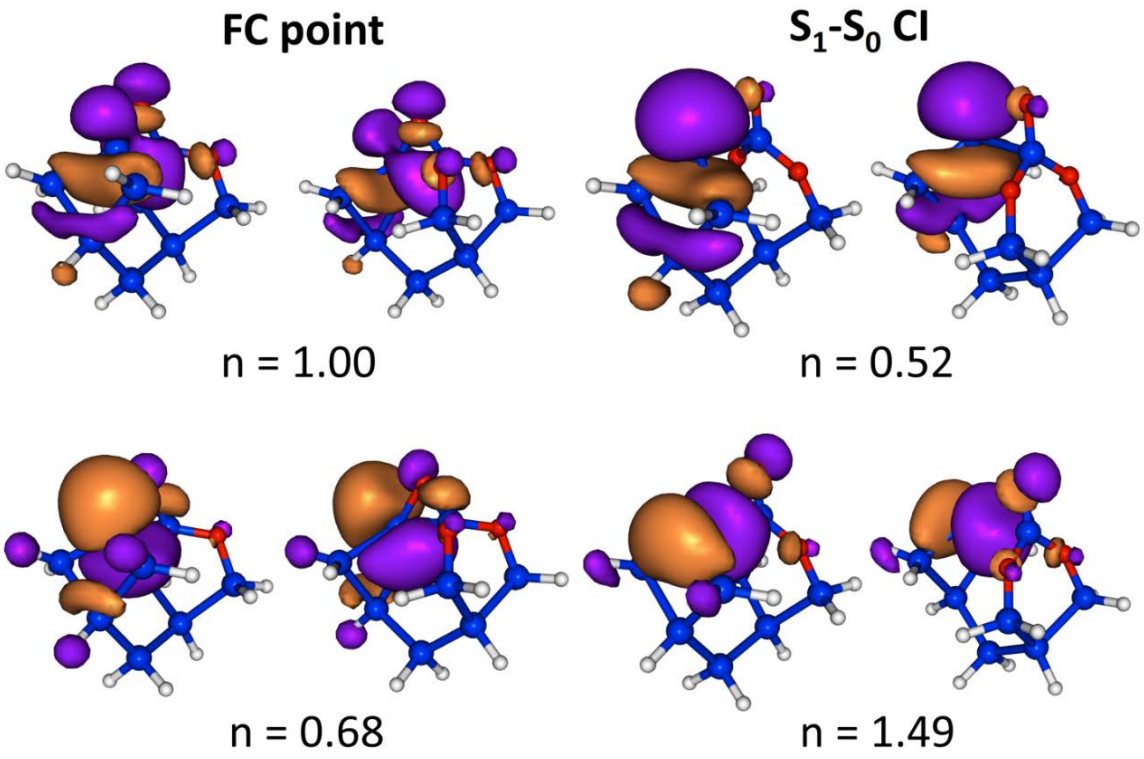


Figure 8-6. Partially occupied CASSCF natural orbitals of the 3O cluster at the FC point (left) and the closed MECI (right) are presented with occupation numbers. The FC natural orbitals are computed by diagonalization of the S_1 density matrix, while the S_1–S_0 CI natural orbitals are computed from the state-averaged S_1–S_0 density matrix. Two views of each orbital are presented for clarity.

Optimization of the closed MECI and S_1 min at the highly accurate MS-CASPT2 level of theory confirms the above described picture of the PES and validates the accuracy of the PES used in our dynamical simulations. As seen in Table 8-2 and Figure 8-5(b), the energies of the MS-CASPT2 optimized structures are generally within 0.1 eV of those optimized at the CASSCF level of theory, whether the energy is computed at the CASSCF,

MS-CASPT2, or MRCI(Q) level of theory. The optimized geometries themselves are qualitatively similar, as can be seen by comparing Figures 8-5(a) and (b).

While much of the population transfer to the ground state observed in the dynamic simulations occurs in the closed region of the PES, significant population transfer was also observed after the entire cluster has broken open (as indicated by the large increase in the Si-Si distance plotted in Figure 8-4(b)). As such, we have identified points of interest in this open region of the PES as well: the open S_1-S_0 MECI and the open S_1 min. These geometries and energies are pictured in Figure 8-5, with the energies presented again in Table 8-2.

The energy of the CASSCF-optimized open MECI is predicted to be 2.64, 2.12, and 1.86 eV above the ground state minimum energy by the CASSCF, CASPT2, and MRCI levels of theory, respectively. The CASSCF-optimized open S_1 min is found to be 2.58, 2.00, and 1.75 eV above the ground state minimum energy according to the same levels of theory. The small barrier that needs to be overcome in order to access the MECI from the S_1 minimum (~0.1 eV) may account for the extremely efficient NR decay observed in the simulations, which we previously noted seemed inconsistent with the large barrier predicted in the closed region. That the dynamical simulations explore the open region of the PES, which is considerably higher in energy than the closed region, is probably either a dynamical effect or an entropic effect; the open structure is far more flexible than the closed and thus fills a larger swath of configuration space. This range of motion is almost surely more constrained in a real material relative to the small cluster model employed in our study, and thus this open region is much less likely to be explored. As

such, the closed region (and its larger barrier to NR decay) are likely more relevant to the behavior of a more extended material. Note also that the CASSCF- and CASPT2-optimized structures are in reasonable agreement, supporting the accuracy of the CASSCF PES used in our AIMS simulations.

In order to further assess the accuracy of the CASSCF surface used in our simulations, we have performed single point CASPT2 calculations at the *spawning points* from the AIMS simulations, those points at which new nuclear basis functions are created. Scatter plots showing the correlation between the S_1-S_0 energy gaps and S_1 energies computed at these points at the CASSCF and MS-CASPT2 theoretical levels are presented Figure 8-7. The fact that the CASSCF and MS-CASPT2 values correspond reasonably well in both panels further validates the CASSCF PES, though a systematic bias towards higher S_1 energy at the CASSCF level can be observed in Figure 8-7(b).

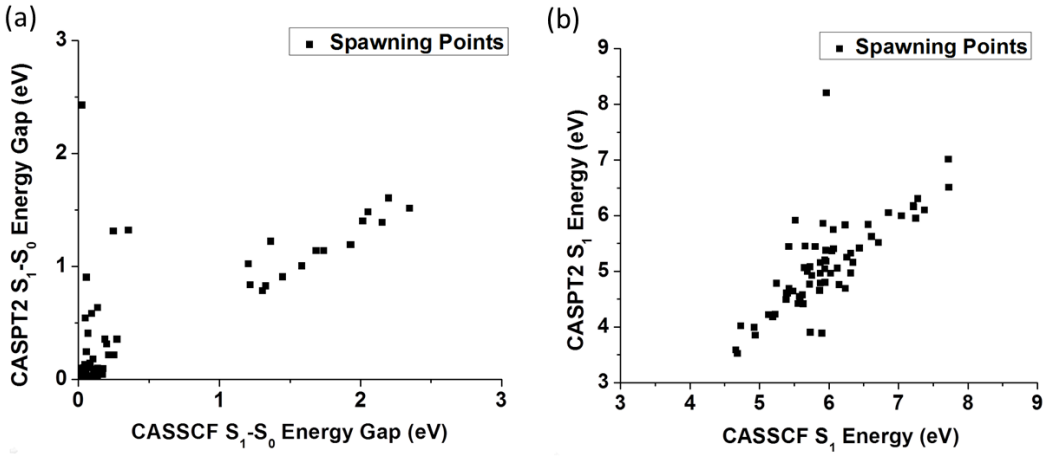


Figure 8-7. Scatter plot of the S_1-S_0 energy gap (a) and S_1 energy (b) of the 3O cluster at all spawning geometries computed at the CASSCF and MS-CASPT2 theoretical levels.

8.2.3 Dynamical simulations of the 5O cluster

We now turn our attention to AIMS dynamical simulations of the 5O cluster. These calculations predict extremely efficient non-radiative decay, with nearly 100% of the excited state population decaying to the ground state within 400 fs after excitation (Figure 8-8). As with the 3O cluster, the non-radiative decay process is far faster than the experimentally observed microsecond PL lifetime of SiNCs. Again the excited state population decay is accompanied by a rapid decrease in the S_1-S_0 energy gap (Figure 8-9). At the FC geometry, the vertical excitation energy is 6.70 eV at the CASSCF level of theory, but the average energy gap drops below 2 eV in the first 25 fs after excitation and oscillates and slowly decays towards zero thereafter. As for the 3O cluster, the fact that many trajectories explore regions of near-zero energy gap suggests the involvement of conical intersections.

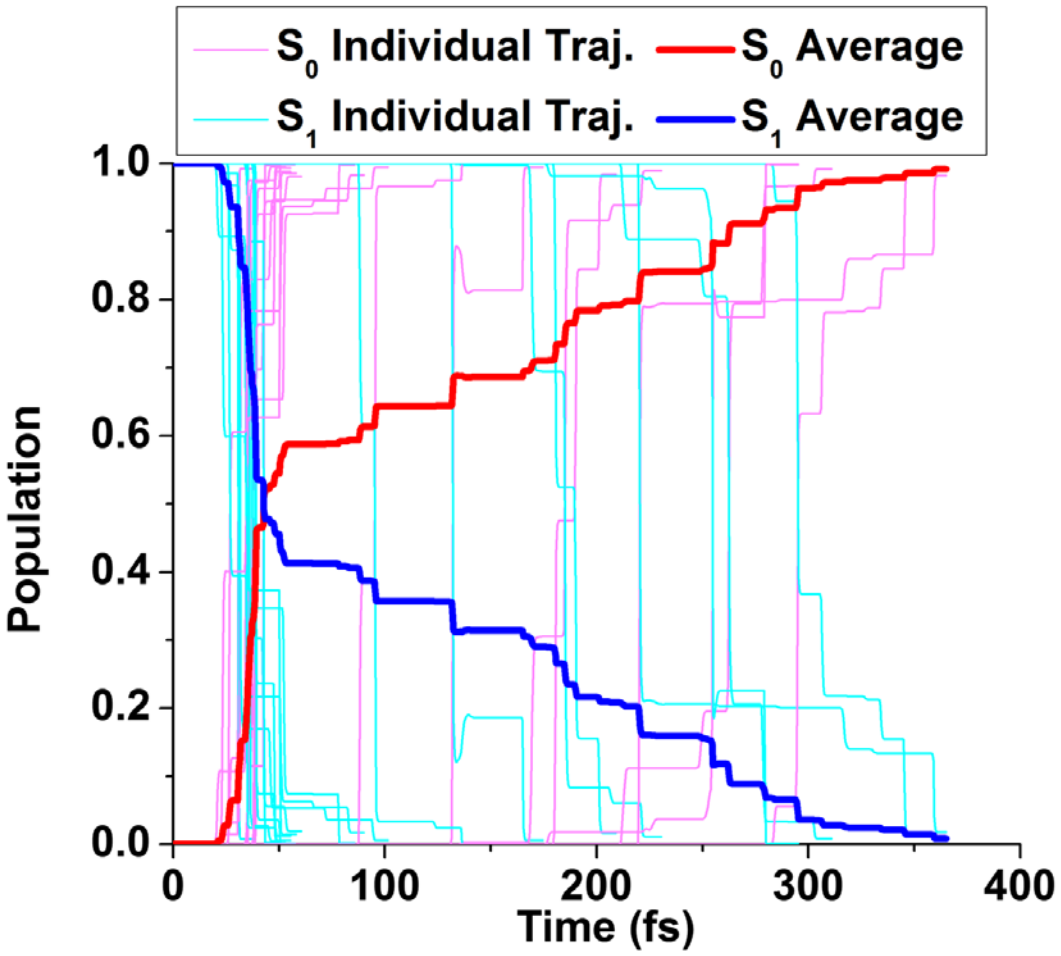


Figure 8-8. Populations of S_1 (bold blue) and S_0 (bold red) of the 5O cluster as a function of time after excitation averaged over all twenty simulations. Thin blue and pink lines report the population corresponding to the twenty individual simulations before averaging.

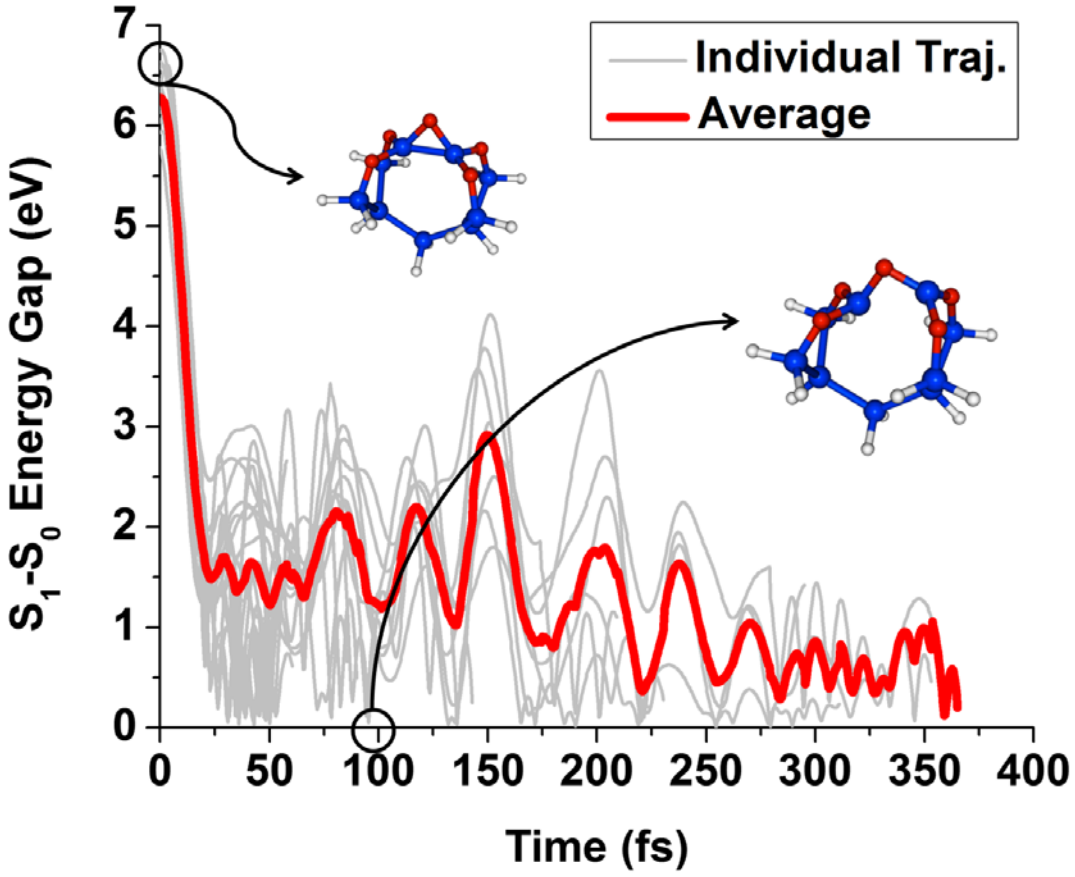


Figure 8-9. The average S_1-S_0 energy gap of the 5O cluster as a function of time after excitation (bold red) with the energy gaps of the twenty individual initial nuclear basis functions shown in gray. Snapshots from representative trajectories are shown in the inset.

Concurrent with the observed non-radiative decay are geometric changes, depicted in Figure 8-10. Unlike the 3O cluster, both Si-O bonds of the epoxide ring remain intact for the duration of our simulations of the 5O cluster. Thus the simulations only explore closed structures which may realistically be accessed under more constrained conditions. Still, the cluster undergoes a noticeable distortion of the O-SiOO pyramidalization angles. In Figure 8-10(a) we present averages of the two O-SiOO pyramidalization angles as a function of time. These two angles are indistinguishable by symmetry, so for each trajectory we sort the two angles by value. The average of the greater angle for each simulation is shown in red, and the average of the lesser is in blue. Both O-SiOO angles increase initially, indicating a weakening of the Si-Si bond of the epoxide ring, but at roughly 25 fs one of the angles returns towards its initial negative value while the other generally remains less negative or becomes positive.

We present the distance between the epoxide Si atoms as a function of time in Figure 8-10(b). The Si-Si bond weakens somewhat, as indicated by the increase in this distance upon excitation. However, unlike in the 3O cluster, the cluster never breaks open completely. Instead, coherent oscillations in the Si-Si distance are observed.

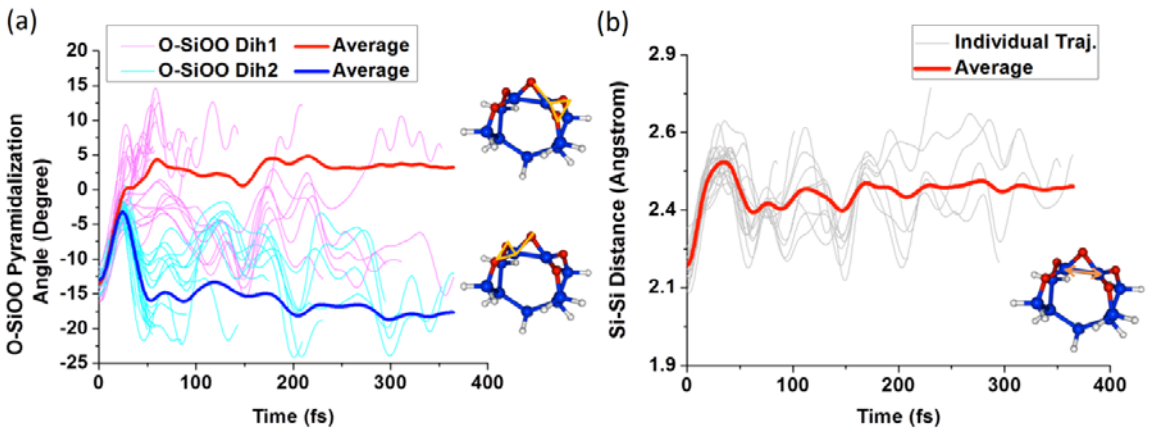


Figure 8-10. (a) Average O-SiOO pyramidalization angle as a function of time (bold lines) with the angles of the twenty individual initial nuclear basis functions (thin lines). The greater angle for each trajectory is shown in red, while the lesser is shown in blue. The inset indicates the atoms that define the pyramidalization angle. (b) The average distance between the epoxide silicon atoms as a function of time (bold red) with the values for the twenty individual initial nuclear basis functions in thin gray lines.

8.2.4 Potential energy surface analysis of the 5O cluster

Here we consider the role that conical intersections may play in the non-radiative decay of the 5O cluster. As for the 3O cluster, we applied geometry optimization at both the CASSCF and MS-CASPT2 levels of theory to identify important points on the PES. Two minimal energy S_1-S_0 conical intersections and one S_1 minimum were found at both theoretical levels. Single point energy calculations were performed at the CASSCF, MS-CASPT2, and MRCI(Q) theoretical levels at the geometries optimized at both levels of theory, with results shown in Table 8-3 and Figure 8-11. Two different regions of the S_1 PES were identified. We term these the stretched (Str) region, in which one of the Si-Si

bonds outside of the defect is significantly stretched, and the unpyramidalized (Unpyr) region, characterized by a near-zero pyramidalization angle about one of the epoxide silicon atoms. An MECI has been identified in each region, while a separate excited state minimum has been identified in the stretched region only. No stable minimum was identified in the unpyramidalized region. We have confirmed that this MECI is peaked by analysis of the surrounding PES, and thus no such minimum should be expected.

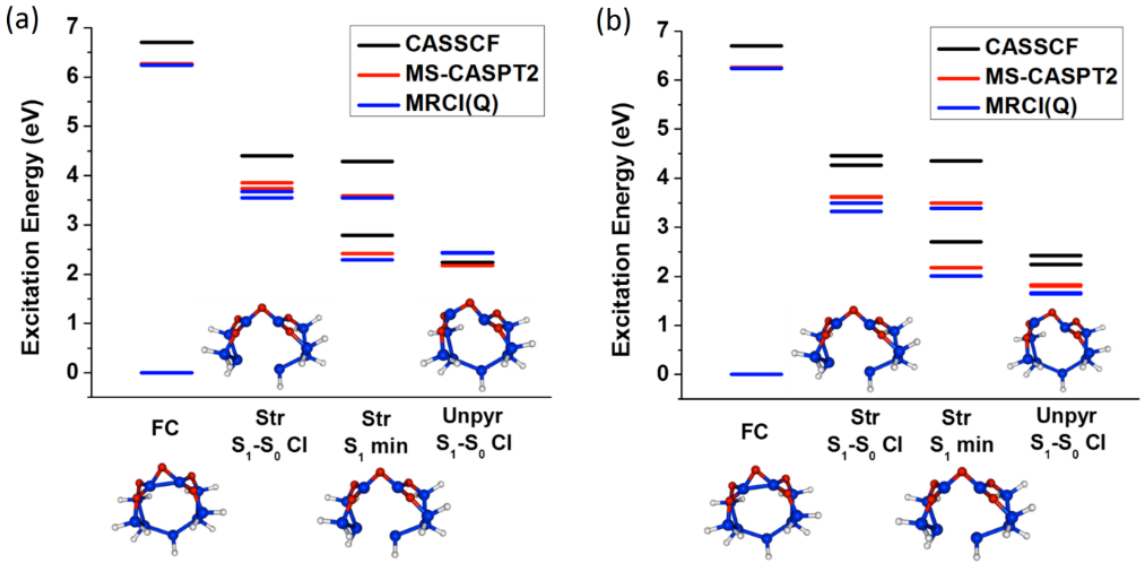


Figure 8-11. (a) The S_0 and S_1 energies of the CASSCF-optimized excited state minimum and S_1-S_0 conical intersections of the 5O cluster as calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory (black, red, and blue bars, respectively). Geometries are shown as insets. Energies at the Franck–Condon (FC) point are shown for comparison. (b) The same as (a) but excited state geometries are optimized at the MS-CASPT2 level of theory.

Table 8-3. The S_0 and S_1 energies (in eV) at four important points on the PES of the 5O cluster calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory. S_1 minimum and S_1-S_0 MECI geometries have been optimized at both the CASSCF and MS-CASPT2 levels of theory.

Geometry	Optimization Method	CASSCF S_0 Energy	CASSCF S_1 Energy	MS-CASPT2 S_0 Energy	MS-CASPT2 S_1 Energy	MRCI(Q) S_0 Energy	MRCI(Q) S_1 Energy
FC point	MP2	0.00	6.70	0.00	6.26	0.00	6.24
Str S_1-S_0 CI	CASSCF	4.40	4.40	3.74	3.85	3.54	3.68
	MS-CASPT2	4.27	4.46	3.61	3.63	3.32	3.50
Str S_1 min	CASSCF	2.79	4.28	2.42	3.58	2.29	3.55
	MS-CASPT2	2.70	4.36	2.18	3.49	2.01	3.39
Unpyr	CASSCF	2.24	2.24	2.17	2.43	2.24	2.44
S_1-S_0 CI	MS-CASPT2	2.24	2.42	1.81	1.82	1.64	1.67

First we will consider the stretched region of the PES. The CASSCF optimized Str S_1-S_0 MECI is found to be 4.40 eV above the ground state minimum energy. Though this intersection is energetically accessible upon excitation of the present cluster, it is far above the visible region and unlikely to influence the visible light photochemistry of SiNCs. Single point calculations at the MS-CASPT2 and MRCI(Q) theoretical levels predict the CASSCF optimized Str S_1-S_0 MECI to be slightly lower in energy—3.85 eV and 3.68eV above the ground state minimum, respectively—but still inaccessible at visible energies. Note that MS-CASPT2 and MRCI(Q) predict small S_1-S_0 energy gaps (0.11eV and 0.14eV) at the CASSCF-optimized MECI, indicating that the CASSCF surface used in our dynamic simulations is providing a qualitatively accurate prediction of the geometry of this intersection. Reoptimization at the MS-CASPT2 level yields similar geometries with similar energies at all three levels of theory, confirming the accuracy of the CASSCF description of this intersection.

A separate S_1 minimum exists in the stretched region of the PES. The CASSCF-optimized Str S_1 min is found to be 4.28 eV above the ground state minimum, just 0.12 eV below the MECI. More accurate MS-CASPT2 and MRCI(Q) calculations predict this feature to be 3.58 and 3.55 eV above the ground state, respectively. At these higher levels of theory, the activation barrier to reach the MECI from the minimum is slightly higher: 0.27 eV and 0.13 eV, respectively. Reoptimization at the MS-CASPT2 level of theory yields similarly small energy differences between the S_1 min and MECI structures: 0.10, 0.14, and 0.11 eV at the CASSCF, MS-CASPT2, and MRCI(Q) levels, respectively.

These small energy differences are consistent with fast and efficient NR decay in this region of the PES.

The PONOs of the CASSCF wavefunction at three points of interest on the PES are shown in Figure 8-12. At the FC point the PONOs are mixtures of the Si-Si σ bonding and antibonding orbitals of the epoxide group, with some oxygen nonbonding character (Figure 8-12 left). Note that the degenerate occupation numbers render the exact mixture of these orbitals meaningless; in fact the excitation is from a bonding HOMO to an antibonding LUMO. Though the excitation at the FC point is localized to the epoxide ring and surrounding oxygen atoms, the character of the electronic excited state in the stretched region of the PES is unique among the conical intersections we have studied in this and previous work in that the excitation is delocalized over the entire cluster. Specifically, the electronic transition facilitated by the Str S_1-S_0 MECI involves the transfer of an electron from a Si-Si σ antibonding orbital on the silicon backbone (center bottom) to a σ bonding orbital in the epoxide defect (center top). Analysis of the Mulliken charges at this geometry confirms that the transition involves charge transfer between the epoxide ring and the stretched bond (see Figure S3). The 3.39 Å Si-Si bond length of this stretched backbone bond (compared to a typical 2.35 Å bond in crystalline silicon) clearly indicates the weakening of this bond induced by the antibonding electron. The involvement of the electrons of the silicon backbone at this MECI suggests that our cluster may be too small to yield accurate energetics, as a larger cluster may allow further delocalization. The energy of this intersection may even be particle-size-dependent.

Further study is underway to assess whether interaction of charge transfer states in which only one charge carrier is localized to the defect with the ground state may be a common mechanism of NR decay in SiNCs, but given the high energy of the Str S_1-S_0 CI, the present work does not directly support this conclusion.

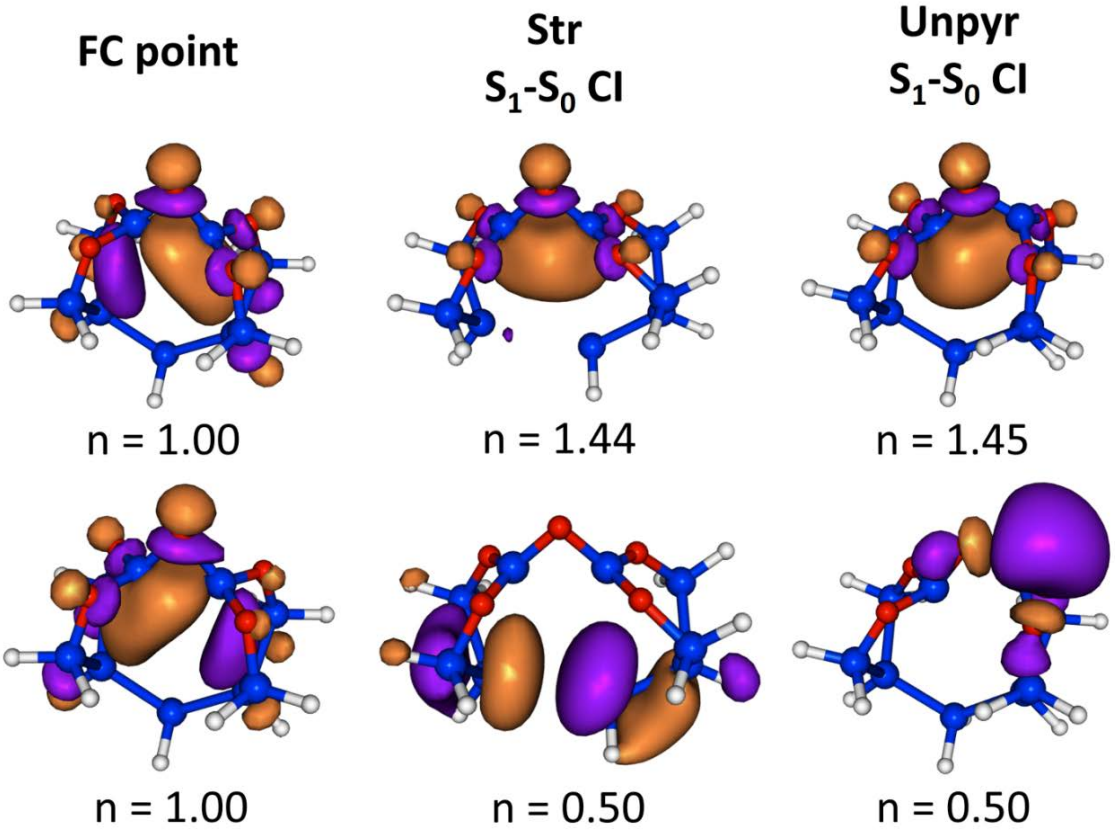


Figure 8-12. Partially occupied CASSCF natural orbitals of the 5O cluster at the FC point (left), Str S_1-S_0 MECI (middle), and Unpyr S_1-S_0 MECI (right) are presented with occupation numbers. The FC natural orbitals are computed by diagonalization of the S_1 density matrix, while those of the S_1-S_0 MECIs are computed using the state-averaged S_1-S_0 density matrix.

The MECI in the unpyramidalized region of the PES is considerably lower in energy than that in the stretched region. Optimized at the CASSCF level of theory, the Unpyr S_1-S_0 MECI is 2.24 eV above the ground state minimum energy. Single point calculations at the MS-CASPT2 and MRCI(Q) levels of theory predict this point to be 2.43 or 2.44 eV above the ground state energy, respectively. These calculations also show a small S_1-S_0 energy gap, confirming the likely existence of a conical intersection in this region of the PES. Reoptimization at the MS-CASPT2 level of theory lowers the energy to 1.82 eV. Single point MRCI(Q) calculations at the MS-CASPT2-optimized structure suggest a yet lower energy of 1.67eV. Thus, there should be little doubt that the energy of the unpyramidalized intersection is low enough to be accessed after excitation at visible energies.

The PONOs of the Unpyr S_1-S_0 CI (Figure 8-12 right) show that the electronic transition promoted by this intersection is localized to the epoxide ring itself; specifically, this transition involves the movement of an electron between the Si-Si bonding orbital of the epoxide (which interacts strongly with all five oxygen atoms; Figure 8-12, top right) and a non-bonding orbital on the unpyramidalized epoxide silicon atom (Figure 8-12, bottom right). Mulliken charge analysis indicates significant charge transfer between the epoxide Si atoms (Figure S3). The modest weakening of the Si-Si bond of the epoxide ring is evident in its 2.47 Å bond length.

As with the 3O cluster, we further assessed the accuracy of the CASSCF surface by performing MS-CASPT2 single point calculations at the spawning geometries from the

AIMS simulations of the 5O cluster and comparing the results to those at the CASSCF level of theory employed in the dynamic simulations. The results of this analysis are shown in Figure 8-13. The correlations between the MS-CASPT2 and CASSCF S_1-S_0 energy gaps (a) and the S_1 energies (b) at the spawning geometries indicate that the CASSCF PES is qualitatively accurate.

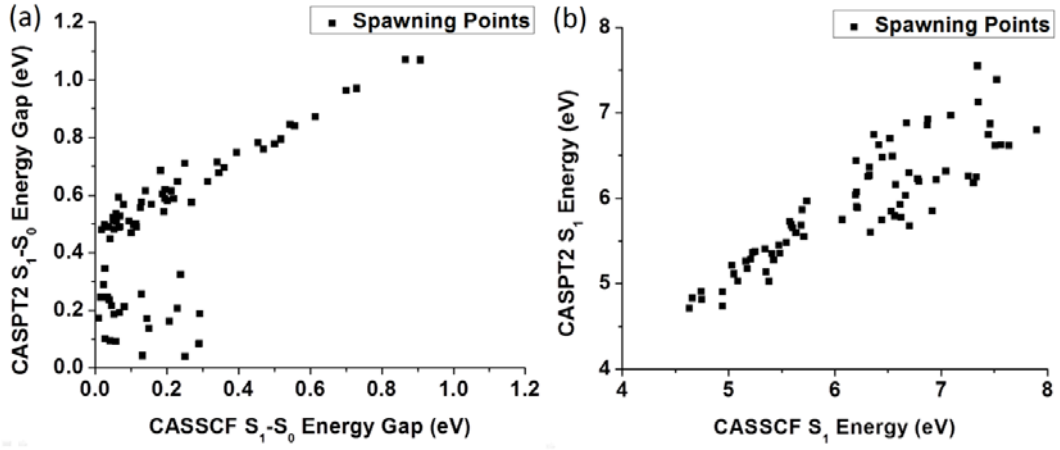


Figure 8-13. Scatter plot of the S_1-S_0 energy gap (a) and S_1 energy (b) of the 5O cluster at all spawning geometries computed at the CASSCF and MS-CASPT2 theoretical levels.

8.3 Discussion

It is worth commenting explicitly on how the present theoretical data should be interpreted in the context of a real SiNC, and how experiments can be designed to gain further insights into non-radiative processes resulting from conical intersections accessed by deformation of oxygen-containing defects on the silicon surface. One can easily infer

excited state lifetimes from Figures 8-2 and 8-8, but it is important to understand what these lifetimes mean. These are the lifetimes of excited defects, which almost surely do not correspond directly to the lifetime of excited SiNCs. Depending on the excitation wavelength, the excitation of a SiNC likely results in the creation of a quantum-confined exciton state. Before this excitation can decay via a defect-localized conical intersection, the excitation must localize to the defect. Given the large vertical excitation energies of the epoxide defects presented here, simple Forster energy transfer is unlikely. Instead, energy transfer is likely an activated process which will require the reorganization of the nuclear structure of the defect (similar in principle to an electron transfer reaction[20]). The activation energy of such a process will depend on the energy of the exciton state, and thus on the size of the SiNC itself. Given the ultrafast non-radiative dynamics predicted here for an excited defect, the localization process is likely the rate limiting step, and thus the excited state lifetimes in this work do not directly correspond to the PL lifetimes of SiNCs. This is not to say that direct experimental investigation of the dynamics of excited defects is not possible. Ultrafast experiments which begin by direct excitation of the defect itself (which, in the present work and our past work on silicon epoxides[37] is predicted to require excitation energies in the UV regardless of particle size) could directly probe the dynamics of excited defects, and would help to elucidate the roles specific defects play in non-radiative recombination.

Assuming that the energies of defect-localized states are reasonably independent of particle size, the energies of the conical intersections reported in this work should

correspond well with those in true nanoscale systems (though work is underway to validate this assumption). As described in the introduction, an intersection whose energy is above the band gap of a particular SiNC is likely not to have any effect on the photochemistry, while an intersection found to be within the gap would likely result in complete non-radiative decay of the excitation prior to PL. As such, we attribute the size-insensitivity of the S-band PL of ensembles of oxidized SiNCs to the preferential quenching of smaller dots with larger gaps. By demonstrating the existence of additional intersections with energies of roughly 2 eV which could be present on the oxidized surface of SiNCs, the present work supports this hypothesis.

In addition to our hypothesis, two other hypotheses for the roots of the size insensitivity of the S-band PL energy were discussed in the introduction: 1) that oxidation introduces emissive defects, or 2) that oxidation renders the energy of the delocalized excited states of SiNCs size-insensitive. Though the presence of non-radiative defects does not exclude these other two possibilities, and it is quite possible that reality involves some mixture of all three mechanisms, the relative importance of our hypotheses can be tested experimentally. Specifically, we suggest measurement of the single-particle PL of individual hydrogen-terminated SiNCs both before and after oxidation. Our proposed mechanism would suggest that those hydrogen-terminated SiNCs with PL energies significantly above 2.1 eV would likely go dark upon oxidation, due to the introduction of conical intersections. In contrast, if oxidation introduces emissive defects or if it renders the energy of delocalized excited states size-insensitive somehow, one would expect that

such SiNCs would show red-shifted PL upon oxidation rather than no PL at all.

Lastly, it is worth discussing similarities and difference between the 3O and 5O defects studied here and the 1O epoxide defect we have previously investigated.[21] All three defects decay via a conical intersection which is accessed after a photochemical ring opening reaction. The 3O cluster is unique in that both the Si-Si and Si-O bonds of the epoxide ring break facilely, whereas only the Si-Si bond breaks readily in the 5O and 1O defects. Even when the 3O cluster remains closed, the preferred decay mechanism is via Si-O rather than Si-Si bond breaking. Regarding energetics, the conical intersection in the 1O cluster was predicted by MS-CASPT2 to be at 2.7 eV above the ground state minimum, compared to 1.9 eV for the closed MECI of the 3O cluster and 1.8 eV for the Unpyr MECI of the 5O cluster. Thus, a trend towards lower energy with increasing oxidation of the epoxide silicon atoms was observed.

8.4 Conclusions

In this work we have demonstrated that silicon epoxide defects containing different numbers of oxygen atoms exhibit conical intersections which may be responsible for non-radiative recombination pathways in oxidized SiNCs. These intersections are accessed upon photochemical opening of the silicon epoxide ring, and are accessible at visible energies (as low as 1.8 eV, according to the MS-CASPT2 level of theory). The presence of these intersections may explain both the decreased PL quantum yield upon oxidation and the red-shifted, size-insensitive nature of the PL of oxidized SiNCs relative to their

hydrogen-terminated counterparts. It is suggested that single particle PL measurements performed on individual nanoparticles both before and after oxidation could be used to test this hypothesis, directly discerning between the non-radiative decay mechanism proposed in this work and alternative hypotheses regarding the origin of the unusual PL of oxidized SiNCs, and we hope to see these experiments realized.

REFERENCES

REFERENCES

(1) Stefanov, B. B.; Gurevich, A. B.; Weldon, M. K.; Raghavachari, K.; Chabal, Y. J. *Phys Rev Lett* 1998, *81*, 3908.

(2) Ben-Nun, M.; Quenneville, J.; Martinez, T. J. *J Phys Chem A* 2000, *104*, 5161.

(3) Levine, B. G.; Coe, J. D.; Virshup, A. M.; Martinez, T. J. *Chem Phys* 2008, *347*, 3.

(4) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. *Chem Phys* 1980, *48*, 157.

(5) Wadt, W. R.; Hay, P. J. *J Chem Phys* 1985, *82*, 284.

(6) Stanton, J. F.; Bartlett, R. J. *J Chem Phys* 1993, *98*, 7029.

(7) Virshup, A. M.; Levine, B. G.; Martinez, T. J. *Theor Chem Acc* 2014, *133*.

(8) Thompson, A. L.; Punwong, C.; Martinez, T. J. *Chem Phys* 2010, *370*, 70.

(9) Finley, J.; Malmqvist, P. A.; Roos, B. O.; Serrano-Andres, L. Chem Phys Lett 1998, 288, 299.

(10) Langhoff, S. R.; Davidson, E. R. *Int J Quantum Chem* 1974, *8*, 61.

(11) Werner, H. J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schutz, M. Wires Comput Mol Sci 2012, 2, 242.

(12) Knowles, P. J.; Werner, H. J. *Chem Phys Lett* 1985, *115*, 259.

(13) Werner, H. J.; Knowles, P. J. *J Chem Phys* 1985, *82*, 5053.

(14) Knowles, P. J.; Werner, H. J. *Chem Phys Lett* 1988, *145*, 514.

(15) Werner, H. J.; Knowles, P. J. *J Chem Phys* 1988, *89*, 5803.

(16) Werner, H. J. *Mol Phys* 1996, *89*, 645.

(17) Celani, P.; Werner, H. J. *J Chem Phys* 2000, *112*, 5546.

(18) Korona, T.; Werner, H. J. *J Chem Phys* 2003, *118*, 3006.

(19) Levine, B. G.; Coe, J. D.; Martinez, T. J. *J Phys Chem B* 2008, *112*, 405.

(20) Marcus, R. A.; Sutin, N. *Biochim Biophys Acta* 1985, *811*, 265.

(21) Shu, Y. A.; Levine, B. G. *J Chem Phys* 2013, *139*.

Chapter 9

Do Excited Silicon-Oxygen Double Bonds Emit Light?

"Science is the belief in the ignorance of experts."
-Richard P. Feynman (1918-1988)

This chapter is reproduced from paper: Do excited silicon-oxygen double bonds emit light? Y. Shu, B. G. Levine, J. Phys. Chem. C, 118, 7669 (2014).

Studies based on tight binding, density functional theory (DFT), time-dependent DFT (TDDFT), and quantum Monte Carlo calculations suggest that excitations localized to individual silicon-oxygen double (Si=O) bonds may be responsible for the S-band PL of oxidized Si-NCs.[1-7] In this chapter, we apply ab initio molecular dynamics and multireference quantum chemistry to investigate the photodynamics of a silicon cluster containing a Si=O bond ($Si_7H_{10}O$; inset of Figure 9-1) to address two questions: What nuclear and electronic motions occur upon excitation of a Si=O bond and how might these motions affect the PL spectra of Si-NCs with Si=O defects on the surface?

9.1 Methods

9.1.1 Nonadiabatic molecular dynamics simulations

To investigate these questions, we have applied ab initio multiple spawning (AIMS), a hierarchy of approximations by which the full molecular (nuclear and electronic) time-

dependent Schrodinger equation of an electronically excited molecule can be solved.[8,9] The adiabatic electronic wavefunction associated with each nuclear basis function is calculated on-the-fly at the complete active space self-consistent field (CASSCF) level of theory.[10] A CASSCF active space of four electrons in six orbitals was used in our dynamical simulations, and a three-state average was employed to eliminate variational bias towards the ground state. The LANL2DZdp basis and effective core potentials (ECPs) were used.[11] These parameters were chosen in part because the excitation energies and state characters of the lowest electronic states of the $Si_7H_{10}O$ cluster computed at this level of theory are in good agreement with those predicted at the trustworthy equation-of-motion coupled cluster with single and double excitations[12] (EOM-CCSD) level of theory using the cc-pVTZ basis[13,14] (see online support information materials Tables S1-S4 http://pubs.acs.org/doi/suppl/10.1021/jp500013g).

All dynamical simulations were initiated on the lowest singlet excited state (S_1) despite the fact that the transition dipole moment between this state and the ground state (S_0) is zero by symmetry in the ground state equilibrium configuration of $Si_7H_{10}O$. This choice of initial conditions is based on the knowledge that excited electrons in a nanoparticle often cool to the lowest excited electronic state on sub-picosecond timescales.[15] Twenty initial nuclear basis functions were chosen with their initial average positions and momenta randomly sampled from the ground state vibrational Wigner distribution calculated in the harmonic approximation at the B3LYP[16,17]/LANL2DZ level of theory. Except where noted, all reported observables were calculated by incoherently averaging over the wavefunctions arising from these twenty initial wave packets. The

nuclear wavefunction is propagated using a mixed timestep approach[9] with the classical and quantum degrees of freedom integrated by adaptive velocity Verlet and adaptive fourth-order Runge-Kutta schemes, respectively. A time step of 20 au (~0.5 fs) was employed. Simulation data was collected for 1 ps after excitation. The widths of the frozen Gaussian basis functions were chosen to be 25.0, 6.0, and 30.0 $Bohr^2$ for silicon, hydrogen, and oxygen atoms, respectively.

9.1.2 Ab-Initio calculations

To further analyze the PES and assess the accuracy of the CASSCF method employed in the AIMS calculations, additional single point computations and geometry optimizations were performed at the multi-state complete active space second-order perturbation (MS-CASPT2) level of theory[18] to add dynamic correlation to the CASSCF reference. The same active space and basis set as used in the dynamical simulations was employed. To accelerate optimization, the 16 lowest energy orbitals were not correlated in these calculations. Conical intersections were optimized using a multiplier-penalty approach[19] which does not require the calculation of nonadiabatic coupling vectors. Additional single point calculations were performance at the multireference configuration interaction level of theory with single and double excitations and Davidson correction[20] (MRCI(Q)) with the same active space and basis set. All dynamical calculations were performed using the FMSMolPro package.[9] Geometry and conical intersection optimizations at the CASSCF and MS-CASPT2 levels of theory were done with MolPro[21] in conjunction with CIOpt.[19] EOM-CCSD calculations were performed in MolPro as well. All computed absolute energies are presented in Tables S1-S4, and all optimized

geometries are presented in Supporting Information as well.

9.2 Results and Discussion

9.2.1 Dynamical simulations

The dynamical simulations predict that 7% of the excited population decays to S_0 within one picosecond after excitation, as can be seen in Figure 9-2 which presents the populations of the various adiabatic electronic states as a function of time. Assuming exponential decay, this rate corresponds to a 9.5 ps excited state lifetime, which is orders of magnitude shorter than the microsecond radiative lifetime of the S-band PL of silicon nanoparticles. If a similar rate of decay were observed in nanoscale systems containing Si=O defects, essentially no PL would be observed.

The onset of non-radiative decay is accompanied by a dramatic decrease in the S_1-S_0 energy gap (Figure 9-1). At the ground state minimum structure, the CASSCF level of theory predicts a 3.0 eV vertical excitation energy (comparable to the value predicted at the EOM-CCSD/cc-pVTZ level of theory: 2.9 eV), but the average energy gap falls below 1.0 eV within the first 55 fs after excitation and oscillates around this value for the remainder of the simulation, with the exception of recurrences at 400 and 800 fs which result from vibrational motions to be discussed below. The fact that individual trajectories approach zero energy gap several times during the simulation suggests the existence of a conical intersection between S_0 and S_1 facilitating the non-radiative decay observed in Figure 9-2, and this possibility will be explored below. A histogram of the S_1-S_0 energy gap integrated over the excited state population of the dynamical simulations is

presented in Figure 9-3. A broad peak in the energy gap is observed between 0.9 and 1.1 eV, which is considerably smaller than the 1.5-2.1 eV S-band PL energy, though this discrepancy is within the range of possible error for energy gaps computed at the CASSCF level. Correlated calculations will be reported below to validate the gap.

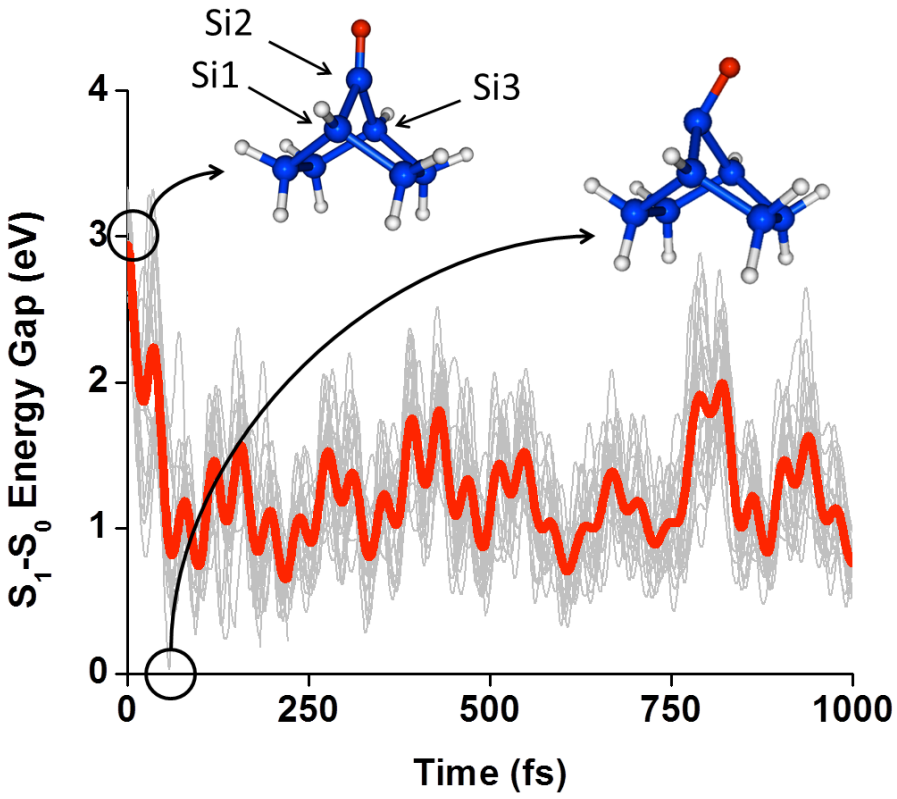


Figure 9-1. The average S_1-S_0 energy gap as a function of time (bold red) with the energy gaps of the twenty initial nuclear basis functions (gray). Snapshots from representative trajectories are shown in the inset, with silicon, oxygen, and hydrogen atoms colored blue, red, and white, respectively.

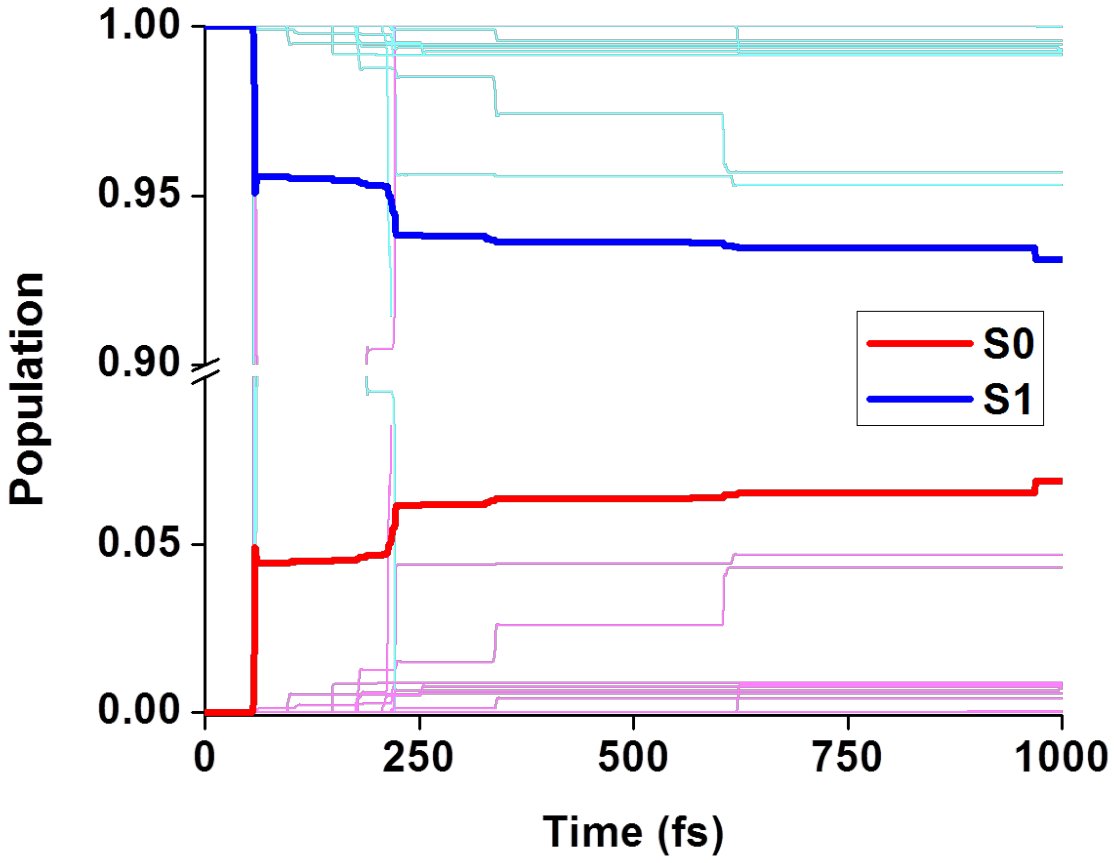


Figure 9-2. The population of S_1 (bold blue) and S_0 (bold red) as a function of time after excitation averaged over all twenty initial nuclear basis functions. Thinner blue and pink lines reflect the population corresponding to the twenty individual basis functions before averaging.

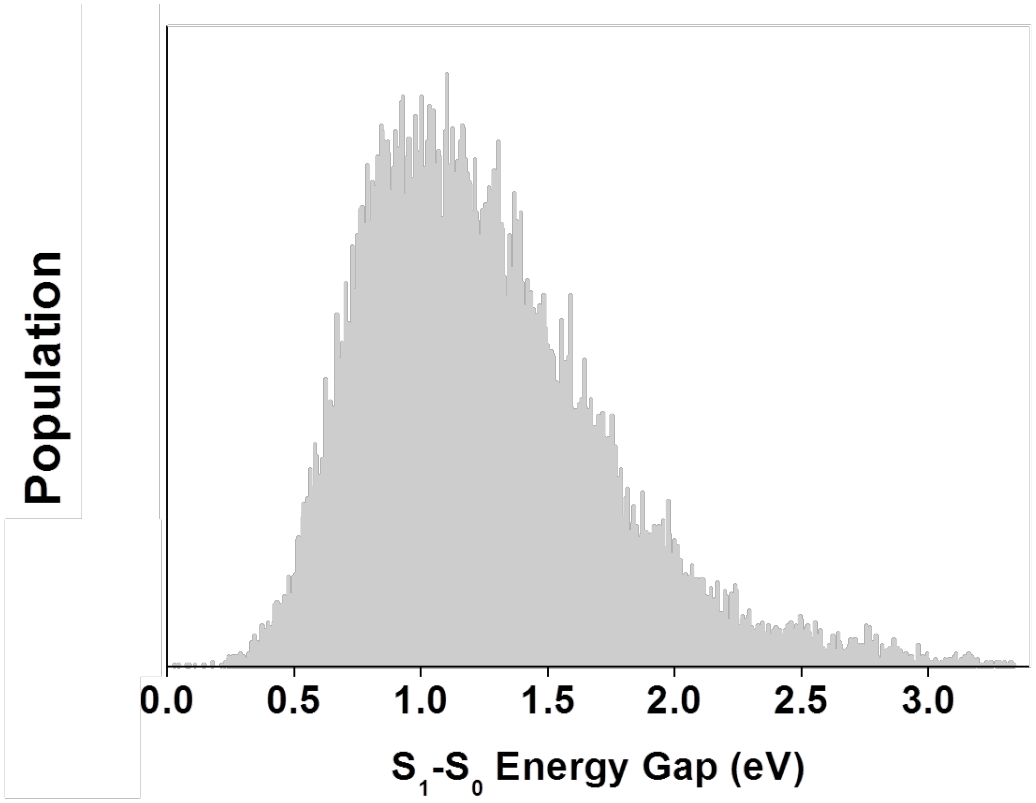


Figure 9-3. A histogram of the S_1-S_0 energy gap of the excited population collected over all twenty initial nuclear basis functions integrated over the duration of the simulations.

The decay of population to the ground state occurs after a significant distortion of the structure of the cluster. At the ground state minimum structure, the double-bonded silicon atom (Si2; see inset of Figure 9-1 for atom labels) is in a planar sp^2 configuration. As the energy gap decreases, the atoms surrounding Si2 take on a pyramidal sp^3 geometry, as can be seen in the inset of Figure 9-1. Figure 9-4 presents the time evolution of the pyramidalization angle about Si2 for the twenty initial basis functions, where this pyramidalization angle is defined as

$$\theta = \left| \arcsin \frac{\vec{C} \bullet \vec{R}_{\mathrm{Si2O}}}{\left\| \vec{C} \right\| \left\| \vec{R}_{\mathrm{Si2O}} \right\|} \right| \tag{8-1}$$

with $\vec{C} = \vec{R}_{\mathrm{Si2Si1}} \times \vec{R}_{\mathrm{Si2Si3}}$, $\vec{R}_{\mathrm{AB}}$ indicating the vector from atom A to atom B, and $\bullet$ and $\times$ indicate the scalar and vector products of two vectors, respectively. A pyramidalization angle of zero indicates a planar sp^2 configuration. Note that atoms Si1 and Si3 are symmetry equivalent, and the definition of the pyramidalization angle does not depend on how these labels are assigned.

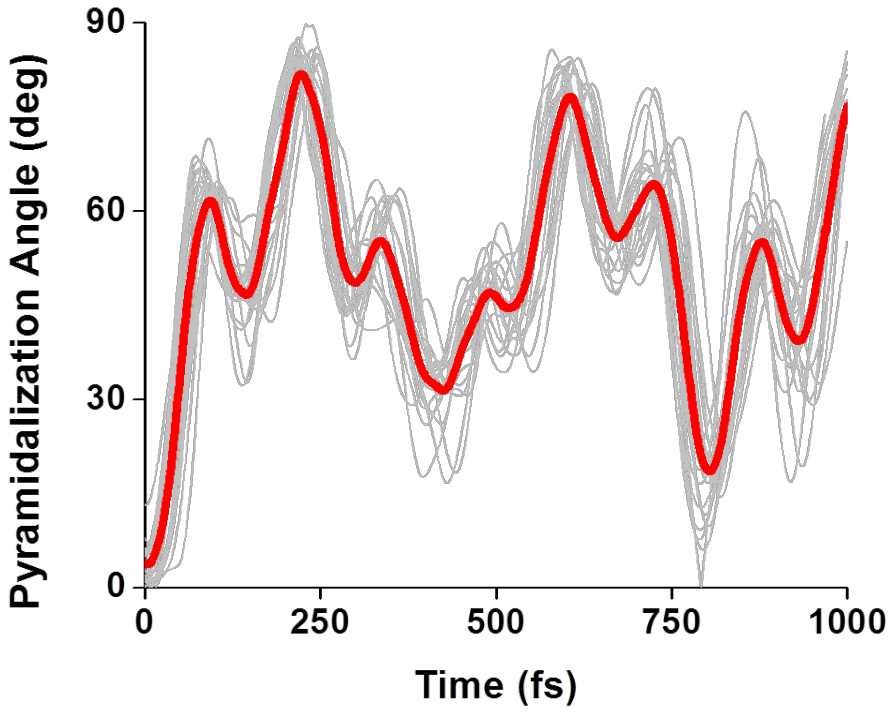


Figure 9-4. The pyramidalization angle (defined in the text) about the double-bonded Si atom as a function of time, with the average over all excited population shown in red and the values corresponding to the twenty individual initial nuclear basis functions shown in gray.

All initial basis functions quickly become strongly pyramidalized, with the average pyramidalization angle rising from an initial value of 4º to a maximum of 82º at 220 fs after excitation. The initial drop in the S_1-S_0 energy gap occurs during this initial approach to the pyramidalized geometry, as does the majority of the observed non-radiative decay. All twenty initial nuclear basis functions exhibit nearly synchronous motion along this coordinate for the duration of the simulations. This coherent motion has a period of approximately 400 fs; the return of the pyramidalization angle to near zero at 400 and 800 fs after excitation corresponds well with the observed recurrences in the S_1-S_0 energy gap discussed above. Given the strong correlation between the pyramidalization angle and the S_1-S_0 energy gap, these oscillations could provide a signature by which to detect the photodynamics of Si=O defects in transient absorption experiments.

The strong correlation between the pyramidalization angle and S_1-S_0 energy gap are consistent with a previous theoretical study reporting pyramidalization of Si=O defects in excited Si-NCs.[4] It is also interesting to consider the present results in the context of a recent nonadiabatic molecular dynamics study of silaethylene, SiH_2=CH_2, which contains a similar double bond between a silicon atom and a more electronegative second row element.[22] After photoexcitation, pyramidalization about both the double-bonded silicon and carbon atoms was observed 20-30% of the time. This pyramidalization leads to internal conversion to the ground state within tens of femtoseconds. Though non-radiative decay is not as fast in the present case, when taken together these results suggest that this may be a general behavior of excited double bonds between silicon and second row elements. The pyramidalized conical intersection is also a common motif in

organic photochemistry (see for example refs.[23-26]).

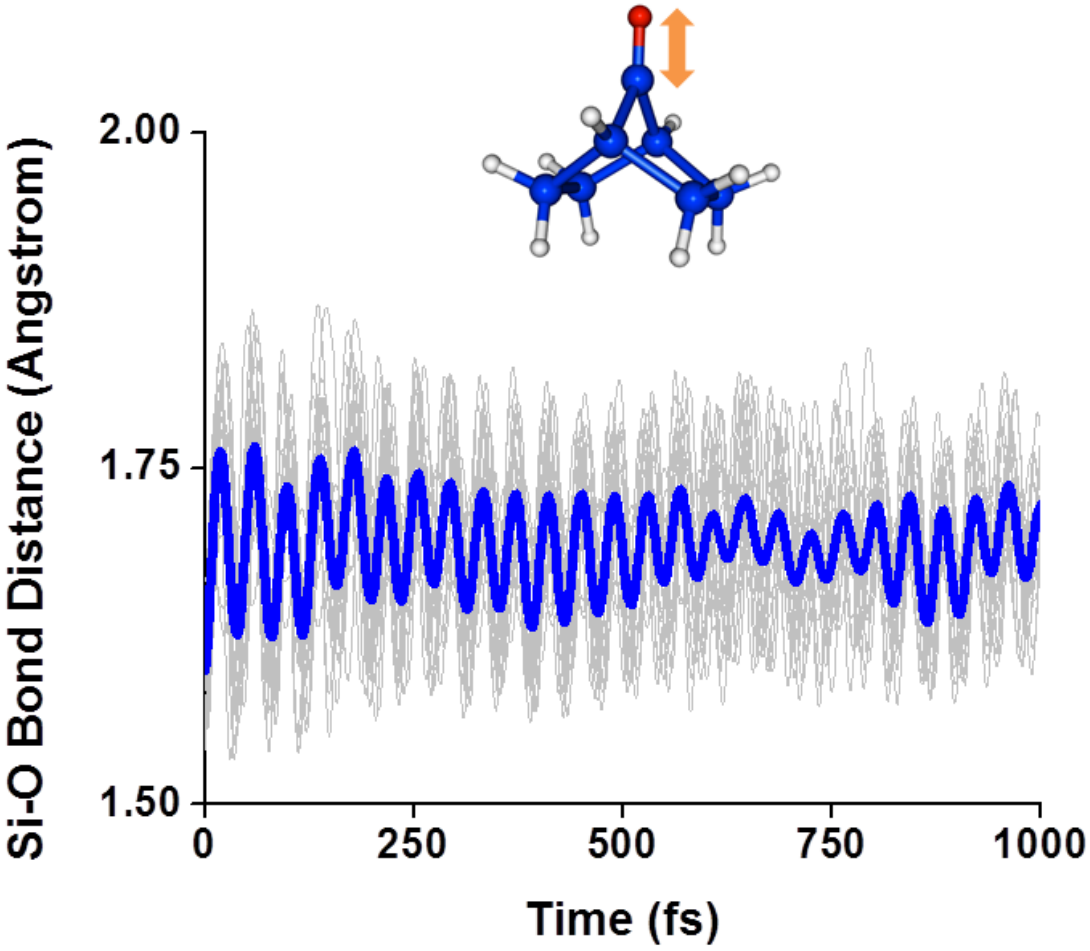


Figure 9-5. The Si=O bond length as a function of time, with the average in blue and the twenty individual initial nuclear basis functions in gray. Coherent oscillations in the bond length are observed for the entire simulation window.

The higher frequency oscillations observed in the energy gap correspond to coherent oscillations in the Si=O bond distance (Figure 9-5). These oscillations span 0.15 Å and have a period of 40 fs. Though the dependence of the energy gap on the Si=O

distance is weaker than the dependence on the pyramidalization angle, these oscillations may still serve as a signature of Si=O defects in time-dependent spectroscopic experiments. In addition, such a large excited state displacement of this totally symmetric mode would likely be visible as a strong resonance Raman peak at roughly 1058 cm^{-1} (the Si=O stretch frequency computed at the B3LYP/LANL2DZ level of theory).

9.2.2 Analysis of the potential energy surface

The fact that non-radiative decay is observed when several nuclear basis functions on S_1 approach regions of near-zero energy gap with S_0 suggests the possibility of non-radiative decay by conical intersection. To investigate this possibility we have applied geometry and conical intersection optimization techniques to explore the CASSCF PESs, the results of which are reported in Figure 9-6 and Table 9-1. The energies of the CASSCF optimized structures computed at the more trustworthy MS-CASPT2 and MRCI(Q) levels of theory are reported as well. A minimal energy conical intersection between S_0 and S_1 (S_1-S_0 CI) and a stable excited state minimum structure (S_1 min) were found in the pyramidalized region of the PES explored by the dynamic simulations. Both of these points are energetically accessible after vertical excitation to S_1 at the Franck-Condon (FC) point. As pictured, these geometries are characterized by pyramidalization about the double bonded silicon atom.

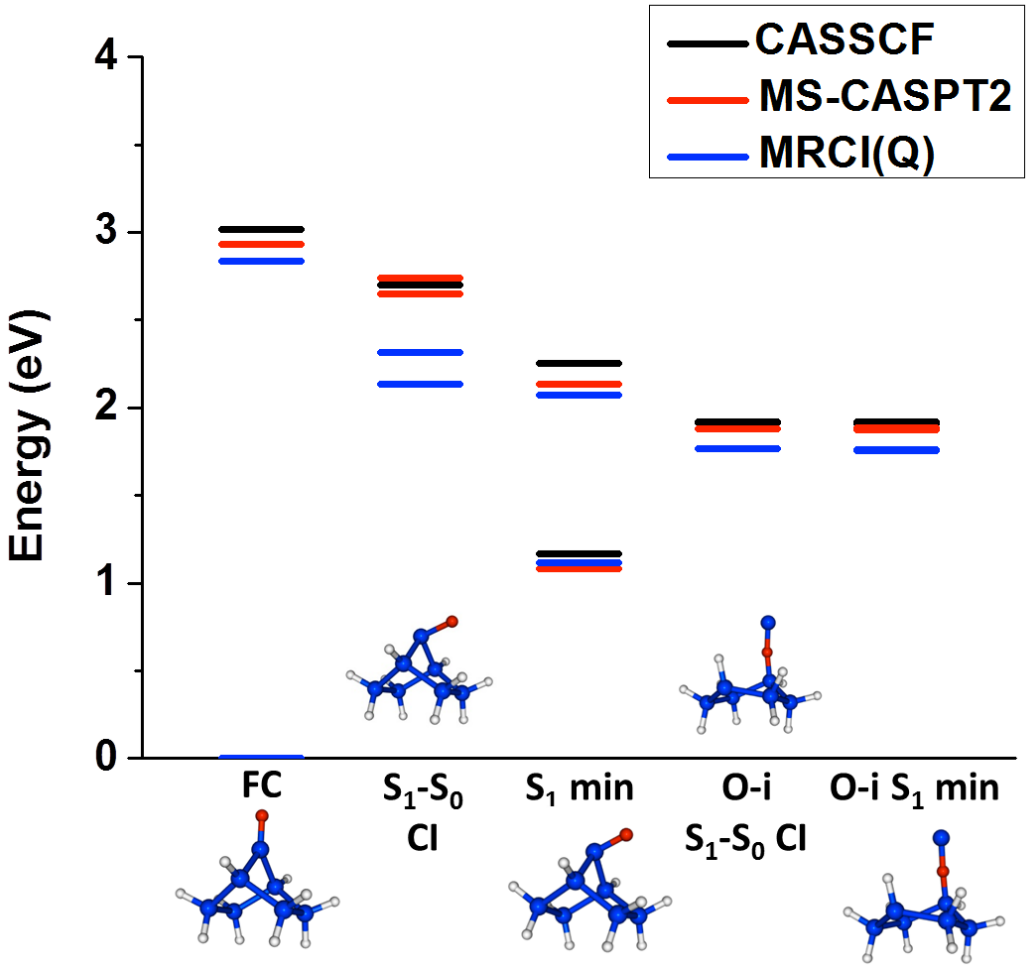


Figure 9-6. The S_0 and S_1 energies of the CASSCF-optimized excited state minima and minimal energy conical intersections as calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory (black, red, and blue bars, respectively). Geometries are shown as insets. Energies at the Franck-Condon (FC) point are shown for comparison. The S_1-S_0 CI and S_1 min correspond to the pyramidalized region of the PES explored by the AIMS simulations. The oxygen insertion (O-i) intersection and minimum were not observed in the dynamical simulations, but are at lower energies than the pyramidalized minimum.

Table 9-1. The S_0 and S_1 energies (in eV) at five important points on the PES of the Si=O bond containing cluster calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory. Excited state geometries are optimized at both the CASSCF and MS-CASPT2 levels of theory. The pyramidalized geometries are referred to generically as S_1-S_0 CI and S_1 min, while additional features in the oxygen insertion region of the PES are prefixed by O-i.

Geometry	Optimization Method	CASSCF S_0 Energy	CASSCF S_1 Energy	MS-CASPT2 S_0 Energy	MS-CASPT2 S_1 Energy	MRCI(Q) S_0 Energy	MRCI(Q) S_1 Energy
FC point	B3LYP	0	3.02	0	2.93	0	2.84
S_1-S_0 CI	CASSCF	2.70	2.70	2.65	2.74	2.13	2.32
	MS-CASPT2	2.62	2.80	2.01	2.03	2.02	2.13
S_1 min	CASSCF	1.17	2.25	1.08	2.14	1.12	2.07
	MS-CASPT2	1.19	2.31	0.81	1.94	0.86	1.86
O-i S_1-S_0 CI	CASSCF	1.91	1.92	1.88	1.88	1.76	1.77
	MS-CASPT2	1.96	1.97	1.81	1.82	1.62	1.63
O-i S_1 min	CASSCF	1.91	1.92	1.87	1.88	1.75	1.76
	MS-CASPT2	1.96	1.97	1.81	1.82	1.62	1.63

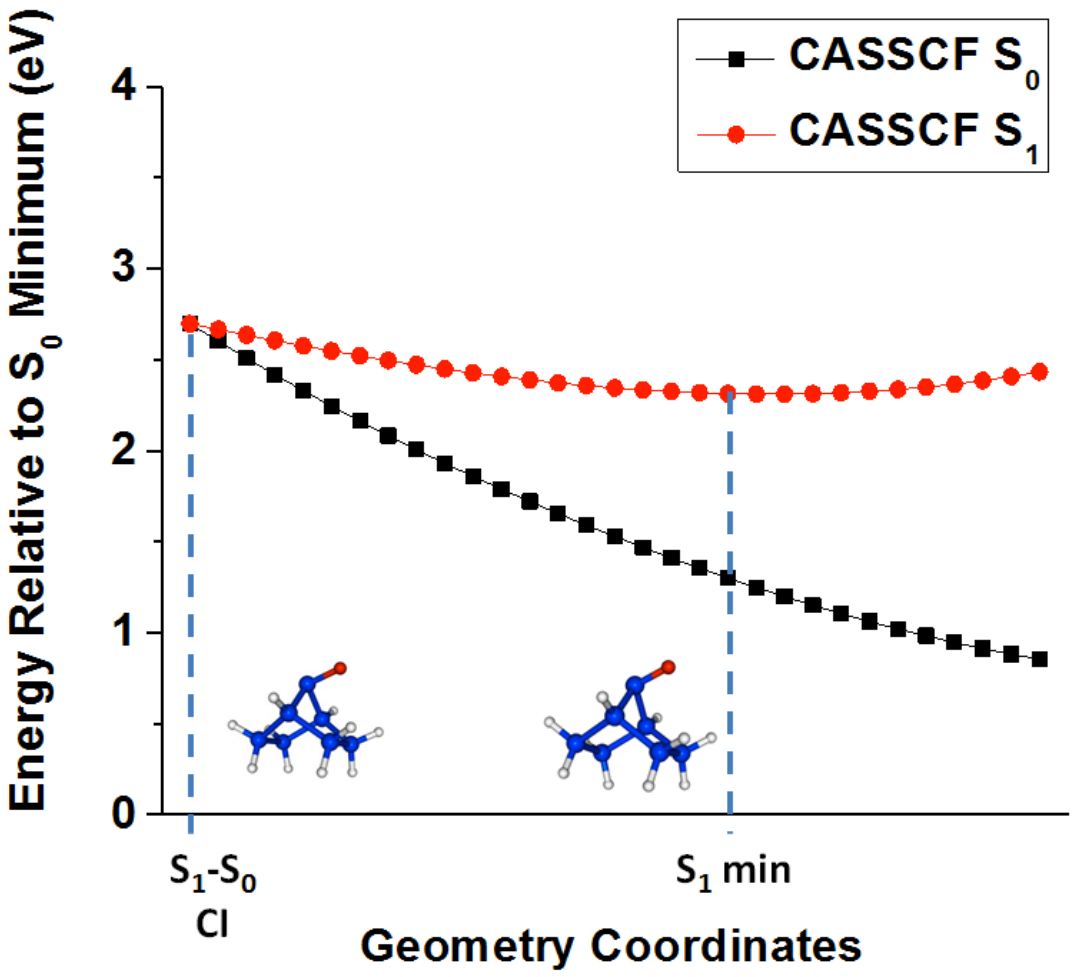


Figure 9-7. The CASSCF S_0 (black) and S_1 (red) energies as a function of a linearly interpolated/extrapolated coordinate between the S_1-S_0 CI and S_1 min geometries.

The S_1 min and S_1-S_0 CI are 2.25 and 2.70 eV above the ground state minimum energy, respectively. The 0.45 eV energy difference between these points can be roughly thought of as an activation energy for internal conversion. This relatively large activation energy is likely the reason that only 7% of the population is observed to decay to S_0 in the 1 ps simulation window. A PES along a linearly interpolated/extrapolated coordinate between the S_1-S_0 CI and the S_1 minimum is presented in Figure 9-7. This PES

demonstrates that no barrier exists between the minimum and the intersection. The fact that it is very rare for trajectories to actually reach zero energy gap in our simulations (see Figure 9-1) suggests that much of the nonadiabatic transfer occurs at lower energy geometries with larger gaps.

Single point calculations at the MS-CASPT2 and MRCI(Q) levels of theory predict small energy gaps (0.09 eV and 0.19 eV, respectively) at the CASSCF-optimized S_1-S_0 CI geometry, suggesting that the CASSCF description of the intersection seam is qualitatively accurate. However, these methods disagree significantly on the activation energy to reach the intersection from the S_1 minimum, with MS-CASPT2 predicting a 0.60 eV energy barrier and MRCI(Q) predicting a 0.25 eV barrier. Such discrepancies have often been observed,[27,28] and here we address them and further analyze the accuracy of the CASSCF surface by reoptimizing these structures at the MS-CASPT2 level of theory. Energies are reported in Figure 9-8 and Table 9-1 with the geometries represented as insets in Figure 9-8. The resulting S_1 minimum and minimal energy conical intersection have significantly lower energies: 1.94 and 2.03 eV above the ground state minimum compared to 2.25 and 2.70 eV at the CASSCF level of theory, respectively. Not only is the energy needed to access these points reduced, the relative energy is also reduced from 0.45 to 0.09 eV, suggesting that the CASSCF surface used in our dynamic simulations may significantly overestimate the activation energy required for internal conversion. Note also that the S_1-S_0 energy gap at the S_0 minimum—which is 1.08 eV at the CASSCF level of theory and is in good agreement with the maximum of the energy gap distribution presented in Figure

9-3—increases negligibly to 1.13 eV at the MS-CASPT2 level of theory. Recalculation at the MRCI(Q) level of theory predicts negligible (<0.1 eV) changes in the S_1-S_0 gap at the CASSCF- and MS-CASPT2-optimized minima relative to the CASSCF and MS-CASPT2 values, as well.

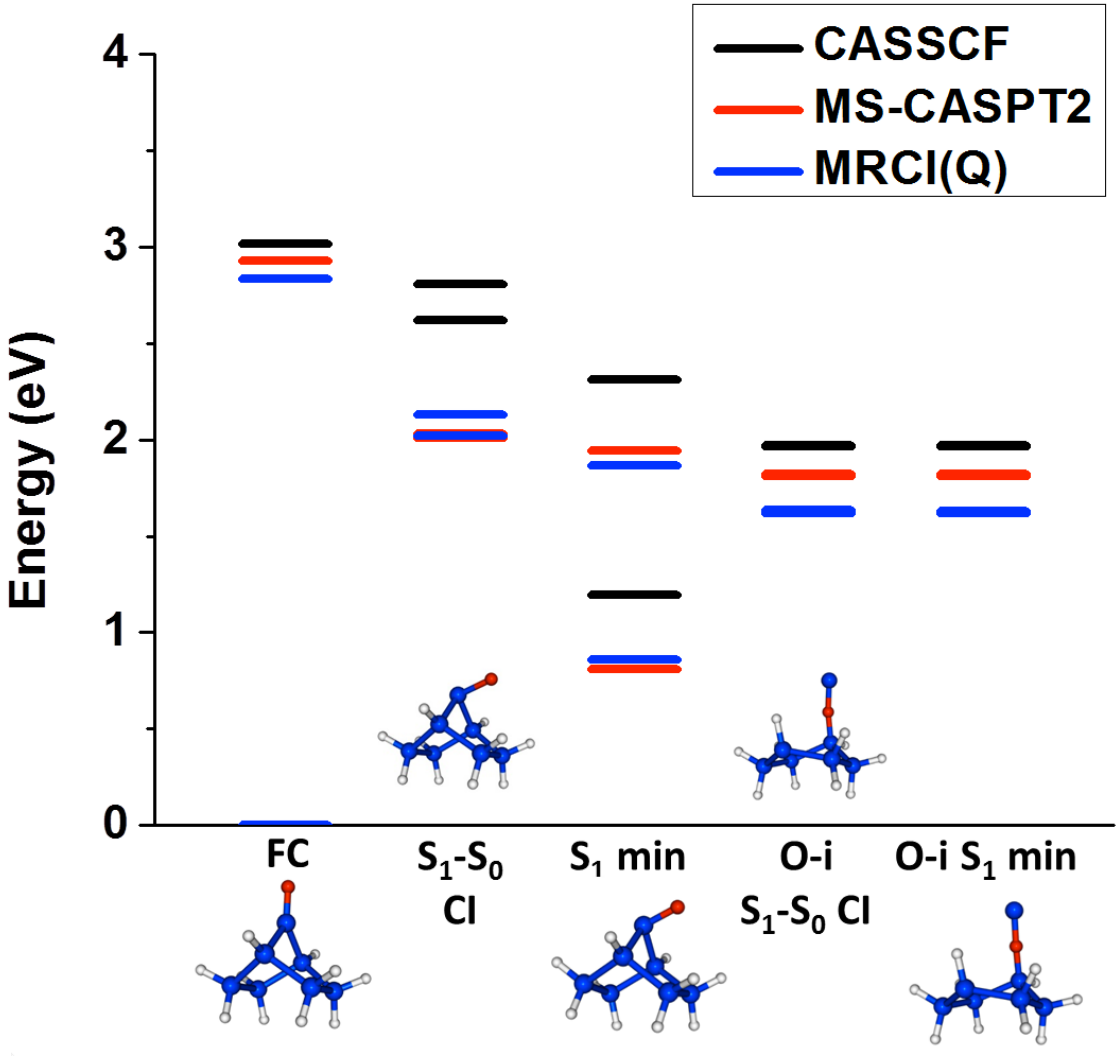


Figure 9-8. The S_0 and S_1 energies of MS-CASPT2 optimized minima and intersections, as calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory (black, red, and blue, respectively). Geometries are shown as insets.

It is noteworthy that the presence of conical intersections at roughly 2.25 eV above the ground state energy is consistent with the temperature and energy dependence of silicon PL. At energies below 2.1 eV, the PL lifetime of Si-NCs has been observed to vary with PL energy (*E*) according to $\tau \propto \exp(-E / 0.25\ \mathrm{eV})$.[29,30] However, at higher energies, NR decay becomes faster than predicted by this proportionality, suggesting a new NR decay mechanism becomes accessible at these energies.[30] The temperature dependence of the PL intensity, which becomes stronger as the PL-energy increases towards 2.0 eV, also points to an activated NR decay pathway with an energy in this range.[31]

Partially occupied natural orbitals at the FC point and the S_1-S_0 CI are presented in Figure 9-9. All active orbitals at all optimized geometries are pictured in Supporting Information. The natural orbitals are determined by diagonalization of the S_1 and state-averaged S_1-S_0 density matrices at the FC point and S_1-S_0 CI geometry, respectively. At the FC point, the transition to S_1 is characterized by an n→π* excitation from an oxygen lone pair orbital to an antibonding superposition of p orbitals on Si2 and the oxygen atom. Upon relaxation to the S_1-S_0 CI geometry the orbitals retain similar character, except that the pyramidalization of Si2 results in the hybridization of the Si2 p component of the antibonding orbital into an sp^3 orbital. This excitation is relatively localized to the Si=O bond, though some density is observed on the other silicon atoms at both structures.

In addition to the pyramidalized intersection and minimum discussed above, additional exploration uncovered a second well in the PES which was not explored in the

dynamical simulations. This region of the PES is characterized by insertion of the oxygen atom between the initially double bonded silicon atom (Si2) and one of the adjacent silicon atoms (Si1), and the simultaneous breaking of the Si2-Si3 bond. The minimum and intersection in this region of the PES are termed the oxygen insertion (O-i) minimum and intersection. At the CASSCF level of theory, these points on the PES are found to be nearly identical in energy, at 1.92 eV above the ground state minimum energy. The flat PES and near-linearity of the Si-O-Si unit suggest that this conical intersection may be a satellite to a nearby Renner-Teller-like intersection.[32] Single point calculations of the CASSCF optimized structures at the MS-CASPT2 and MRCI(Q) levels of theory predict a small reduction of the energy of the intersection (less than 0.2 eV), as does optimization at the MS-CASPT2 level of theory. The fact that the dynamic simulations do not explore the O-i region of the PES suggests the existence of a substantial barrier between it and the pyramidalized region of the PES, and as such it is difficult to assess the role that the O-i intersection may play in the photodynamics of a real Si-NC containing a Si=O bond.

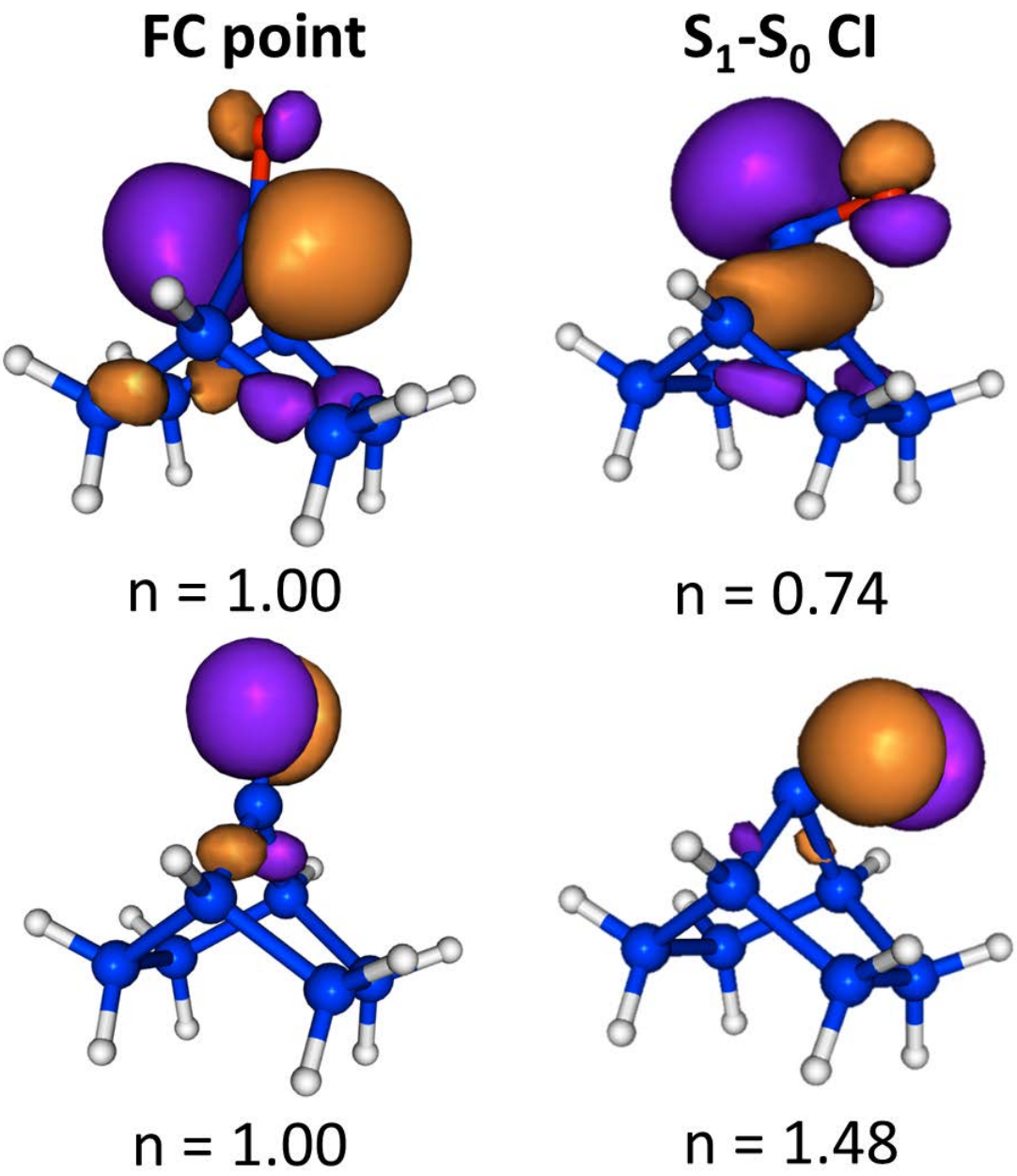


Figure 9-9. Partially occupied CASSCF natural orbitals at the FC point (left) and the S_1-S_0 CI (right) are presented with occupation numbers. The FC natural orbitals are computed by diagonalization of the S_1 density matrix, while the S_1-S_0 CI natural orbitals are computed from the state-averaged S_1-S_0 density matrix.

9.2.3 Discussion of excited state lifetime

As noted above, the observation of the decay of 7% of the excited population within 1 ps after excitation corresponds to an excited state lifetime of 9.5 ps, and such decay would result in internal conversion fast enough to preclude the possibility that silicon-oxygen double bonds are responsible for the experimentally observed S-band PL of Si-NCs (which exhibits a lifetime on the order of microseconds[30,33]). Here we consider the various approximations and assumptions made in this analysis and how they might affect the simulated excited state lifetime. First we consider the choice of a relatively small cluster model of the Si=O defect rather than a whole Si-NC (the simulation of which is not presently feasible at a multireference level of theory). Previous calculations on Si-NCs have demonstrated the strong localization of the electronic excitation to these Si=O bonds and an insensitivity of excitation energy to the size of the surrounding cluster.[1-4,7] A similar localization is observed here, and as such extension of the cluster is unlikely to dramatically affect the energetics of the decay path. It is possible that upon excitation to a delocalized exciton in a larger Si=O defect-containing Si-NC the excited state lifetime would be longer than predicted by our simulations because time would be required for the excitation to localize to the defect. However, the Si=O defect itself cannot emit before localization, so it is the lifetime of the excited defect itself which determines the probability of observing Si=O emission, not the total excited state lifetime. Thus the lifetime of the exciton state is not relevant to the question of whether emission is observed from Si=O defects. Though we see no reason to expect that expansion of our

model would increase the excited state lifetime by the five orders of magnitude necessary to reach the microsecond regime, work is underway to extend the system size and timescales accessible via our dynamical simulations to give a more definitive assessment of the validity of the chosen cluster model.

Another factor affecting the rate of non-radiative decay is the relative energy of the conical intersection and the excited state minimum. As noted above, the energy of a conical intersection relative to the minimum energy point on the PES can be thought of as the activation energy for decay, and a large energy difference will therefore significantly slow internal conversion. The reoptimization of the conical intersection and excited state minimum at the MS-CASPT2 level of theory suggests that the prediction of a 0.45 eV barrier at the CASSCF level of theory may be a significant overestimation of the true activation energy. In addition, no activations energy to reach the O-i intersection is predicted at either the CASSCF or MS-CASPT2 levels of theory, thus suggesting that internal conversion would be efficient if this region of the PES were accessible on longer timescales than those explored by our dynamical simulations.

One final consideration is that the simulated decay may not be exponential. The decay in Figure 9-2 appears to have a fast component in the first ~100 fs after excitation followed by a slower decay. It is possible that this arises due to the limited sampling of our simulations, but it is also possible that intermolecular vibrational energy redistribution slows the decay as energy flows out of the reactive pyramidalization mode. Though such a bi-exponential behavior would result in slower decay than suggested above, the slower

decay observed in the second half of the simulation (500-1000 fs) corresponds to an excited state lifetime of 92 ps, which is still far shorter than the μs experimental lifetime. Taken together, these facts suggest that the dynamical simulations provide a rough upper bound to the non-radiative lifetime of an excited Si=O bond, and thus excited state lifetimes much shorter than the experimentally observed μs radiative lifetime should be expected.

9.3 Conclusions

In this work we have applied ab initio multiple spawning simulations to model the photochemistry of a Si=O bond-containing cluster to investigate the photodynamics of Si=O defects, which have been implicated as emissive defects responsible for the S-band PL of Si-NCs. In contrast to the hypothesis that these defects exhibit PL with a lifetime of microseconds, our simulations predict a non-radiative lifetime of less than 10 ps, and thus suggest that Si=O double bonds are not responsible for the S-band PL. Instead, if present, a Si=O defect may contribute to the experimentally observed decrease in the PL yield of well passivated Si-NCs upon oxidation.

This is the second example we have reported of non-radiative decay due to conical intersections related to oxygen-containing defects in silicon. In conjunction with our previous results, which suggest that conical intersections preferentially quench large-gapped nanocrystals and thus render the ensemble PL energies of oxidized Si-NCs insensitive to particle size, the present result indirectly supports the hypothesis that S-

band PL occurs from a delocalized, quantum-confined exciton state. Given the fact that defects tend to localize electronic excitations into a molecule-sized regions of space, and that excited molecules often decay via such intersections, it would not be surprising to find that decay via conical intersections resulting from the deformation of semiconductor defects is a general mechanism for non-radiative decay. Work is underway to investigate this possibility further in silicon and other materials.

REFERENCES

REFERENCES

(1) Wolkin, M. V.; Jorne, J.; Fauchet, P. M.; Allan, G.; Delerue, C. *Physical Review Letters* **1999**, *82*, 197.

(2) Puzder, A.; Williamson, A. J.; Grossman, J. C.; Galli, G. Journal of Chemical Physics 2002, 117, 6721.

(3) Vasiliev, I.; Chelikowsky, J. R.; Martin, R. M. *Phys. Rev. B* **2002**, *65*.

(4) Puzder, A.; Williamson, A. J.; Grossman, J. C.; Galli, G. Journal of the American Chemical Society 2003, 125, 2786.

(5) Chen, X. B.; Pi, X. D.; Yang, D. R. *Journal of Physical Chemistry C* **2010**, *114*, 8774.

(6) Pennycook, T. J.; Hadjisavvas, G.; Idrobo, J. C.; Kelires, P. C.; Pantelides, S. T. Phys. Rev. B 2010, 82.

(7) Luppi, M.; Ossicini, S. *Phys. Rev. B* **2005**, *71*.

(8) Ben-Nun, M.; Quenneville, J.; Martinez, T. J. Journal of Physical Chemistry A 2000, 104, 5161.

(9) Levine, B. G.; Coe, J. D.; Virshup, A. M.; Martinez, T. J. *Chemical Physics* **2008**, *347*, 3.

(10) Roos, B. O.; Taylor, P. R. *Chemical Physics* **1980**, *48*, 157.

(11) Wadt, W. R.; Hay, P. J. *Journal of Chemical Physics* **1985**, *82*, 284.

(12) Stanton, J. F.; Bartlett, R. J. *The Journal of chemical physics* **1993**, *98*, 7029.

(13) Dunning Jr, T. H. *The Journal of chemical physics* **1989**, *90*, 1007.

(14) Woon, D. E.; Dunning Jr, T. H. *The Journal of chemical physics* **1993**, *98*, 1358.

(15) Klimov, V. I.; McBranch, D. W. *Physical Review Letters* **1998**, *80*, 4028.

(16) Becke, A. D. *Journal of Chemical Physics* **1993**, *98*, 5648.

(17) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. J Phys Chem-Us 1994, 98, 11623.

(18) Finley, J.; Malmqvist, P. A.; Roos, B. O.; Serrano-Andres, L. *Chemical Physics Letters* **1998**, *288*, 299.

(19) Levine, B. G.; Coe, J. D.; Martinez, T. J. *Journal of Physical Chemistry B* **2008**, *112*, 405.

(20) Langhoff, S. R.; Davidson, E. R. *Int J Quantum Chem* **1974**, *8*, 61.

(21) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schutz, M.; Celani, P.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Koppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; O'Neill, D. P.; Palmieri, P.; Pfluger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stol, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M.; Wolf, A. MOLPRO, version 2010.1, a package of ab initio programs, 2010.

(22) Zechmann, G.; Barbatti, M.; Lischka, H.; Pittner, J.; Bonacic-Koutecky, V. Chemical Physics Letters 2006, 418, 377.

(23) Michl, J.; Bonačić-Koutecký, V. Electronic aspects of organic photochemistry; Wiley New York etc., 1990.

(24) Ismail, N.; Blancafort, L.; Olivucci, M.; Kohler, B.; Robb, M. A. Journal of the American Chemical Society 2002, 124, 6818.

(25) Levine, B. G.; Martinez, T. J. *Annu Rev Phys Chem* **2007**, *58*, 613.

(26) Migani, A.; Blancafort, L.; Robb, M. A.; Debellis, A. D. Journal of the American Chemical Society 2008, 130, 6932.

(27) Sinicropi, A.; Page, C. S.; Adam, W.; Olivucci, M. Journal of the American Chemical Society 2003, 125, 10947.

(28) Gozem, S.; Melaccio, F.; Lindh, R.; Krylov, A. I.; Granovsky, A. A.; Angeli, C.; Olivucci, M. J Chem Theory Comput 2013, 9, 4495.

(29) Vial, J. C.; Bsiesy, A.; Gaspard, F.; Herino, R.; Ligeon, M.; Muller, F.; Romestain, R.; Macfarlane, R. M. Phys. Rev. B 1992, 45, 14171.

(30) Huisken, F.; Ledoux, G.; Guillois, O.; Reynaud, C. *Advanced Materials* **2002**, *14*, 1861.

(31) Wen, X. M.; Van Dao, L.; Hannaford, P. *J Phys D Appl Phys* **2007**, *40*, 3573.

(32) Halasz, G. J.; Vibok, A.; Baer, R.; Baer, M. *J Phys a-Math Theor* **2007**, *40*, F267.

(33) Bley, R. A.; Kauzlarich, S. M.; Davis, J. E.; Lee, H. W. H. Chemistry of Materials 1996, 8, 1881.

Chapter 10

Defect-Induced Conical Intersections Promote Non-Radiative Recombination: A Study at Nano-Sized Cluster

"If I have seen further it is by standing on the sholders of Giants"
-Isaac Newton (1643-1727)

This chapter is reproduced from paper: Defect-Induced Conical Intersections Promote Nonradiative Recombination, Y. Shu, B. S. Fales, and B. G. Levine, Nano Lett., 15, 6247 (2015) (Y. Shu and B. S. Fales contributed equally).

There has been much recent interest in developing a more detailed microscopic understanding of non-radiative recombination in semiconductors, the fundamental physical process which converts electronic energy into heat, thus limiting the efficiencies of many optoelectronic devices.[1-9] It has long been known that defects introduce pathways for such recombination.[10,11] Through a combination of experimental and theoretical work it can be established that specific defects, e.g. dangling bonds in silicon,[5,12-21] can act as non-radiative recombination centers, but the a priori theoretical prediction of which defects facilitate non-radiative transitions remains an important area

of development in theoretical nanoscience.[5,7,22-28] In molecular photochemistry, on the other hand, a theory capable of predicting which molecules will undergo fast and efficient non-radiative decay is well established. Over the last several decades it has been demonstrated that conical intersections (CIs), points of degeneracy between the potential energy surfaces (PESs) of different electronic states, provide pathways for such decay.[29-32] In general, a molecule will non-radiatively decay quickly and efficiently to the electronic ground state after excitation when three conditions are met simultaneously: a) a CI is present between the ground and lowest excited state, b) the energy of that CI is below the vertical excitation energy of the lowest electronic excited state, and c) there is not a large energetic barrier in the excited state PES preventing access from the Franck-Condon point.

Defects in semiconductors are similar to molecules in that their electronic excitations involve similar amounts of energy (a few eV) and are localized to approximately the same volume (that of a few atoms). Thus, it is reasonable to hypothesize that defects will undergo non-radiative decay via CIs just as molecules do, and that the presence of a CI meeting the above criteria may indicate that a particular defect facilitates non-radiative recombination. Though a recent groundbreaking study has attributed the cooling of hot electrons in semiconductor nanoparticles to cascades of CIs between excited states,[33] to date, the hypothesis that defect-induced CIs between the ground and first excited state facilitate recombination has only been investigated in molecule-sized cluster models (<10

silicon atoms).[24,34,35] This is due to the lack of an appropriate theoretical method to investigate CIs in nanoscale systems; the most widely used and efficient single reference electronic structure methods, such as density functional theory (DFT)[36] and time-dependent DFT,[37,38] can be readily applied to nanoscale systems, but cannot properly describe the PES in the region surrounding a CI involving the electronic ground state.[39] Multireference electronic structure methods are the state of the art for describing such CIs, but it is only with recent advances in methodology[40] and computer hardware[41] that applying such methods to nanomaterials has become realistic.

In the previous chapters (chapter 6 - 8), we have reported the non-radiative decay mechanism of small silicon cluster with oxygen containing defects. However, those systems are small cluster models. One of the question remains to be answer is whether the CI energy involves system size effects. Here, we address this question by investigating large silicon nanocluster, with our new method and implementation, it is the first time that people are being able to apply multi-reference electronic structure theory to nanometer clusters.

10.1 Cluster Models

In this work we demonstrate that oxygen-containing defects on the surface of silicon nanocrystals (SiNCs) introduce CIs between the ground and first excited electronic states. The structures of low-lying CIs in three SiNCs containing different defects are

shown in Figure 10-1. In panels A and B SiNCs containing silicon epoxide surface defects (three-membered rings composed of one oxygen and two silicon atoms) are shown. Such silicon epoxide rings have been observed in the vibrational spectrum of the silicon-silicon oxide interface.[42] The defects in panels A and B are differentiated by how they are bound to the remainder of the SiNC; the silicon atoms of the epoxide ring in panel A are connected to the remainder of the nanoparticle via three Si-Si bonds, while those in panel B are bound via only two Si-Si bonds. These defects will be referred to as type 1 and 2 epoxide defects, respectively, and their chemical structures are pictured in the first column of Figure 10-1, with the remainder of the SiNC represented by gray silicon atoms. The SiNCs in panels A and B contain 44 and 50 silicon atoms, respectively. They are slightly oblate with an average diameter of 1.3 nm and an aspect ratio of 1.3. All surface dangling bonds in all systems studied in this work are capped with hydrogen atoms. As reported in Table 1, our multireference calculations predict that the type 1 and 2 epoxide CIs reside 1.6 and 1.8 eV above the ground state minimum energy, respectively, and are therefore both accessible upon excitation by visible light.

The third defect studied in this work is the silicon-oxygen double (Si=O) bond. This defect has frequently been implicated as the source of the size-insensitive slow (or S-) band emission of oxidized SiNCs at 1.5-2.1 eV,[43-49] though this assignment has been quite controversial.[34,50-54] Indeed a SiNC of 44 Si atoms (1.3 nm average diameter) containing a Si=O defect exhibits a low-lying CI, pictured in Figure 10-1C. This CI is found to be 2.5

eV above the ground state minimum and therefore is accessible upon excitation at visible wavelengths. This suggests that Si=O defects could act as efficient non-radiative centers.

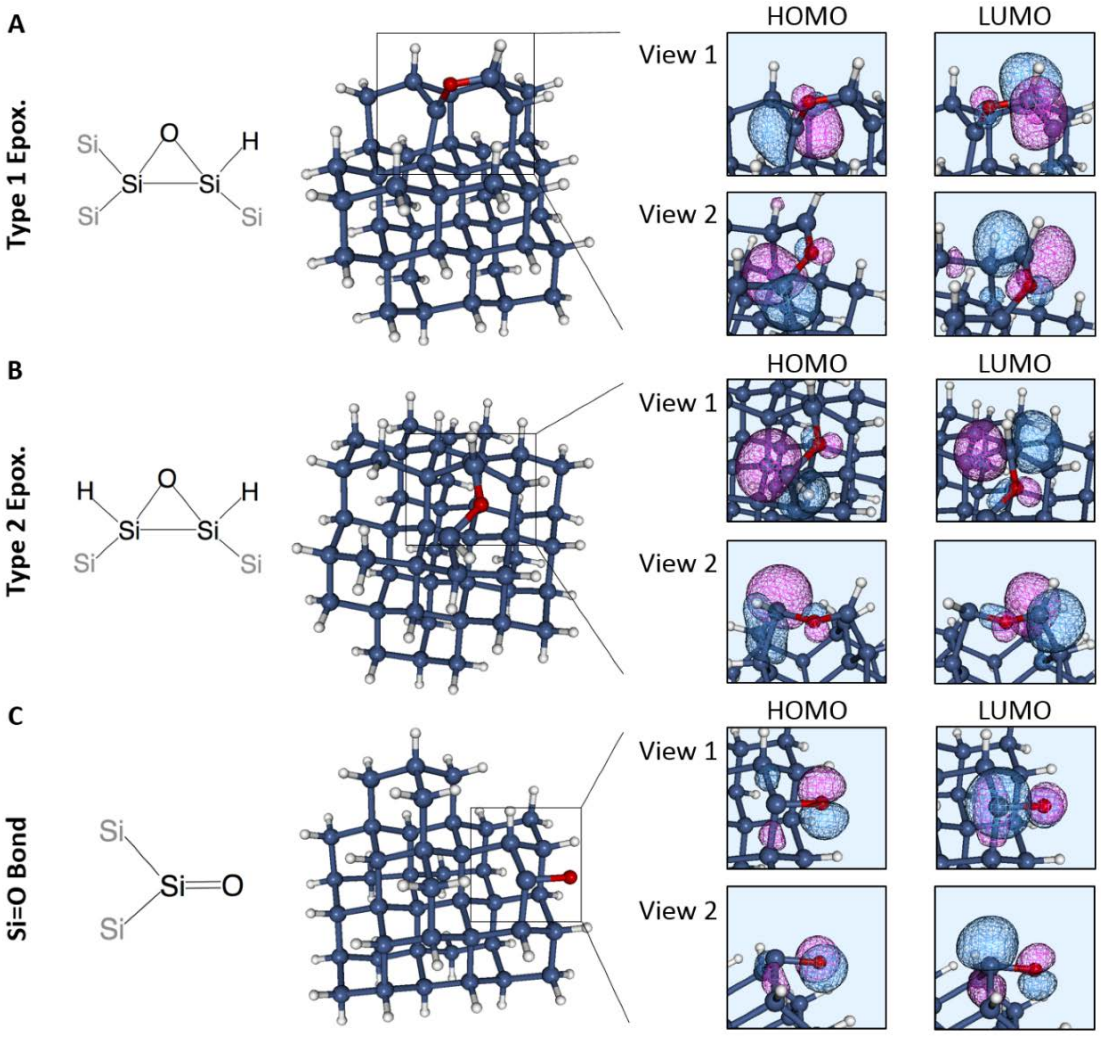


Figure 10-1. Panels A, B, and C show key features of the three CIs discussed in this work. The defects are defined in the first column, the optimized CI structures are shown in the second column, and two different views of the HOMOs and LUMOs involved in the non-radiative transitions at each CI are shown in the third column.

10.2 Computational Details

All of the energies reported above are computed at a multireference level of theory using a best-estimate approach, which includes both static and dynamic electronic correction. Our computational scheme is outlined here, but a fully explicit description can be found in supporting information. Intersections were optimized at the configuration interaction singles natural orbital complete active space configuration interaction (CISNO-CASCI) level of theory[40] with the LANL2DZ basis.[55] The variant of CISNO based on a relaxed CIS density matrix (CISNOr) was used. Active spaces for each particle were chosen based on prior work on small cluster models (abbreviated *M*/*N* where *M* is the number of active electrons and *N* is the number of active orbitals): 2/4 and 4/3 for the epoxide and Si=O defects, respectively. In all cases, three states were included in the CISNOr density matrix used to determine the orbitals. In each optimization, only the coordinates of those atoms in the immediate region of the defect are allowed to vary, while the remainder of the atoms are frozen in their CAM-B3LYP[56]/LANL2DZ ground state minimum structure. This constrained optimization approach provides a rigorous upper bound to the energies of the true minimal energy CIs. For each defect, three CIs were optimized in particles of different size, as described above. A multi-state second order perturbation correction for dynamic electron correlation (CASPT2)[57] was applied to the smallest particles in each set. The CASPT2 correction to the CI energy computed for the smallest particle with each defect was added to the CI

energies of the larger particles to provide our best estimate of the CI energy. In all cases this energy correction is quite small: 0.3 eV or less. For the smallest particle containing each defect, fully variational complete active space self-consistent field (CASSCF)[58] optimizations and CASPT2 calculations were performed with the polarized LANL2DZdp basis set for comparison, and the results obtained were nearly identical to those computed using this CISNOr-CASCI approach and the smaller LANL2DZ basis. Raw energies from all of these calculations are reported in supporting information. The orbitals presented in Figure 10-1 are state-averaged (S_0/S_1) natural orbitals, with the HOMO and LUMO having occupation numbers of 1.5 and 0.5 in all cases. Vibrational frequencies were computed at the MP2/6-31G** level using cluster models defined in the Supporting Information. CISNOr-CASCI optimizations were performed using a development version of the TeraChem graphics processing unit (GPU-) accelerated electronic structure software package,[41,59] in conjunction with the CIOpt intersection optimizer.[60,61] Parallel numerical energy derivatives were used to accelerate the optimization. The GPU-accelerated implementation of CASCI will be discussed in a subsequent publication. CASSCF optimizations and single point CASPT2 corrections were computed using the MolPro electronic structure package.[62-66]

10.3 Size Insensitive Defects Conical Intersections

In order to investigate the possibility that hydrogen-terminated SiNCs exhibit low-lying CIs similar to those introduced by oxygen-containing defects, we have performed an intersection optimization for a small cluster of hydrogen-terminated silicon. This cluster was found to exhibit CIs, but only at energies above 3.8 eV, well outside the visible region and thus not accessible after excitation by visible light. Thus, the above defects introduce CIs at lower energies than are present in pristine silicon. See online Supporting Information for details (http://pubs.acs.org/doi/suppl/10.1021/acs.nanolett.5b02848).

One key feature of all three defect-induced CIs is that accessing them requires significant distortion of their nuclei relative to their ground state equilibrium structures. As pictured in Figure 10-2, the CIs in the epoxide-containing SiNCs are reached only after photochemical ring opening of the epoxide ring (with the distance between epoxide silicon atoms increasing from 2.2 Å at the ground state minimum to 2.8 Å at the CI for the type 1 defect and from 2.2 to 3.0 Å for type 2). Similarly, the Si=O CI is reached by pyramidalization (sp^3 hybridization) of the double-bonded silicon atom, which has a planar sp^2 structure in the ground state. A distinct feature of the Si=O cluster is that this intersection is not the lowest point on the PES of the first singlet excited state (S_1). It exhibits an S_1 minimum with a similarly pyramidalized structure that falls below the CI by 0.9 eV. This minimum is illustrated in Figure S1. In contrast, for the epoxide containing

SiNCs no minima lower than their respective intersection energies was observed on S_1. The absence of a minimum indicates that the epoxide intersections are "peaked" intersections which efficiently funnel population to the electronic ground state.[67] In contrast the presence of a separate minimum in the Si=O case indicates that the non-radiative transition may be much slower, because the PES slopes away from the intersection. Radiative decay from this minimum may be possible at low temperature (as is the case for dangling bond defects[16]), but as will be discussed below, our calculations are not consistent with the assignment of the experimentally observed photoluminescence of SiNCs to Si=O defects.

Another key feature of these three CIs is that their electronic excitations are localized to a small number of atoms surrounding their defects. The highest occupied and lowest unoccupied molecular orbitals (HOMOs and LUMOs) describing the electronic transition at each CI are pictured as insets in Figure 10-1. In the epoxide defects, passage through the CI involves transfer of an electron between silicon non-bonding orbitals (created during the ring opening reaction), both of which include some oxygen non-bonding (lone pair) character. Similarly, passage through the Si=O defect CI involves transfer of an electron between a silicon non-bonding orbital (an sp^3 orbital created upon pyramidalization) with some oxygen non-bonding character and an oxygen non-bonding (lone pair) orbital. These intersections, each of which involves an electron transitioning between a LUMO and HOMO which are rendered nearly degenerate by geometric

distortion, are examples of biradicaloid structures, which are one of the primary classes of conical intersections in organic photochemistry and have been described in detail by Michl and Bonaĉić-Koutecký.[68]

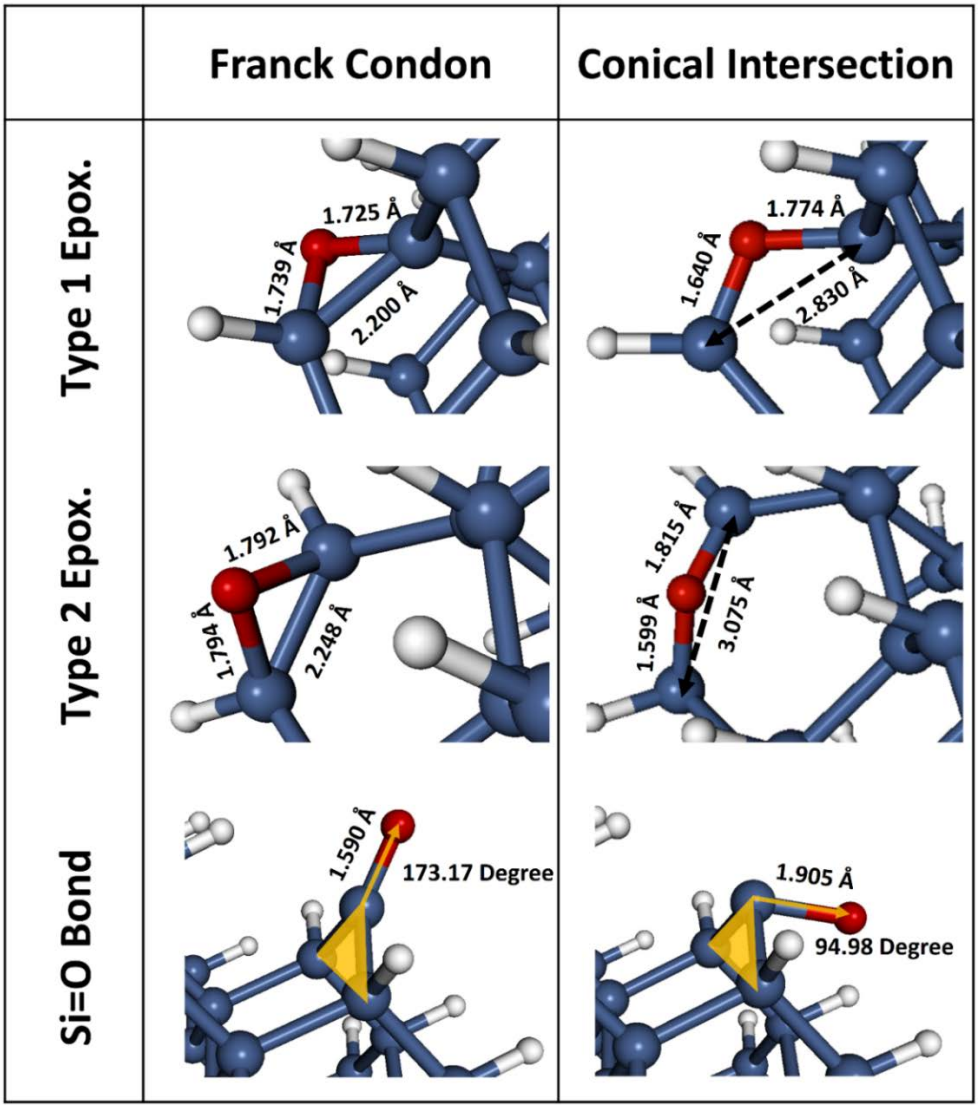


Figure 10-2. Selected geometric parameters at the Franck-Condon (left) and optimized conical intersection (right) structures for each defect. For the Si=O defect we report the pyramidalization angle, defined as the angle between the Si=O bond and the plane defined by the double-bonded silicon atom and the two silicon atoms to which it is bound.

That the electronic processes associated with passage through these CIs are localized is supported by the fact that their energies are not strongly particle-size-dependent. We have investigated the particle size dependence of the CI energies by optimizing CIs in three particles of varying size for each defect, ranging from 14 to 44 silicon atoms for the type 1 epoxide and Si=O defects, and from 15 to 50 silicon atoms for the type 2 epoxide defect. For each defect class, the energies of the conical intersections as a function of particle size are presented in Figure 10-3. In all cases only a weak dependence on particle size is observed. For the type 1 and 2 epoxides, the energies decrease slightly with increasing system size, from 1.8 to 1.6 eV and 2.1 to 1.8 eV, respectively. This would be consistent with very weak quantum confinement effects, but given the local nature of the orbitals and the charge transfer character of the non-radiative process it is more likely due to increased dielectric stabilization of the excited state by the surrounding silicon. If dielectric stabilization plays a role in the observed size dependence of the CI energy, we would expect the energy to depend on the environment (e.g. solvent) as well. An experimentally observed environment dependence of the PL yield and lifetime thus may provide insight into exactly which defects are responsible for decay. In contrast to the epoxides, the energy of the Si=O CI increases weakly with increasing system size, from 2.3 to 2.5 eV. This behavior may arise from dielectric stabilization of the ground state or due to strain induced by the larger, more rigid clusters. Note that the characters of

the orbitals involved in the CI are also insensitive to particle size, as can be seen in Figure S2.

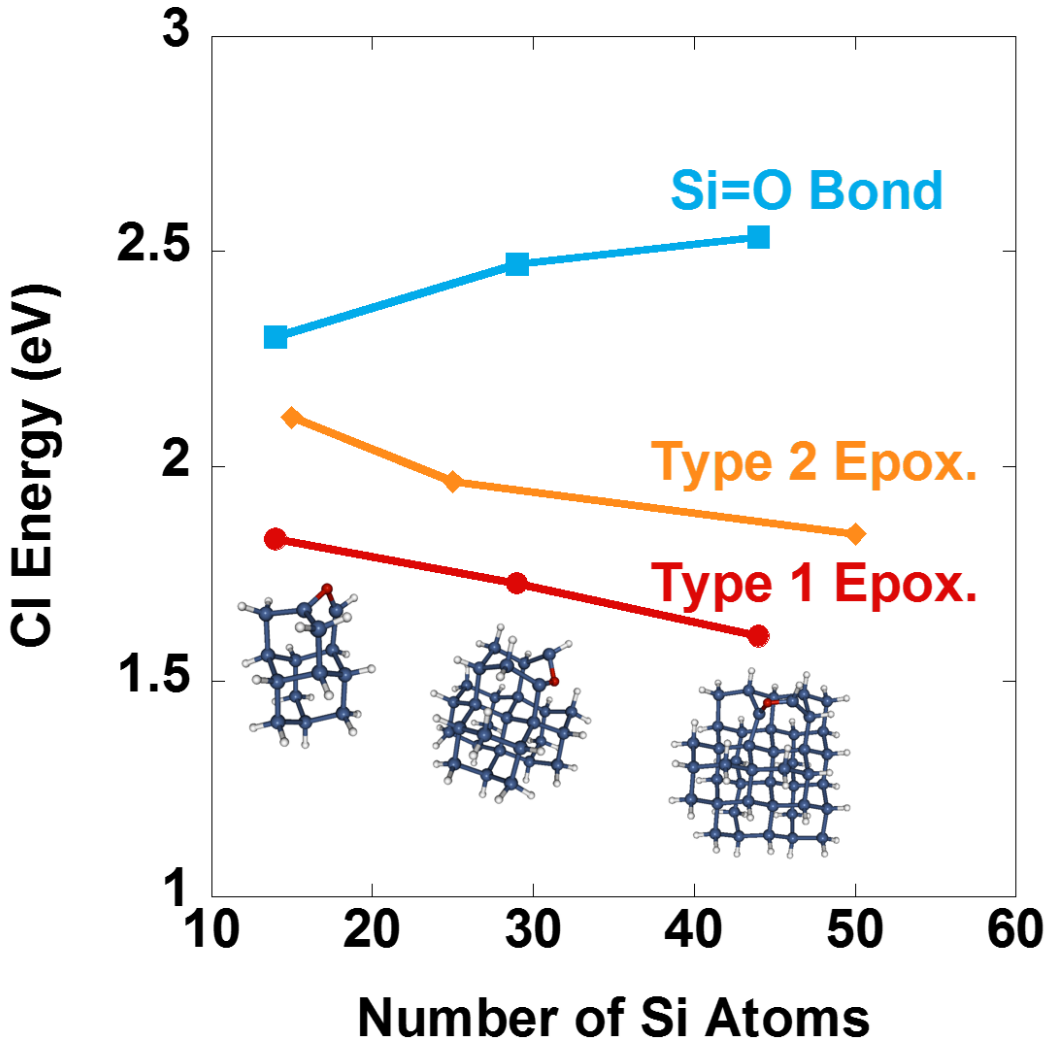


Figure 10-3. Particle size dependence of the CI energy (relative to the ground state minimum energy) of the type 1 epoxide (red), type 2 epoxide (orange), and Si=O (blue) defects.

10.4 Recombination Rate Analysis

In the context of non-radiative recombination, defects are often thought of as static structural features that introduce one-electron states within the band gap through which charge carriers can shuttle between the valence and conduction bands. Considering that significant nuclear motion is necessary to reach the CIs reported here, a very different picture emerges in the present case. Consistent with an important recent theoretical study of the blinking of SiNCs,[5] we expect that in many cases non-radiative recombination occurs by an activated process, analogous to a Marcus electron transfer reaction.[69,70] As illustrated in Figure 10-4, at least three potential energy surfaces play a role in the recombination process: those of the ground state (gray), the first quantum-confined excited state (green), and the defect-localized excited state (red). (Note that the defect-localized state is that in which an electron-hole pair remains excited, but both charge carriers are localized to the defect.) Due to the relatively delocalized nature of the exciton, the surface of the quantum-confined state and ground state are nearly parallel, but as clearly seen in this work, the defect-localized state can be strongly displaced along some nuclear coordinate corresponding to distortion of the defect. (In fact, the vertical excitation energies of the epoxide defects are quite high—4.4 and 5.7 eV for type 1 and 2, respectively—but distortion reduces the energy of the defect-localized state, thus allowing CIs which are accessible at visible energies. The vertical excitation energy of the Si=O defect, 2.9 eV, is considerably lower to begin with, but its energy is still

decreases upon displacement.) In this study we have directly computed only the crossing between the defect-localized state and the ground state; it is the energy of this CI that determines at what energy non-radiative recombination via CI becomes energetically possible.

However, the rate of non-radiative recombination is likely determined by the crossing between the quantum-confined and defect-localized states, which corresponds to the barrier for localization. Given that the energy of the quantum-confined state can be tuned by changing the size of the SiNC, but the energy of the defect-localized state (as reflected in the CI energies reported above) is relatively size-insensitive, the particle size will determine the thermodynamic driving force ($-\Delta G^{\circ}$)—and therefore the barrier, $\Delta G^{\ddagger}$—for localization. Smaller SiNCs (Figure 10-4, upper left) will have a larger driving force, and therefore a faster rate of localization and subsequent recombination. Larger SiNCs (Figure 10-4, upper right) will have a smaller driving force resulting in a larger barrier to recombination. The increased rate of recombination in smaller SiNCs is analogous to the acceleration of an oxidation or reduction reaction due to increased overpotential, except the potential in this case is the energy of the entire many-electron system rather than that of an individual charge carrier. Estimated barriers to localization ($\Delta G^{\ddagger}$) are reported in the lower panel of Figure 10-4. These barriers are computed via Marcus theory, using the energies reported in Table 10-1. Because we have not computed the quantum-confined state explicitly, and the energy of this state is needed to determine the driving

force for localization, we report the activation energy over a range of possible quantum-confined state energies corresponding to a range of different particle sizes (see Supporting Information for methodological details). We see that the activation energy is strongly size-dependent, and can become quite low at the upper end of the visible range. Thus, vertical excitation directly to the defect-localized state is not necessary for efficient recombination, because the microsecond excited state lifetime of SiNCs leaves enough time for thermal nuclear motion over a barrier.

Table 10-1. Energies of the first singlet excited states of the three SiNCs pictured in Figure 10-1 at their respective Franck-Condon points (FC) and optimized conical intersections (CI). Energies are reported relative to the ground state minimum energy.

Defect	FC energy (eV)	CI energy (eV)
Type 1 epox.	4.41	1.61
Type 2 epox.	5.56	1.84
Si=O bond	2.91	2.53

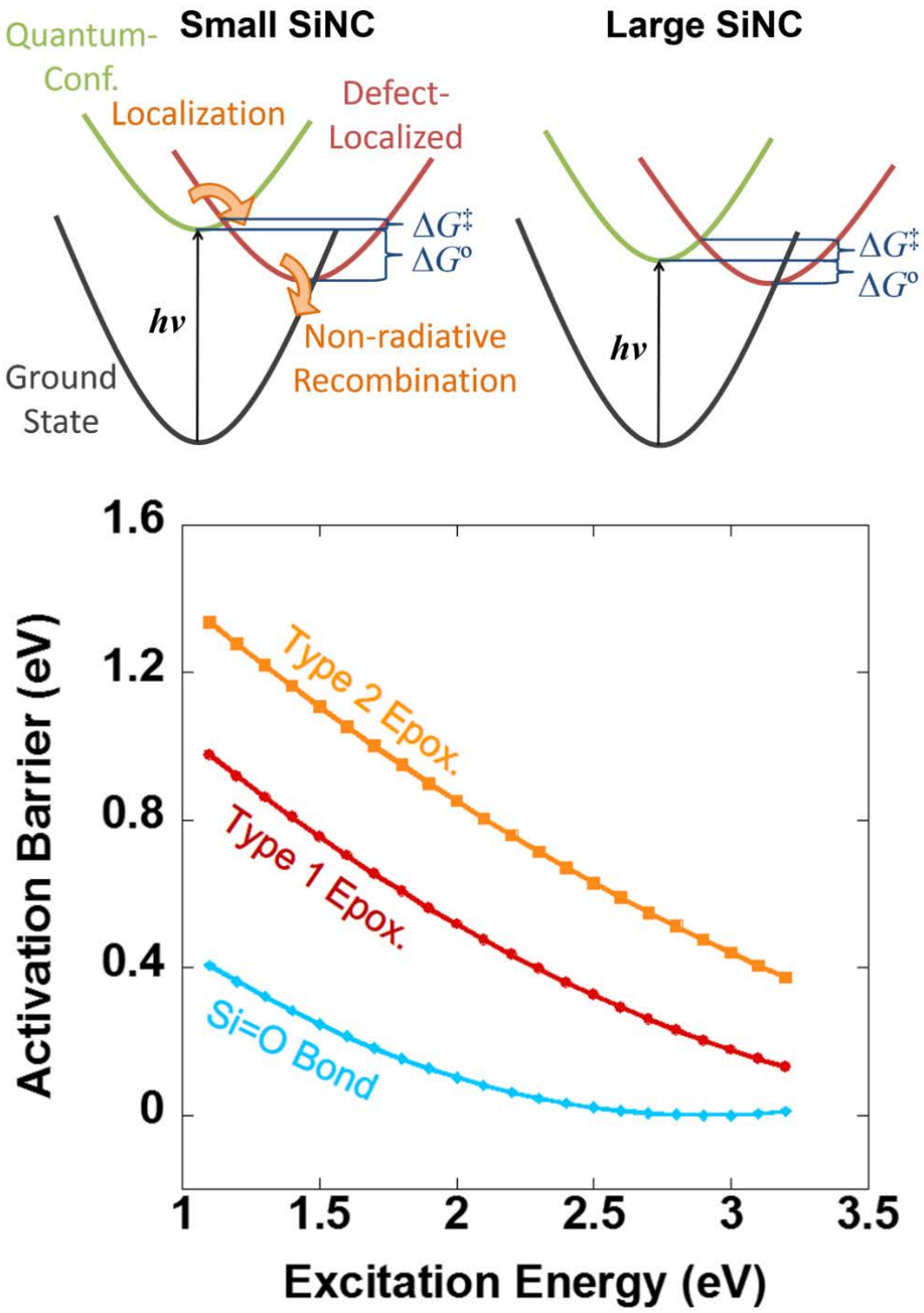


Figure 10-4. Upper panel: Idealized representation of the PESs of the ground (gray), quantum-confined (green), and defect-localized (red) states of SiNCs of different sizes illustrating how the barrier to localization (crossing between the red and green curves),

Figure 10-4 (cont'd). and thus the rate of recombination, can be strongly size-dependent, though the energy of the defect-localized conical intersection (crossing between the red and gray curves) is not. PESs are represented as a function of a coordinate which locally distorts the nuclear structure of the defect so as to reduce the energy of the defect-localized excited state (e.g. epoxide ring opening or Si=O pyramidalization). Lower panel: Estimated activation barriers for localization to the three defects studied here, $\Delta G^{\ddagger}$, are plotted as a function of the excitation energy to the quantum-confined state.

This prediction of a size-dependent non-radiative recombination rate is consistent with several experimental observations of SiNC photoluminescence; higher-gap samples of oxidized SiNCs exhibit smaller quantum yields than lower-gap ones,[71] and higher energy emission observed from ensembles of SiNCs exhibits a shorter lifetime than lower energy emission.[72] In fact, the photoluminescence lifetime begins to drop sharply at 2.0-2.2 eV, suggesting the introduction of a new mechanism for non-radiative recombination accessible only above this energy.[72] This energy is comparable to the energies of the CIs reported in this work.

The size-dependent rate of recombination also relates to a longstanding debate about the source of the S-band (1.5-2.1 eV) photoluminescence of SiNCs, the energy of which is observed to be insensitive to the average size of the SiNCs in ensemble PL measurements.[43] As noted above, some have suggested that this size-insensitive

emission arises from an emissive surface defect, but the identity of this defect remains elusive. The Si=O defect is often cited as the emitter, but the presence of a conical intersection associated with this defect calls this assignment into question, and past simulations have shown that Si=O defects introduce non-radiative rates several orders of magnitude faster than the microsecond radiative lifetime of the S-band.[34] Others have argued on the basis of its long lifetime and other features that the emission arises from a quantum-confined state,[73-75] but this hypothesis has been hard to reconcile with the size-insensitivity of the emission energy. The existence of non-radiative defects which preferentially quench the excitations of smaller SiNCs with higher gaps allows us to reconcile this apparent inconsistency. Note that these experiments probe the photoluminescence of ensembles of nanoparticles which are not uniform in size. When such an ensemble is oxidized, defects may be introduced which preferentially quench smaller SiNCs with larger gaps, leaving only larger SiNCs with smaller gaps to emit. Thus, an ensemble of dots which emits in the blue prior to oxidation would shift strongly to the red upon oxidation, because only the smaller-gapped subpopulation remains emissive. The result is that ensembles that differ significantly in their average particle size may have photoluminescence maxima at similar energies after oxidation because the subset of the SiNCs which emit efficiently are roughly the same size.

The specificity of the displacements required to reach the conical intersections indicates that specific vibrations may be excited upon return to the ground state. These

vibrational excitations, which would likely be localized to the defects, can potentially be observed as anti-Stokes Raman scattering lines. We have computed the frequencies of vibrational modes corresponding to the displacements required to reach the CIs. Characteristic frequencies may include epoxide Si-Si stretching at 537 and 520 cm^{-1} and Si=O stretching at 1137 cm^{-1}. Si=O out-of-plane bending (pyramidalization) at 254 cm^{-1} is a direction of strong displacement, but the local symmetry of the sp^2 Si atom suggests the fundamental line would be weak. Still, given the magnitude of the displacement, to the extent that Si=O bonds exist overtones may be observable. Note that dangling bonds likely play a key role in the recombination of oxidized SiNCs, and a similar analysis of their anti-Stokes signature is forthcoming.

The above analysis is based on the assumption that the system cools to the lowest quantum-confined state prior to localization and recombination regardless of which state is created upon absorption. Excitation directly into the defect-localized excited state could change this picture, leading to diabatic passage over the barrier and thus ultrafast recombination. This is particularly interesting in the case of the Si=O defect, whose vertical excitation energy of 2.91 eV is within the visible range. Though diabatic passage over the barrier seems likely, some transfer of excited population from the defect-localized state back to one of the quantum-confined states is possible. The complexity of such dynamics is beyond the scope of the present work, but we hope to study their extent via non-adiabatic molecular dynamics (NAMD) methods currently under

development. The above analysis also only considers the barrier for simultaneous transfer of the electron and hole. Sequential transfer is a possibility which can also be investigated by subsequent NAMD.

10.5 Conclusions

We have shown that defects introduce conical intersections between the ground and first singlet excited state PESs of three defective SiNCs. These defect-induced CIs are energetically accessible upon excitation at visible energies. The locality of the excitations which arise at these CIs and the large nuclear displacements relative to the ground state structures needed to access them suggest that localization may be an activated process in many cases, and this argument is consistent with several experimentally observed features of the S-band photoluminescence of SiNCs. We hypothesize that defect-induced CIs are likely to arise at other defects in silicon—e.g. dangling bonds—and other semiconductors, such as II-VI materials which exhibit blinking behavior attributed to the charging of surface defects.[1] We will apply our optimization scheme to investigate these possibilities in future studies.

REFERENCES

REFERENCES

(1) Galland, C.; Ghosh, Y.; Steinbruck, A.; Sykora, M.; Hollingsworth, J. A.; Klimov, V. I.; Htoon, H. *Nature* **2011**, *479*, 203.

(2) Kamat, P. V. *Acc, Chem. Res.* **2012**, *45*, 1906.

(3) Rao, A.; Chow, P. C. Y.; Gelinas, S.; Schlenker, C. W.; Li, C. Z.; Yip, H. L.; Jen, A. K. Y.; Ginger, D. S.; Friend, R. H. Nature 2013, 500, 435.

(4) Gabriel, M. M.; Grumstrup, E. M.; Kirschbrown, J. R.; Pinion, C. W.; Christesen, J. D.; Zigler, D. F.; Cating, E. E. M.; Cahoon, J. F.; Papanikolas, J. M. Nano Lett. 2014, 14, 3079.

(5) Brawand, N. P.; Voros, M.; Galli, G. *Nanoscale* **2015**, *7*, 3737.

(6) Najafi, E.; Scarborough, T. D.; Tang, J.; Zewail, A. *Science* **2015**, *347*, 164.

(7) Wang, L. J.; Long, R.; Prezhdo, O. V. *Annu. Rev. Phys. Chem.* **2015**, *66*, 549.

(8) Wang, H. N.; Zhang, C. J.; Rana, F. *Nano Lett.* **2015**, *15*, 339.

(9) Frost, J. M.; Butler, K. T.; Brivio, F.; Hendon, C. H.; van Schilfgaarde, M.; Walsh, A. Nano Lett. 2014, 14, 2584.

(10) Hall, R. N. *Phys. Rev.* **1952**, *87*, 387.

(11) Shockley, W.; Read, W. T. *Phys. Rev.* **1952**, *87*, 835.

(12) Poindexter, E. H.; Caplan, P. J.; Deal, B. E.; Razouk, R. R. *J. Appl. Phys.* **1981**, *52*, 879.

(13) Poindexter, E. H.; Gerardi, G. J.; Rueckel, M. E.; Caplan, P. J.; Johnson, N. M.; Biegelsen, D. K. J. Appl. Phys. 1984, 56, 2844.

(14) Hane, M.; Miyamoto, Y.; Oshiyama, A. *Phys. Rev. B* **1990**, *41*, 12637.

(15) Lannoo, M.; Delerue, C.; Allan, G. *J Lumin* **1993**, *57*, 243.

(16) Meyer, B. K.; Hofmann, D. M.; Stadler, W.; Petrovakoch, V.; Koch, F.; Omling, P.; Emanuelsson, P. Appl. Phys. Lett. 1993, 63, 2120.

(17) Stesmans, A. *Phys. Rev. Lett.* **1993**, *70*, 1723.

(18) Kageshima, H.; Shiraishi, K. *Surface Science* **1998**, *407*, 133.

(19) Stesmans, A.; Nouwen, B.; Afanas'ev, V. V. *Phys. Rev. B* **1998**, *58*, 15801.

(20)Bratus, V. Y.; Yukhimchuk, V. A.; Berezhinsky, L. I.; Valakh, M. Y.; Vorona, I. P.; Indutnyi, I. Z.; Petrenko, T. T.; Shepeliavyi, P. E.; Yanchuk, I. B. Semiconductors 2001, 35, 821.

(21) Pereira, R. N.; Rowe, D. J.; Anthony, R. J.; Kortshagen, U. *Phys. Rev. B* **2011**, *83*.

(22) Meng, S.; Kaxiras, E. *Nano Lett.* **2010**, *10*, 1238.

(23) Inerbaev, T. M.; Hoefelmeyer, J. D.; Kilin, D. S. *J. Phys. Chem. C* **2013**, *117*, 9673.

(24) Shu, Y.; Levine, B. G. *J. Chem. Phys.* **2013**, *139*.

(25) Agiorgousis, M. L.; Sun, Y. Y.; Zeng, H.; Zhang, S. B. *J. Am. Chem. Soc.* **2014**, *136*, 14570.

(26) Liu, J.; Neukirch, A. J.; Prezhdo, O. V. *J. Phys. Chem. C* **2014**, *118*, 20702.

(27)Yin, W. J.; Wu, Y. L.; Wei, S. H.; Noufi, R.; Al-Jassim, M. M.; Yan, Y. F. Adv. Energy Mater. 2014, 4.

(28) Brown, E.; Sheng, C. Y.; Shimamura, K.; Shimojo, F.; Nakano, A. J. Appl. Phys. 2015, 117.

(29) Bernardi, F.; Olivucci, M.; Robb, M. A. *Chem. Soc. Rev.* **1996**, *25*, 321.

(30) Yarkony, D. R. *Rev. Mod. Phys.* **1996**, *68*, 985.

(31) Levine, B. G.; Martinez, T. J. *Annu. Rev. Phys. Chem.* **2007**, *58*, 613.

(32) Domcke, W.; Yarkony, D. R. *Annu. Rev. Phys. Chem.* **2012**, *63*, 325.

(33) Berardo, E.; Hu, H. S.; van Dam, H. J. J.; Shevlin, S. A.; Woodley, S. M.; Kowalski, K.; Zwijnenburg, M. A. J. Chem. Theory Comput. 2014, 10, 5538.

(34) Shu, Y.; Levine, B. G. *J. Phys. Chem. C* **2014**, *118*, 7669.

(35) Shu, Y.; Levine, B. G. *J. Phys. Chem. C* **2015**, *119*, 1737.

(36) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, 1133.

(37) Casida, M. E.; Jamorski, C.; Casida, K. C.; Salahub, D. R. *J. Chem. Phys.* **1998**, *108*, 4439.

(38) Gross, E. K. U.; Kohn, W. *Phys. Rev. Lett.* **1985**, *55*, 2850.

(39) Levine, B. G.; Ko, C.; Quenneville, J.; Martinez, T. J. *Mol. Phys.* **2006**, *104*, 1039.

(40) Shu, Y.; Hohenstein, E. G.; Levine, B. G. *J. Chem. Phys.* **2015**, *142*.

(41) Ufimtsev, I. S.; Martinez, T. J. *J. Chem. Theory Comput.* **2009**, *5*, 2619.

(42) Stefanov, B. B.; Gurevich, A. B.; Weldon, M. K.; Raghavachari, K.; Chabal, Y. J. Phys. Rev. Lett. 1998, 81, 3908.

(43) Wolkin, M. V.; Jorne, J.; Fauchet, P. M.; Allan, G.; Delerue, C. Phys Rev Lett 1999, 82, 197.

(44) Puzder, A.; Williamson, A. J.; Grossman, J. C.; Galli, G. *J. Chem. Phys.* **2002**, *117*, 6721.

(45) Vasiliev, I.; Chelikowsky, J. R.; Martin, R. M. *Phys. Rev. B* **2002**, *65*.

(46) Puzder, A.; Williamson, A. J.; Grossman, J. C.; Galli, G. J. Am. Chem. Soc. 2003, 125, 2786.

(47) Luppi, M.; Ossicini, S. *Phys. Rev. B* **2005**, *71*.

(48) Chen, X. B.; Pi, X. D.; Yang, D. R. *J. Phys. Chem. C* **2010**, *114*, 8774.

(49) Pennycook, T. J.; Hadjisavvas, G.; Idrobo, J. C.; Kelires, P. C.; Pantelides, S. T. Phys. Rev. B 2010, 82.

(50) Zhou, Z. Y.; Brus, L.; Friesner, R. *Nano Lett.* **2003**, *3*, 163.

(51) Ramos, L. E.; Furthmuller, J.; Bechstedt, F. *Phys. Rev. B* **2004**, *70*.

(52) Konig, D.; Rudd, J.; Green, M. A.; Conibeer, G. Sol. Energy Mater. Sol. Cells 2009, 93, 753.

(53) Seino, K.; Bechstedt, F.; Kroll, P. *Nanotechnology* **2009**, *20*.

(54) Degoli, E.; Guerra, R.; Iori, F.; Magri, R.; Marri, I.; Pulci, O.; Bisi, O.; Ossicini, S. C. R. Phys. 2009, 10, 575.

(55) Wadt, W. R.; Hay, P. J. *J. Chem. Phys.* **1985**, *82*, 284.

(56) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51.

(57) Finley, J.; Malmqvist, P. A.; Roos, B. O.; Serrano-Andres, L. Chem. Phys. Lett. 1998, 288, 299.

(58) Roos, B. O.; Taylor, P. R. *Chem. Phys.* **1980**, *48*, 157.

(59) Isborn, C. M.; Luehr, N.; Ufimtsev, I. S.; Martinez, T. J. J. Chem. Theory Comput. 2011, 7, 1814.

(60) Levine, B. G.; Coe, J. D.; Martinez, T. J. *J Phys Chem B* **2008**, *112*, 405.

(61) Hansen, J. A.; Piecuch, P.; Levine, B. G. *J. Chem. Phys.* **2013**, *139*.

(62) Werner, H. J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schutz, M. WIREs Comput. Mol. Sci. 2012, 2, 242.

(63) Knowles, P. J.; Werner, H. J. *Chem. Phys. Lett.* **1985**, *115*, 259.

(64) Werner, H. J.; Knowles, P. J. *J. Chem. Phys.* **1985**, *82*, 5053.

(65) Werner, H. J. *Mol. Phys.* **1996**, *89*, 645.

(66) Celani, P.; Werner, H. J. *J. Chem. Phys.* **2000**, *112*, 5546.

(67) Yarkony, D. R. *J. Chem. Phys.* **2001**, *114*, 2601.

(68) Michl, J.; Bonacic-Koutecky, V. Electronic Aspects of Organic Photochemistry; Wiley: New York, 1990.

(69) Marcus, R. A. *J. Chem. Phys.* **1956**, *24*, 966.

(70) Barbara, P. F.; Meyer, T. J.; Ratner, M. A. *J. Phys. Chem.* **1996**, *100*, 13148.

(71) Jurbergs, D.; Rogojina, E.; Mangolini, L.; Kortshagen, U. *Appl. Phys. Lett.* **2006**, *88*.

(72) Huisken, F.; Ledoux, G.; Guillois, O.; Reynaud, C. *Adv. Mater.* **2002**, *14*, 1861.

(73) Garcia, C.; Garrido, B.; Pellegrino, P.; Ferre, R.; Moreno, J. A.; Morante, J. R.; Pavesi, L.; Cazzanelli, M. Appl. Phys. Lett. 2003, 82, 1595.

(74) Sychugov, I.; Juhasz, R.; Valenta, J.; Linnros, J. *Phys. Rev. Lett.* **2005**, *94*.

(75) Wen, X. M.; Van Dao, L.; Hannaford, P. *J. Phys. D-Appl. Phys.* **2007**, *40*, 3573.

Chapter 11

Silicon Cluster Terminated with H Atoms

"Concepts that have proven useful in ordering things easily achieve such authority over us that we forget their earthly origins and accept them as unalterable givens."
-Albert Einstein (1879-1955)

Previous chapters (chapter 7-10) have emphasized on the study of the non-radiative recombination processes of silicon nanocrystal (SiNC) with various oxygen containing defects. The results suggest that the defect induced conical intersection (CI) do not just play a role in determining the PL quantum yield, but also affect the observed peak energy of PL. CI tends to quench nanocrystals (NCs) with larger band gaps and hence result in the insensitivity of PL peak energy to the particle size. It has been hypothesized in Chapter 1 that local nuclear deformation can be treated as a type of defect as well. In this chapter, we will explore such possibility and investigate whether the existence of CIs arising from deformation of semiconductor local structure is a general non-radiative decay pathway for materials without defects. One of the model systems to study is the pristine SiNC capped with H atoms (we will simply denote it as pristine SiNC in the following text). As the red shift of the PL peaks of the oxidized SiNC is compared with the pristine SiNC, it suggests that the energy level of the minimal energy conical intersection (MECI) of the

pristine SiNC should be higher than that of oxidized SiNC according to our non-radiative recombination mechanism. The PL energy of the pristine SiNC can be predicted by quantum confinement effect.[1-6] In this chapter, we will discuss the following two questions, 1) What is the energy level of the minimal energy conical intersection of the pristine SiNC? 2) Is there any experimental evidence to confirm that the PL energy of pristine SiNC deviates from what has been predicted by quantum confinement effect due to the existence of the conical intersection?

Previous experiments show that ultra-small SiNCs have been synthesized with small size distribution. However, the PL maximum locates between 320 nm to 500 nm depend on the methods of synthesis.[2,7-21] To name a few, Gennadiy et al.[8] synthesized SiNC capped with H atoms with size around 1 nm while the PL maximum locates around 390 nm. Elena and co-workers[10] claimed observation of H-passivated SiNC with 1 nm size has PL maximum around 320 nm with high quantum yield, which is also the shortest PL maximum shown in the literature according to the best of our knowledge. Jonghoon et al.[13] have prepared the ultra-small SiNC by dissolution process in the acid mixture. However, the PL maximum never drops below 500 nm with size around 1 nm. Melanie et al.[21] have reported SiNC with 1 nm size with the PL peak is above 500 nm and low quantum yield.

The experimental PL energy deviates from that have been predicted by theoretical calculations based on quantum confinement effect.[22-26] A clear summary of the difference is provided by Hill and Whaley[22] in which they compared the experimental PL

energy with theoretical predictions at different theoretical levels (including methods improved over the effective mass approximations and tight bonding theoretical levels) with difference sizes of the pristine SiNC. The deviation is small at large pristine SiNC size while it becomes not negligible at a small size. Especially, the calculations predicted the exciton energy of pristine SiNC with 1 to 2 nm diameter size is about 2.5 to 5 eV depend on the theoretical levels applied, however, the experimental PL energy of the pristine SiNC is about 2 eV. Wolkin and Fauchet et al.[25] have shown the calculated PL energy is about 4 eV for 1 nm diameter pristine SiNC.

Hence, it remains to answer whether this deviation is caused by the existence of MECI. In this chapter, we will try to understand the proposed first question with our small cluster model (Si_7H_{12}) and give reasonable explanations of the difference between existing experimental observations and theoretical predictions according to our proposed non-radiative recombination mechanism.

11.1 Methods

11.1.1 Nonadiabatic molecular dynamics simulations

To model the non-radiative recombination process of the silicon clusters, one needs a careful description of the complex electronic wavefunctions as well as nuclear motion. We applied *ab initio* multiple spawning (AIMS) algorithm,[27,28] a hierarchy of approximations that capture the quantum mechanical effect of both electronic and nuclear motions. The nuclear wavefunction is approximated by expanding it in a time-

dependent basis of frozen Gaussian functions[29] while the adiabatic electronic potential energy surfaces (PESs) are calculated on-the-fly at the state averaged complete active space self-consistent field (SA-CASSCF)[30] level of theory. State average two and two electrons in two orbitals active space were employed for SA-CASSCF method. The LANL2DZdp basis set and effective core potentials (ECP)[31] were used in all calculations. These parameters were chosen in part due to a good agreement of the low-lying excited state characters with equation-of-motion coupled cluster with single and double excitations (EOM-CCSD)[32] level of theory using the 6-31G** basis set. Twenty initial nuclear basis were placed on the S_1 state of the clusters with their initial positions and momenta randomly sampled from the ground state vibrational Wigner distribution, with the ground state minimum geometry optimized at MP2/6-31G** level of theory. The nuclear wavefunction is propagated using an adaptive multiple time step algorithm with classical and quantum degrees of freedom integrated by velocity Verlet and second-order Runge-Kutta schemes with a 0.5fs time step.[28] The widths of the frozen Gaussian basis functions were chosen to be 25.0 and 6.0 $Bohr^2$ for silicon and hydrogen atoms, respectively.

11.1.1 Quantum chemical calculations

To gauge the accuracy of the SA-CASSCF PES that was employed in the AIMS calculations, minimal energy conical intersection (MECI) and S_1 minimum (S_1 min) geometries were optimized at SA-CASSCF and multi-state complete active space second-order perturbation (MS-CASPT2) level of theories.[33] The same active space and basis set

were used as that has been employed in the dynamical simulations, and six occupied orbitals besides the active orbitals are correlated for the MS-CASPT2 calculations. Additional single point calculations at multireference configuration interaction with single and double excitations and Davidson correction[34] (MRCI(Q)) level of theory were performed on Frank-Condon (FC), MECI, and S_1 min geometries. Further analysis of the PES was performed by applying single point calculations on *spawning points* (geometries on where an addition basis is spawned to another state in AISM simulations) at SA-CASSCF and MS-CASPT2 level of theories.

All dynamical calculations were performed with FMSMolPro package[28]. MP2, EOM-CCSD, SA-CASSCF, MS-CASPT2 and MRCI(Q) were performed in MolPro.[35-41] MECI and S_1 min geometries were optimized with MolPro in conjunction with CIOpt.[36,42,43]

11.2 Results and Discussion

11.2.1 Dynamical simulation of the silicon cluster

At the FC geometry of the model cluster, the SA2-CASSCF(2/2)/LANL2DZdp level of theory predicts the S_1 vertical excitation energy at 6.606 eV (as comparable to EOM-CCSD/6-31G** level of theory: 6.076 eV). The dynamical simulation shows that 89% of the excited state population decays to the ground state in one picosecond after excitation as shown in Figure 11-1. It is noteworthy that 13 out of 20 individual trajectories transfer more than 98% of the excited state population to the ground state in 350 fs, this time scale indicates a very fast non-radiative decay process. Assume an exponential decay, this

corresponds to a lifetime that is much shorter than the radiative lifetime of the PL emission of silicon nanoparticles (a typical band-to-band recombination happens in 10-100 μs[44,45]). The S_1-S_0 population decay is accompanied by a decreasing of the S_1-S_0 energy gap as shown in Figure 11-2. The averaged S_1-S_0 energy gap falls below 1.64 eV in 100 fs after excitation and oscillates around this value for the rest of the simulation. Individual trajectories are hitting zero S_1-S_0 energy gap indicates the conical intersection may play a role in the non-radiative decay process.

For most of the trajectories (19 out of 20 of our initial nuclear basis), one of the Si-Si bonds undergoes bond breaking process during the dynamical simulation. This suggests that Si-Si bond breaking motion may facilitate the non-radiative recombination of the excited cluster. This hypothesis will be validated in detail in the next section. Figure 11-3(a) shows one Si-Si bond (the one that has been lengthened distinctly compared with the rest of the Si-Si bonds) length as a function of time for the 19 trajectories. In these trajectories, besides the bond breaking process that might be related to the non-radiative recombination process, there are various other dynamical behaviors also have been observed during our simulations. Two out of these 19 trajectories do not transfer population to the ground state during the 1 ps simulation time. In these two trajectories, one trajectory shows the silicon model cluster opens up as shown in the insets of Figure 11-3(b), which indicates that the cluster is trapped into another region of the potential energy surface where the conical intersection is not energetically accessible. However, this will be less likely to happen in the real size NC due to the steric effect of the

surrounding Si atoms. While the other trajectory stays near the Frank-Condon region as shown in Figure 11-3(c), longer simulation time might be required to observe the cluster undergo non-radiative decay. Figure 11-3(d) shows that one of the 19 trajectories undergoes hydrogen transfer process during the simulation. The hydrogen transfer process is shown in Figure 11-3(e). The hydrogen is initially bonded to Si(1), and it transfers to Si(2) in 100 fs after photoexcitation and leaving a bi-radical Si atom behind. Si-Si bond breaking was observed during the simulation in this trajectory as well. However, it looks like the hydrogen transfer process provides the cluster a chance to reach another region of the PES and opens up the cluster as shown in the right one of the insects of Figure 11-3(d). Compare the time scale of population transfer (significant population transfer was observed at 580 fs) in this specific trajectory and the time scale of hydrogen transfer process, it suggests that hydrogen transfer may not be responsible for non-radiative decay, hence Si(1)-Si(2) bond breaking could still be the main mechanism of excited state population transfer. It is also suggested from the time scale that the cluster opens up might be important to reach the conical intersection region of the PES as the Si(1)-Si(2) bond breaking happened much earlier than population transfer. Although hydrogen transfer might not cause non-radiative recombination of the excited cluster, this could be a reasonable way to introduce dangling bond which is a type of known defect that quenches PL.[1,46-50]

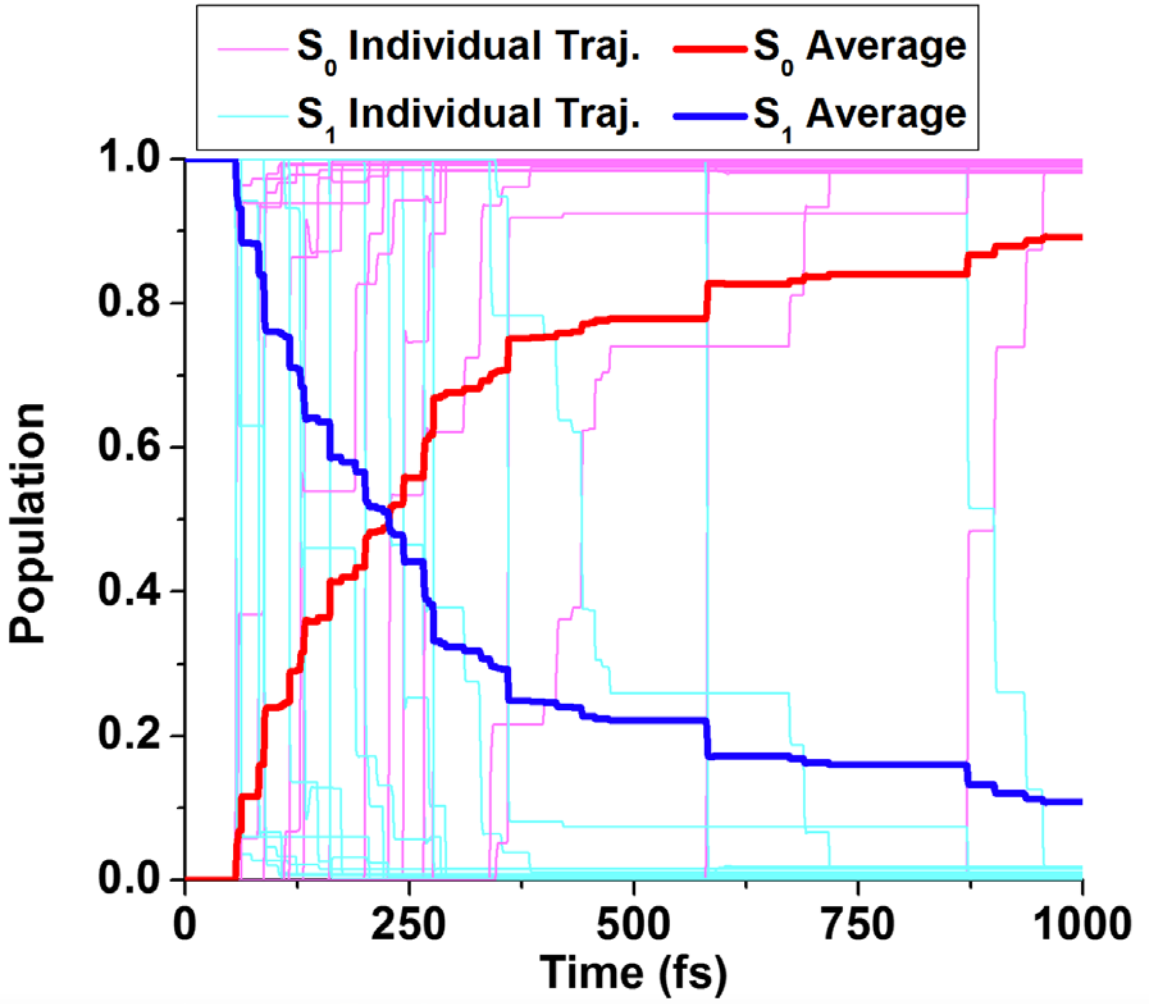


Figure 11-1. Population of S_1 (bold blue) and S_0 (bold red) as a function of time after excitation averaged over all 20 initial nuclear basis functions. Thinner blue and pink lines reflect the population corresponding to the 20 individual basis functions before averaging.

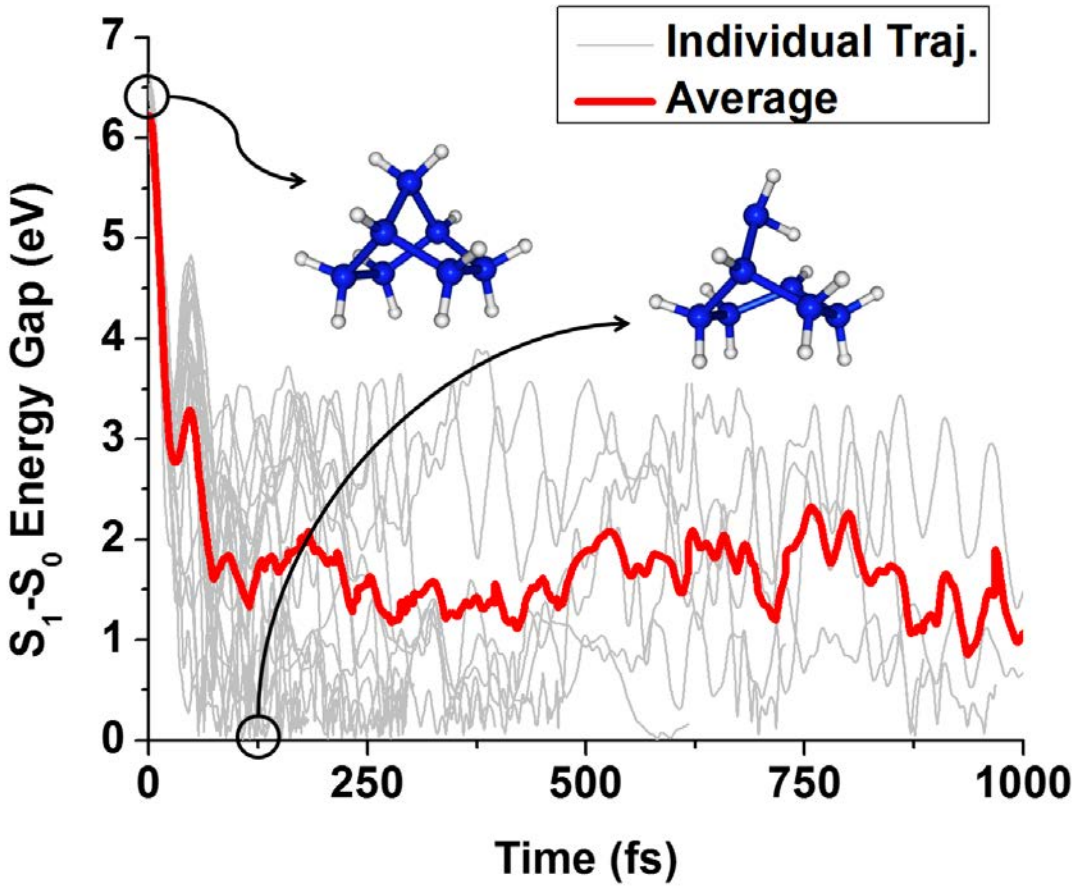


Figure 11-2. Average S_1-S_0 energy gap of silicon cluster as a function of time (bold red) with the energy gaps of 20 initial nuclear basis functions (gray). Snapshots from representative trajectories are shown in the inset, with silicon and hydrogen atoms colored blue and white, respectively.

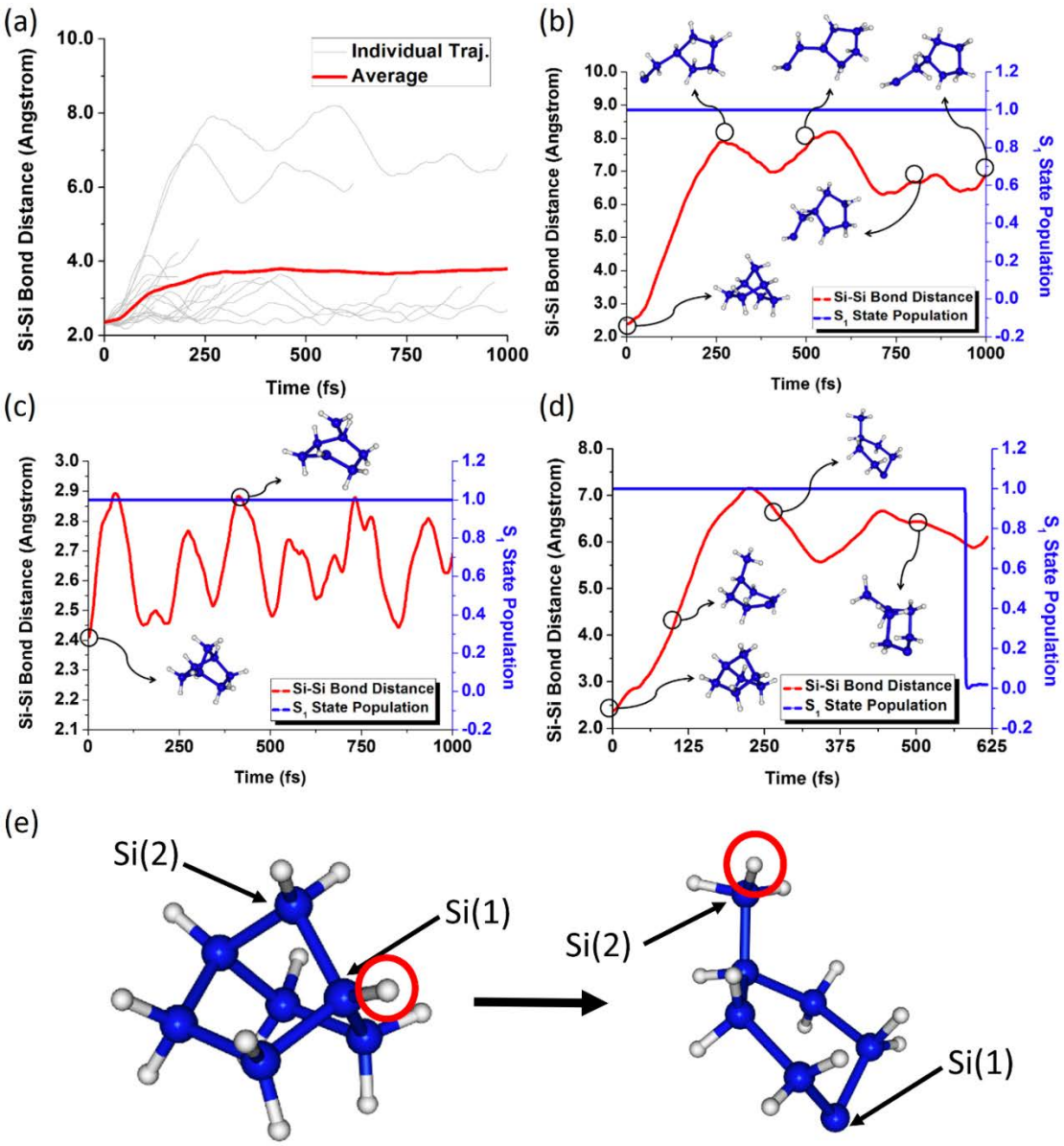


Figure 11-3. (a) Averaged Si-Si bond distance as a function of time (bold red) with bond distance of 19 initial nuclear basis functions (gray), (b) One trajectory that undergoes cluster opening process and no population transfer was observed during the simulation with the snap shots from the dynamics, (c) One trajectory stays near Frank-Condon region in the whole simulation with the snap shots from the dynamics, (d) One trajectory undergoes hydrogen transfer process with the snap shots from the dynamics, (e) indicates

Figure 11-3 (cont'd). the hydrogen transfer process in (d) with the transferred hydrogen shown in red circle. (b),(c) and (d) graphs show the Si-Si bond distance as a function of time (red) as well as S_1 state population as a function of time (blue) for this specific trajectory.

Besides the bond breaking as a main dynamical process associated with non-radiative decay as we observed in these AIMS simulations, another trajectory in the 20 initial nuclear basis undergoes non-radiative decay through the re-hybridization of one of the Si atoms. The H-SiSiSi pyramidalization angles as a function of time are shown in Figure 11-4 with the inserted pictures show the corresponding two pyramidalization angles. The H-SiSiSi pyramidalization angle is defined in eq. (11-1),

$$\theta = \left| \arcsin \frac{\vec{C} \bullet \vec{R}_{Si(X)H(1or2)}}{\left\| \vec{C} \right\| \left\| \vec{R}_{Si(X)H(1or2)} \right\|} \right| \qquad (11\text{-}1)$$

where $\vec{C} = \vec{R}_{Si(X)Si(Y)} \times \vec{R}_{Si(X)Si(Z)}$, $\vec{R}_{AB}$ is a vector from atom A to atom B, with $\bullet$ and $\times$ indicating the scalar and vector products of the two vectors respectively. It is observed that the two Si-H bonds rotate during the simulation. The pyramidalization angles H(1)-SiSiSi and H(2)-SiSiSi decrease 24.5 and 27.0 degrees as shown in Figure 11-4. This suggests the Si(X) atom changes its coordination from tetrahedral to square planar type, which corresponds to sp^3 and sp^3d^2 hybridization types respectively. The non-radiative decay process corresponds to this re-hybridization mechanism is very fast (significant population transfer was observed around 115 fs).

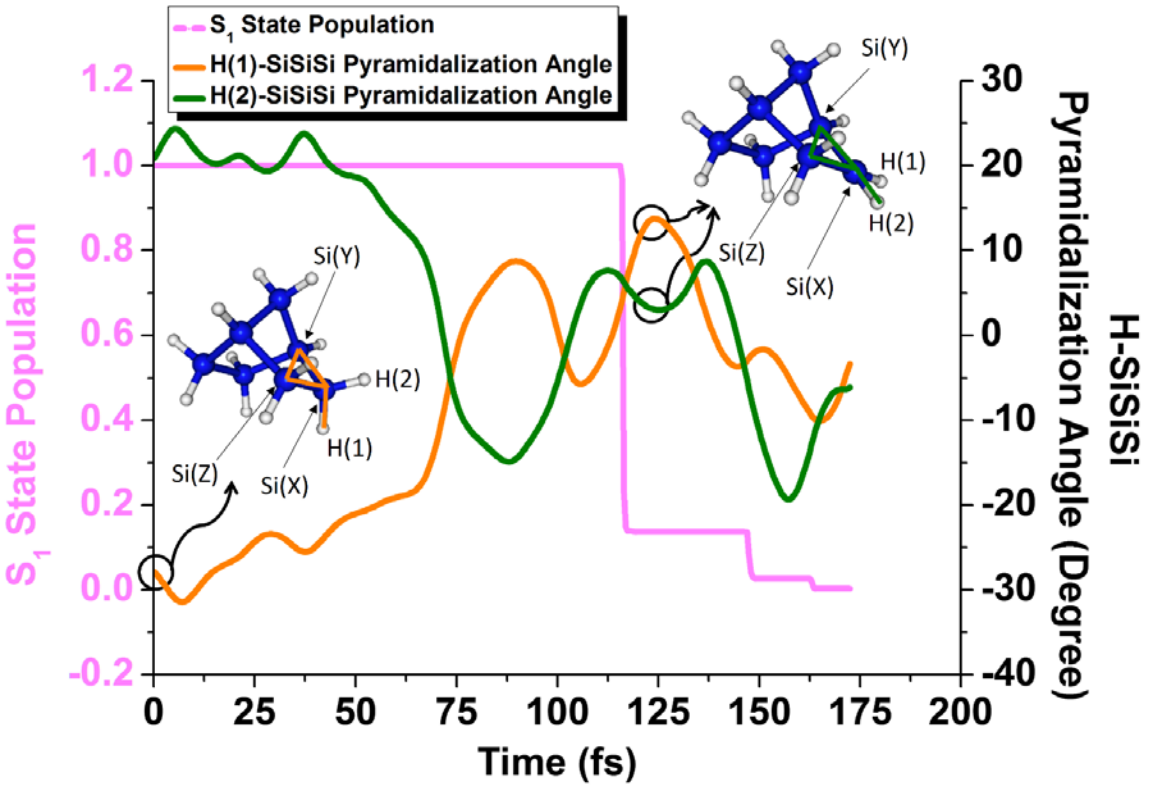


Figure 11-4. Two H-SiSiSi pyramidalization angle as a function of time with snapshots from dynamics and indicating corresponding atoms associated with the pyramidalization angle, H(1) and H(2) are the labels for the two different hydrogen atoms. Si(X), Si(Y) and Si(Z) are three silicon atoms associated with the definition of pyramidalization angle. Noticed that the two hydrogen atoms rotated during the simulation. The graph also shows the S_1 state population as a function of time (pink color).

11.2.2 Potential energy surface analysis of the silicon cluster

To further investigate the possibility that conical intersection plays a role in the non-radiative decay process, we applied SA-CASSCF and MS-CASPT2 theoretical levels to explore the PESs. The minimal energy conical intersection (MECI) geometries optimized

with both methods have one broken Si-Si bond. Geometries similar to the optimized MECIs were observed in the dynamical simulations. We will note this type of MECI as bond dissociation MECI (BD MECI) in the following text. This validates our hypothesis in the previous section that Si-Si bond breaking may facilitate the non-radiative decay process. The energies of SA-CASSCF and MS-CASPT2 optimized geometries are computed at SA-CASSCF, MS-CASPT2 and MRCI(Q) theoretical levels to account for the dynamical correlations. These results are presented in Table 11-1 and Figure 11-5.

Optimized BD MECIs at both SA-CASSCF and MS-CASPT2 theoretical levels are energetically accessible after vertical excitation at Frank-Condon (FC) geometry. However, despite significant effort towards searching the S_1 PES, we were not able to locate a stable minimum. In addition, no geometries that have lower S_1 energy than BD MECI was found. This suggests that the MECI may be a peaked intersection, and therefore the MECI is the lowest energy point in the local region of the PES. Hence efficient non-radiative decay through the MECI is expected. Most individual trajectories have fast population transfer process as noted in the previous section supports the conclusion that BD MECI is a peaked intersection. At SA-CASSCF optimized BD MECI geometry, the energy is 3.87, 4.06 and 3.94 eV above the ground state of FC geometry at SA-CASSCF, MS-CASPT2 and MRCI(Q) theoretical levels, respectively. This indicates a 0.29 eV or 0.07 eV error compared between SA-CASSCF and MS-CASPT2 or MRCI(Q) theoretical levels. This amount of discrepancy is common for SA-CASSCF method. Single point calculations on SA-CASSCF optimized BD MECI predicts 0.35 eV and 0.20 eV

S_1-S_0 energy gaps at MS-CASPT2 and MRCI(Q) theoretical levels. These energy gaps are not perfect, however, is under a reasonable range. The S_1 state energies of the MS-CASPT2 optimized BD MECI geometry predicted at SA-CASSCF, MS-CASPT2 and MRCI(Q) theoretical levels are 3.90, 3.89 and 3.90 eV above the ground state minimum. With the 0.01 and 0.17 eV S_1-S_0 energy gaps at SA-CASSCF and MRCI(Q) levels respectively. These results validate our SA-CASSCF PES employed in AIMS dynamical simulations. The energy level of BD MECI is consistent with the experiment fact that no PL maximum shorter than 320 nm (corresponds to 3.875 eV) is observed for the ultra-small SiNC as conical intersections are energetically accessible by excitons that have energy higher than 320 nm. To further investigate the accuracy of SA-CASSCF PES, single point calculations are performed at the spawning points. Figure 11-6 shows the S_1-S_0 energy gap and S_1 excitation energy correlations between SA-CASSCF and MS-CASPT2 theoretical levels. The roughly linear relation on both figures further confirms that the SA-CASSCF PES is qualitatively correct.

Table 11-1. The S_0 and S_1 energies (in eV) at three important points on the PES of the silicon cluster calculated at the SA- CASSCF, MS-CASPT2, and MRCI(Q) levels of theory. S_1-S_0 MECI geometries are optimized at both the SA-CASSCF and MS-CASPT2 levels of theory.

Geometry	Optimization Method	SA-CASSCF S_0 Energy	SA-CASSCF S_1 Energy	MS-CASPT2 S_0 Energy	MS-CASPT2 S_1 Energy	MRCI(Q) S_0 Energy	MRCI(Q) S_1 Energy
FC geometry	KS	0	6.61	0	6.19	0	6.31
BD S_1-S_0 CI	SA-CASSCF	3.87	3.87	3.71	4.06	3.74	3.94
	MS-CASPT2	3.89	3.90	3.87	3.89	3.73	3.90
SqP S_1-S_0 CI	SA-CASSCF	4.21	4.21	4.14	4.52	4.28	4.46
	MS-CASPT2	4.19	4.30	4.29	4.30	4.28	4.31

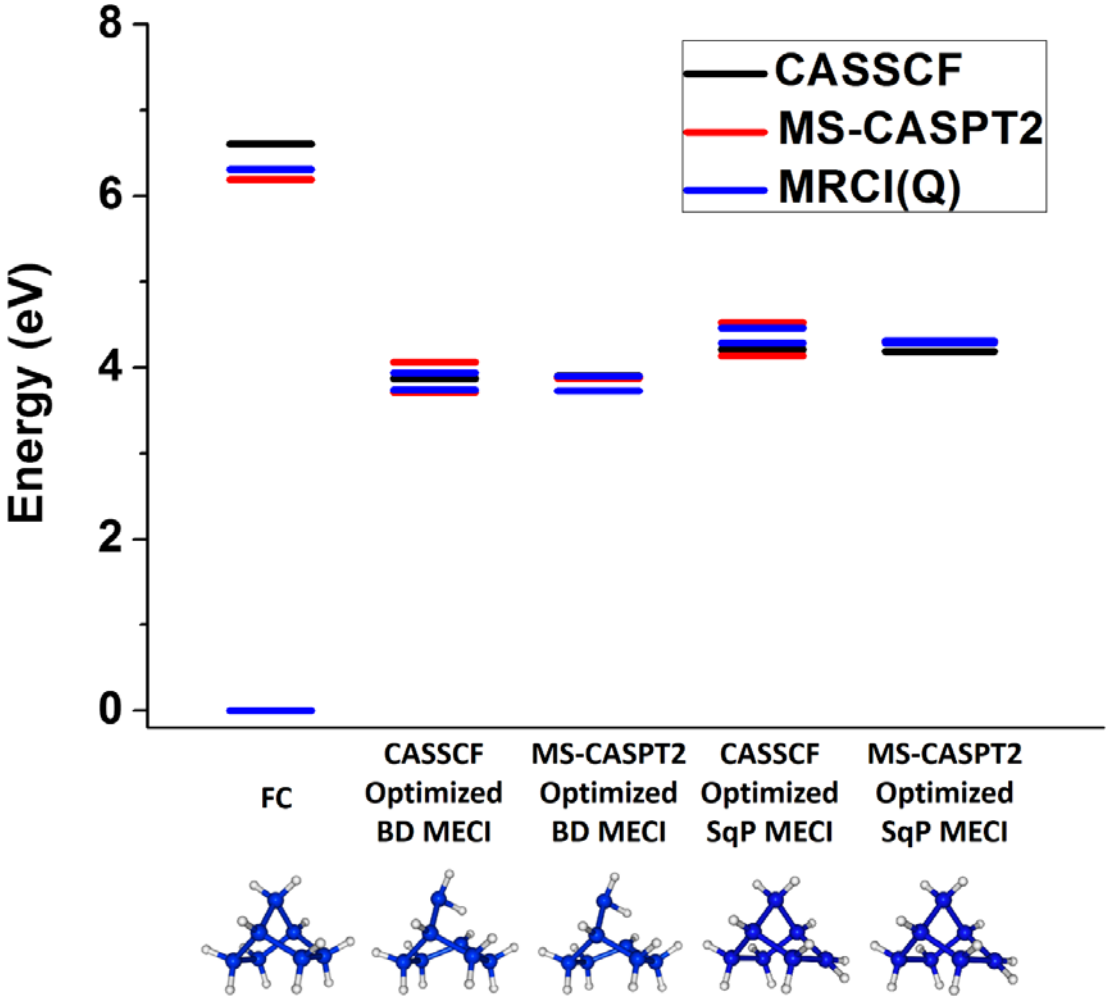


Figure 11-5. S_0 and S_1 energies of SA-CASSCF and MS-CASPT2 optimized bond breaking (BD) and square splanar (SqP) minimal energy conical intersections of silicon cluster as calculated at the SA-CASSCF, MS-CASPT2, and MRCI(Q) levels of theory (black, red, and blue bars, respectively). Geometries are shown as insets. Energies at the Franck–Condon (FC) geometry are shown for comparison.

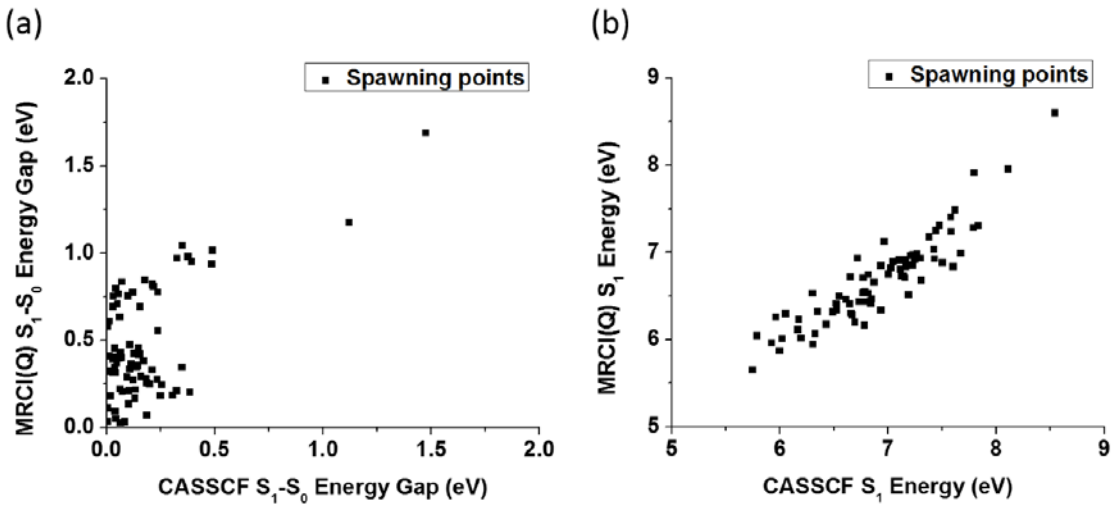


Figure 11-6. (a) The S_1-S_0 energy gap correlation between SA-CASSCF and MRCI(Q) theoretical levels with 6 orbitals correlated in the MRCI(Q) calculations. (b) The adiabatic S_1 excitation energy correlation between SA-CASSCF and MRCI(Q) theoretical levels.

We also find a MECI geometry corresponds to the re-hybridization mechanism. This type of MECI has one of the Si atoms forms a square planar type of coordination and will be noted as square planar MECI (SqP MECI) in the following discussion. SqP MECI geometry is optimized with both SA-CASSCF and MS-CASPT2 methods. The results are shown in Table 11-1 and Figure 11-5. Optimized SqP MECIs at both SA-CASSCF and MS-CASPT2 theoretical levels are energetically accessible after vertical excitation at Frank-Condon (FC) geometry. Single point calculations performed at SA-CASSCF, MS-CASPT2 and MRCI(Q) theoretical levels predict 4.21, 4.52 and 4.46 eV excitation energies at SA-CASSCF optimized SqP MECI geometry. This implies an error of 0.31 and 0.25 eV comparing SA-CASSCF with MS-CASPT2 and MRCI(Q) methods. The SqP MECI optimized

at MS-CASPT2 theoretical level shows S_1 state energy at 4.30 eV, 4.30 eV and 4.31 eV above the ground state minimum computed with SA-CASSCF, MS-CASPT2 and MRCI(Q) methods respectively. These results have good agreement with those calculated on SA-CASSCF optimized SqP MECI geometry, which further supports our simulation that involves this distinct non-radiative decay mechanism. We were not able to locate the excited state minimum at this region of the PES. Combined with the fact that the population transfer is fast, it suggests that SqP MECI is also a peaked MECI.

To understand the re-hybridization mechanism, we investigated the electronic structures at two important geometries from that specific AIMS trajectory. The first geometry corresponds to the initial geometry which is sampled from ground state vibrational Wigner distribution (we will denote this geometry as Initial Geometry (IG) in the following text and figures). The other geometry is the first spawning geometry (as defined in the previous section, we will denote the First Spawning Geometry as FSG in the following text and figures) in that trajectory. The FSG was hit around 115 fs in the simulation. The electronic structures of IG and FSG are computed at SA-CASSCF theoretical level with the hole and particle orbitals shown in Figure 11-7.

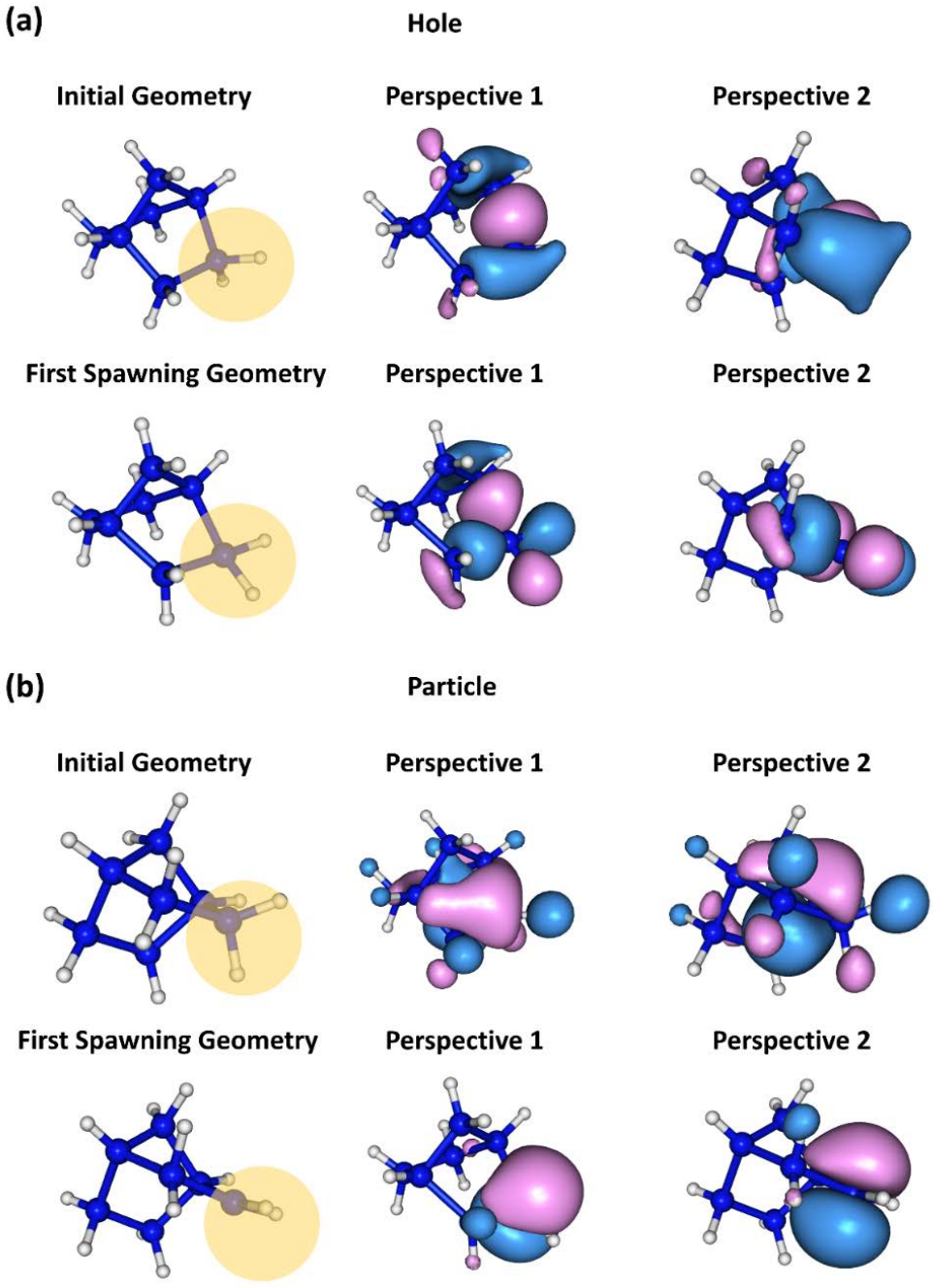


Figure 11-7. Hole and particle orbitals at IG and FSG are shown in (a) and (b) respectively, with the fist and second columns show the perspective 1 and 2 at IG and FSG, the yellow shade area indicates the local area where re-hybridization happens. The hole and particle orbitals correspond to the HOMO and LUMO SA-CASSCF natural orbitals.

The hole and particle orbitals correspond to the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) of SA-CASSCF natural orbitals. SA-CASSCF natural orbitals are generated by diagonalizing the averaged S_1 and S_0 first order reduced density matrices. Figure 11-7 shows two perspectives of the hole and particle orbitals at IG and FSG. For the purpose of clarity, the yellow shade indicates the local area where re-hybridization happens. By investigating the evolving of the electronic structures during the dynamics, we hypothesize that this type of mechanism is possibly initialized by a vertical excitation involves a charge transfer character from Si-Si σ bonding orbital to H s orbitals. The particle orbital at IG has an anti-bonding interaction character between Si p orbital and H s orbitals as shown in Figure 11-8(a). This anti-bonding interaction drives the rotation of the Si-H bonds. As the anti-bonding interaction is reduced when the coordination type of the Si atom changes from tetrahedral to square planar, the excited state is stabilized. This mechanism is illustrated in Figure 11-8(b) and (c). Figure 11-8(b) shows the tetrahedral and square planar coordination type of the Si atom at IG and FSG respectively. Red arrows in Figure 11-8(b) indicates the rotation directions of the Si-H bonds. Figure 11-8(c) shows the simplified orbital characters to clarify the driving force of the Si-H bonds rotation. We hypothesize that this type of mechanism may happen for the Si atoms bonded to two or more H atoms. The flexibility of the Si-H bonds allows them to rotate around and form the SqP MECI. While those Si atoms bonded to less than two H atoms, the strong steric effect of the Si network in SiNC will possibly prohibit the rotation of the Si-Si bonds in the cluster and

hence Si-Si bond breaking may become the mechanism for non-radiative decay process. Hence re-hybridization mechanism is more likely a surface NC photophysical process.

Partially occupied natural orbitals (PONOs) at FC, BD MECI and SqP MECI geometries computed at SA-CASSCF theoretical level are presented in Figure 11-9. The PONOs are generated by diagonalizing the state-averaged S_1-S_0 density matrices at all geometries. At FC geometry, the orbitals are a mixture of Si p, Si d and H s orbitals. Mulliken population analysis shows the transition can mostly be described by Si-Si σ bonding orbitals to Si d orbitals. At BD MECI geometry, one orbital has Si p nonbonding orbital character due to the broken bond at the BD MECI geometry; the other orbital is characterized by two Si p orbitals form a σ anti-bonding orbital character. The occupation number indicates that at BD MECI geometry, an electron from the σ anti-bonding orbital transfers to the Si p non-bonding orbital. At SqP MECI geometry, the two orbitals correspond to Si d orbital and Si p orbital.

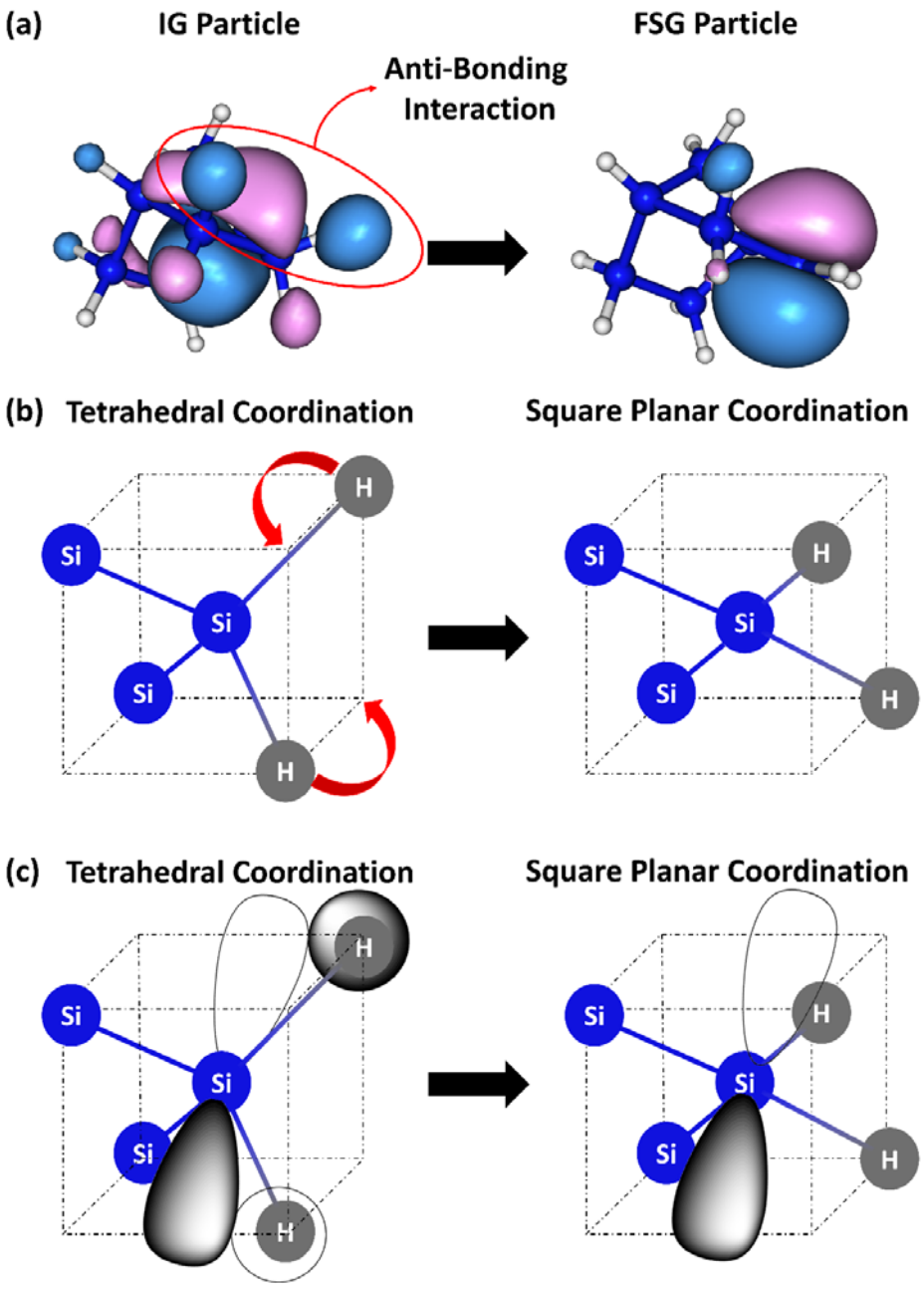


Figure 11-8. (a) The hole and particle orbitals at IG and FSG from with the red circle indicates the anti-bonding interaction between the Si p and H s orbitals. The hole and particle orbitals are the same as those shown in Figure 11-7(c). (b) A schematic graph shows the coordination of the Si atom, these atoms correspond to the yellow shade area as shown in Figure 11-7. The dashed cubic is used as a reference to clarify the positions

Figure 11-8 (cont'd). of each atoms. Red arrow on the left panel shows the direction of rotation of the Si-H bonds. (c) A simplified cartoon to illustrate the anti-bonding character between Si p and H s orbitals at tetrahedral coordination as shown in the left panel. While this anti-bonding interaction does not exist in the right panel which shows the square planer coordination.

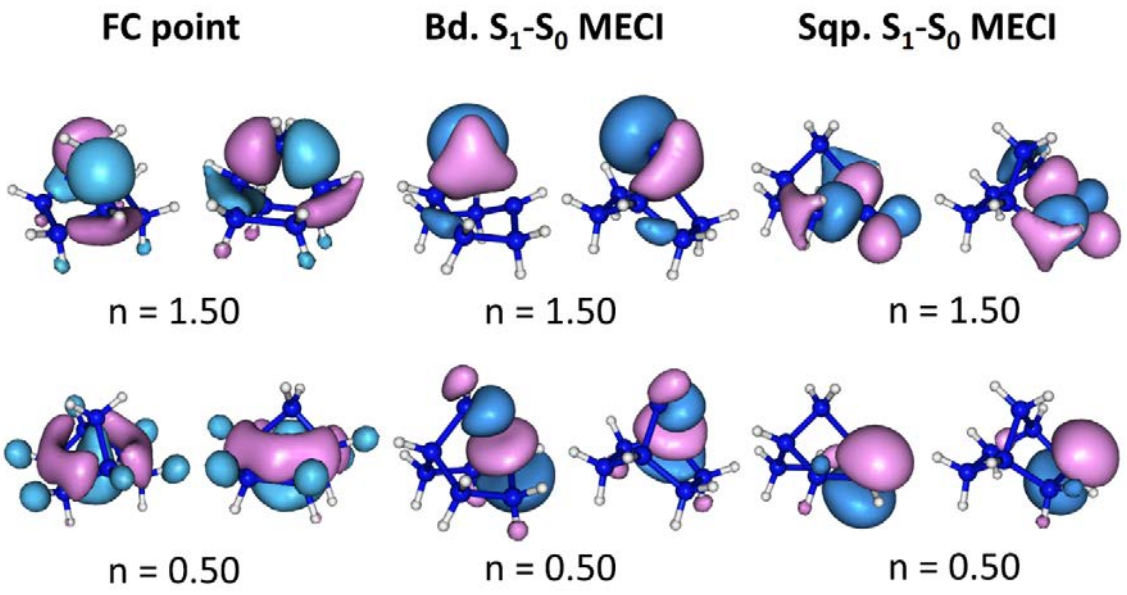


Figure 11-9. Partially occupied SA-CASSCF natural orbitals of silicon cluster on two different views at FC geometry (left), BD MECI (middle) and SqP MECI (right) are presented with occupation numbers. All natural orbitals are computed from the state-averaged S_1–S_0 density matrix.

11.3 Conclusions

In this work, we have applied AIMS to simulate the excited state dynamics of pristine

SiNC to investigate the possible mechanism for non-radiative decay process. We found two possible mechanisms. The first one corresponds to Si-Si bond breaking process. It is demonstrated from our simulations that the MECI geometry is reached by local nuclear deformation involving breaking one Si-Si bond, this BD MECI has energy level around 3.8 eV above the ground state minimum. The other mechanism involves a re-hybridization of the surface Si atom bonded to more than one hydrogen atoms with the coordination type changes from tetrahedral to square planar. This photophysical process is driven by an anti-bonding interaction between the Si p orbital and H s orbitals. The associated SqP MECI has energy level around 4.5 eV above the ground state minimum. Because both MECI is reached by local nuclear deformation, the energy levels of these MECIs are believed to be insensitive to the NC size. This assumption has been discussed in detail in Chapter 10.

The energy level of the BD MECI and SqP MECI correspond to 325 and 275 nm PL peak, which is in the ultraviolet range. Hence they have no effect on the PL of the pristine SiNC with the visible range band gap. This is also consistent with the fact that no existing PL peak shorter than 320nm has been observed despite the various efforts in synthesizing the ultra-small SiNCs. Excitons of smaller size SiNC with band gap higher than 3.8 eV as predicted by quantum confinement effect will be quenched by the BD MECI. However, various other types of Si-Si bonds with a different number of H atoms or Si atoms bonded to the Si atom exist in the SiNC; it is important in the future to investigate different types of Si-Si bonds and associated energy levels of these BD MECIs.

APPENDIX

APPENDIX

Table 11-S1. The absolute energies of the FC and various optimized geometries of the S_0/S_1 MECIs at the SA-CASSCF level of theory (with active space and basis sets as described in the text).

SA-CASSCF	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point (MP2)	-33.44213702	-33.19936611
SA-CASSCF optimized BD S_0/S_1 MECI	-33.30001128	-33.30000696
MS-CASPT2 optimized BD S_0/S_1 MECI	-33.29928596	-33.29873547
SA-CASSCF optimized SqP S_0/S_1 MECI	-33.28745160	-33.28741899
MS-CASPT2 optimized SqP S_0/S_1 MECI	-33.28829552	-33.28411057

Table 11-S2. The absolute energies of the FC and various optimized geometries of the S_0/S_1 MECIs at the MS-CASPT2 level of theory (with active space and basis sets as described in the text).

MS-CASPT2	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point (MP2)	-33.5998520	-33.3722791
SA-CASSCF optimized BD S_0/S_1 MECI	-33.4635762	-33.4505864
MS-CASPT2 optimized BD S_0/S_1 MECI	-33.4574604	-33.4567455
SA-CASSCF optimized SqP S_0/S_1 MECI	-33.4478655	-33.4336377
MS-CASPT2 optimized SqP S_0/S_1 MECI	-33.4422031	-33.4419929

Table 11-S3. The absolute energies of the FC and various optimized geometries of the S_0/S_1 MECIs at the MRCI(Q) level of theory (with active space and basis sets as described in the text).

MRCI(Q)	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point (MP2)	-33.60724730	-33.37167601
SA-CASSCF optimized BD S_0/S_1 MECI	-33.46729463	-33.46180294
MS-CASPT2 optimized BD S_0/S_1 MECI	-33.46744360	-33.46267053
SA-CASSCF optimized SqP S_0/S_1 MECI	-33.46967542	-33.46325990
MS-CASPT2 optimized SqP S_0/S_1 MECI	-33.46963981	-33.46851120

Table 11-S4. The absolute energies of the FC geometry at the EOM-CCSD level of theory.

MRCI(Q)	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point (MP2)	-2029.95135163	-2029.95135163

Optimized Geometries (In Angstrom)

Optimized FC geometry

Si	19.74234072	17.32285985	20.46703571
Si	17.75016860	16.26712121	19.74497106
Si	16.28456761	18.08745344	19.29928651
Si	17.59173979	19.99318384	19.81502261
Si	19.24903090	20.19240629	18.13585537
Si	20.71020010	18.36718180	18.57543833
Si	18.87931525	19.19296789	21.63272362
H	20.80899889	17.44581166	17.40784110
H	22.08237089	18.86187013	18.88515076
H	19.98022698	21.48557184	18.26395619
H	18.67670401	20.12407834	16.76128373
H	16.80109126	21.23548922	20.04777836
H	15.82083676	18.08609622	17.88261294
H	15.08054867	18.01451799	20.17567158
H	17.18950404	15.39609231	20.81745145
H	17.96413943	15.42391416	18.53464130
H	20.67179076	16.43251343	21.21935351
H	19.93642359	20.14965968	22.06745423
H	18.05627176	18.81803668	22.81716863

SA-CASSCF Optimized BD S_0/S_1 MECI geometry

Si	19.73252643	17.56119212	20.57065790
Si	17.87408469	16.34040108	19.82686596
Si	16.15053199	17.91236480	19.50094193
Si	17.16628253	20.01183630	19.42459612
Si	19.01567473	20.27080249	18.02506570
Si	20.63636927	18.70954407	18.73383674
Si	18.82724380	19.14604800	22.04607285
H	20.97046936	17.81350532	17.60391524
H	21.85847343	19.44022515	19.13098982
H	19.50101881	21.65902310	18.16164496
H	18.62404286	20.04010972	16.61748797
H	16.34961876	21.19964680	19.72020445

H	15.42956842	17.66693677	18.23096276
H	15.17352450	17.90931804	20.60592721
H	17.44406209	15.34674099	20.83274254
H	18.13325294	15.63707413	18.55017323
H	20.73145013	16.67616367	21.21001711
H	19.85387989	19.78044785	22.91585109
H	18.48979349	20.26787270	21.03618026

MS-CASPT2 Optimized BD S_0/S_1 MECI geometry

Si	19.72948504	17.58832655	20.55553577
Si	17.87226791	16.30518240	19.87657340
Si	16.16114472	17.88551883	19.48862089
Si	17.16127579	19.99653842	19.40398754
Si	19.04024278	20.33827989	18.04584624
Si	20.61691179	18.68728510	18.66829986
Si	18.82206224	19.15315440	22.05919623
H	20.84219801	17.76057741	17.54690245
H	21.88301345	19.34503071	19.01989779
H	19.55238607	21.68662973	18.33068666
H	18.67229065	20.25923119	16.62545278
H	16.29210680	21.15510918	19.63772040
H	15.50757639	17.64229178	18.19540447
H	15.15509504	17.90486605	20.55693370
H	17.46060914	15.38290046	20.94637253
H	18.13521013	15.54315320	18.64528038
H	20.75300468	16.70917139	21.18608721
H	19.82369494	19.79196378	22.92731357
H	18.48129252	20.25404261	21.02802198

SA-CASSCF Optimized SqP S_0/S_1 MECI geometry

Si	19.70790975	17.34037907	20.52009386
Si	17.73154459	16.03863486	19.83527429
Si	16.35488710	18.05180790	19.43855758

Si	17.60524568	19.98475272	19.87360823
Si	19.20250459	20.16805913	18.16680998
Si	20.59340396	18.30291887	18.55957087
Si	18.89870267	19.21693132	21.66799094
H	20.56765252	17.35361595	17.42544063
H	21.99239006	18.73380782	18.77369313
H	19.97928946	21.41965033	18.29858465
H	18.59254573	20.12955077	16.82060822
H	16.78970308	21.19724177	20.10174331
H	15.84656672	17.96705483	18.04967140
H	15.19421196	17.93453305	20.35329793
H	16.51962564	15.23650867	19.38061692
H	18.52447977	14.76578372	20.06871857
H	20.67279605	16.47200793	21.23194552
H	19.96723123	20.15648824	22.07198754
H	18.09142961	18.85868273	22.85379845

MS-CASPT2 Optimized SqP S_0/S_1 MECI geometry

Si	19.72268166	17.35374207	20.52651673
Si	17.72212072	16.00293907	19.83144999
Si	16.34256762	18.05955510	19.43645155
Si	17.61172436	19.98890002	19.87835967
Si	19.20592304	20.16912973	18.15664238
Si	20.60382337	18.30123434	18.54587445
Si	18.89970319	19.22382337	21.68542475
H	20.56704293	17.35342563	17.42637734
H	21.98964195	18.73531264	18.77594168
H	19.97698301	21.41644593	18.30235136
H	18.59146746	20.12905403	16.82374551
H	16.78713700	21.19176933	20.10115116
H	15.85294851	17.96401349	18.05097433
H	15.19654428	17.93431300	20.35534719
H	16.50246939	15.24486151	19.37546521
H	18.53931760	14.77935789	20.07278137
H	20.66516010	16.46963442	21.23087875
H	19.96657785	20.15517694	22.06583626
H	18.08828611	18.85572118	22.85044233

REFERENCES

REFERENCES

(1) Ledoux, G.; Gong, J.; Huisken, F.; Guillois, O.; Reynaud, C. *Appl Phys Lett* **2002**, *80*, 4834.

(2) Kim, T. Y.; Park, N. M.; Kim, K. H.; Sung, G. Y.; Ok, Y. W.; Seong, T. Y.; Choi, C. J. Appl Phys Lett 2004, 85, 5355.

(3) Delley, B.; Steigmeier, E. F. *Phys Rev B* **1993**, *47*, 1397.

(4) English, D. S.; Pell, L. E.; Yu, Z. H.; Barbara, P. F.; Korgel, B. A. *Nano Lett* **2002**, *2*, 681.

(5) Pavesi, L.; Dal Negro, L.; Mazzoleni, C.; Franzo, G.; Priolo, F. *Nature* **2000**, *408*, 440.

(6) Ogut, S.; Chelikowsky, J. R.; Louie, S. G. *Phys Rev Lett* **1997**, *79*, 1770.

(7) Holmes, J. D.; Ziegler, K. J.; Doty, R. C.; Pell, L. E.; Johnston, K. P.; Korgel, B. A. J Am Chem Soc 2001, 123, 3743.

(8) Belomoin, G.; Therrien, J.; Nayfeh, M. *Appl Phys Lett* **2000**, *77*, 779.

(9) Matsumoto, T.; Arata, G.; Nair, S. V.; Masumoto, Y. *Jpn J Appl Phys 1* **1999**, *38*, 589.

(10) Rogozhina, E. V.; Eckhoff, D. A.; Gratton, E.; Braun, P. V. *J Mater Chem* **2006**, *16*, 1421.

(11) Nayfeh, M. H.; Barry, N.; Therrien, J.; Akcakir, O.; Gratton, E.; Belomoin, G. Appl Phys Lett 2001, 78, 1131.

(12) Belomoin, G.; Therrien, J.; Smith, A.; Rao, S.; Twesten, R.; Chaieb, S.; Nayfeh, M. H.; Wagner, L.; Mitas, L. Appl Phys Lett 2002, 80, 841.

(13) Choi, J.; Wang, N. S.; Reipa, V. *Langmuir* **2007**, *23*, 3388.

(14) Kim, T. W.; Cho, C. H.; Kim, B. H.; Park, S. J. *Appl Phys Lett* **2006**, *88*.

(15) Alkis, S.; Okyay, A. K.; Ortac, B. *J Phys Chem C* **2012**, *116*, 3432.

(16) Dasog, M.; Yang, Z. Y.; Regli, S.; Atkins, T. M.; Faramus, A.; Singh, M. P.; Muthuswamy, E.; Kauzlarich, S. M.; Tilley, R. D.; Veinot, J. G. C. Acs Nano 2013, 7, 2676.

(17) de Boer, W. D. A. M.; Timmerman, D.; Dohnalova, K.; Yassievich, I. N.; Zhang, H.; Buma, W. J.; Gregorkiewicz, T. Nat Nanotechnol 2010, 5, 878.

(18) Eckhoff, D. A.; Sutin, J. D. B.; Clegg, R. M.; Gratton, E.; Rogozhina, E. V.; Braun, P. V. J Phys Chem B 2005, 109, 19786.

(19) Intartaglia, R.; Bagga, K.; Genovese, A.; Athanassiou, A.; Cingolani, R.; Diaspro, A.; Brandi, F. Phys Chem Chem Phys 2012, 14, 15406.

(20) Wilcoxon, J. P.; Samara, G. A.; Provencio, P. N. *Phys Rev B* **1999**, *60*, 2704.

(21) Mastronardi, M. L.; Maier-Flaig, F.; Faulkner, D.; Henderson, E. J.; Kübel, C.; Lemmer, U.; Ozin, G. A. Nano Lett 2011, 12, 337.

(22) Hill, N. A.; Whaley, K. B. *Phys Rev Lett* **1995**, *75*, 1130.

(23) Wang, L. W.; Zunger, A. *J Phys Chem-Us* **1994**, *98*, 2158.

(24) Delerue, C.; Allan, G.; Lannoo, M. *Phys Rev B* **1993**, *48*, 11024.

(25) Wolkin, M. V.; Jorne, J.; Fauchet, P. M.; Allan, G.; Delerue, C. Phys Rev Lett 1999, 82, 197.

(26) Luppi, E.; Iori, F.; Magri, R.; Pulci, O.; Ossicini, S.; Degoli, E.; Olevano, V. Phys Rev B 2007, 75.

(27) Ben-Nun, M.; Quenneville, J.; Martinez, T. J. *J Phys Chem A* **2000**, *104*, 5161.

(28) Levine, B. G.; Coe, J. D.; Virshup, A. M.; Martinez, T. J. *Chem Phys* **2008**, *347*, 3.

(29) Heller, E. J. *J Chem Phys* **1981**, *75*, 2923.

(30) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. *Chem Phys* **1980**, *48*, 157.

(31) Wadt, W. R.; Hay, P. J. *J Chem Phys* **1985**, *82*, 284.

(32) Stanton, J. F.; Bartlett, R. J. *The Journal of chemical physics* **1993**, *98*, 7029.

(33) Finley, J.; Malmqvist, P. A.; Roos, B. O.; Serrano-Andres, L. Chem Phys Lett 1998, 288, 299.

(34) Langhoff, S. R.; Davidson, E. R. *Int J Quantum Chem* **1974**, *8*, 61.

(35) Knowles, P. J.; Werner, H. J. *Theor Chim Acta* **1992**, *84*, 95.

(36) Shiozaki, T.; Gyorffy, W.; Celani, P.; Werner, H. J. *J Chem Phys* **2011**, *135*.

(37) Hampel, C.; Peterson, K. A.; Werner, H. J. *Chem Phys Lett* **1992**, *190*, 1.

(38) Werner, H. J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schutz, M. Wires Comput Mol Sci 2012, 2, 242.

(39) Celani, P.; Werner, H. J. *J Chem Phys* **2000**, *112*, 5546.

(40) Korona, T.; Werner, H. J. *J Chem Phys* **2003**, *118*, 3006.

(41) Werner, H. J.; Knowles, P. J. *The Journal of chemical physics* **1985**, *82*, 5053.

(42) Levine, B. G.; Coe, J. D.; Martinez, T. J. *J Phys Chem B* **2008**, *112*, 405.

(43) Busch, T.; Degli Esposti, A.; Werner, H. J. *J Chem Phys* **1991**, *94*, 6708.

(44) Littau, K. A.; Szajowski, P. J.; Muller, A. J.; Kortan, A. R.; Brus, L. E. J Phys Chem-Us 1993, 97, 1224.

(45) Wilson, W. L.; Szajowski, P. F.; Brus, L. E. *Science* **1993**, *262*, 1242.

(46) Batstone, J. L.; Tischler, M. A.; Collins, R. T. *Appl Phys Lett* **1993**, *62*, 2667.

(47) Kim, B. H.; Cho, C. H.; Kim, T. W.; Parka, N. M.; Sung, G. Y.; Park, S. J. Appl Phys Lett 2005, 86.

(48) Lannoo, M.; Delerue, C.; Allan, G. *J Lumin* **1993**, *57*, 243.

(49) Prokes, S. M.; Carlos, W. E.; Bermudez, V. M. *Appl Phys Lett* **1992**, *61*, 1447.

(50) Brandt, M. S.; Stutzmann, M. *Appl Phys Lett* **1992**, *61*, 2569.

Chapter 12

Silicon Nanocrystal with Hydroxide Group at Surface

"I'd rather be silly than lazy"
-James C. Maxwell (1831-1879)

In Chapter 11 we have studied the excited state dynamics of H passivated pristine SiNC (will be denoted as pristine SiNC in the following text), the minimal energy conical intersection (MECI) energy level locates around 3.8 eV (325 nm) above the ground state minimum. And hence the MECI has little effect on the PL spectrum of the pristine SiNCs with visible range band gap. The stability of pristine SiNC in water and air is poor and thus it is believed that hydroxide defect will be introduced at the surface of the SiNC.[1-5] In contrast to the oxide defects that has been investigated in the previous chapters (Chapter 7-10), experiments show that the PL intensity increases as the SiNC being passivated by OH groups from the SiNC with dangling bonds defects.[2,5] Due to the fact that there exists various synthetic methods[6-12] and wide applications of water soluble or water dispersible SiNC,[7,10,12-23] it is important to investigate the energy level of hydroxide group introduced conical intersection and its role in photophysics of SiNC.

12.1 Methods

12.1.1 Nonadiabatic molecular dynamics simulations

We applied ab initio multiple spawning (AIMS) algorithm[24,25] to simulate the excited state dynamics of our small silanol cluster model ($Si_7H_{12}O$). In AIMS hierarchy, the nuclear wavefunction is expanded by time-dependent frozen Gaussian basis functions[26] while the adiabatic potential energy surfaces is modeled on-the-fly at the *ab initio* multi-reference electronic structure theory. In our simulation, state averaged complete active space self-consistent field (SA-CASSCF) theory[27] is applied to obtain the electronic wavefunction. State average three with two electrons in three orbitals active space were employed for SA-CASSCF method. The LANL2DZdp basis set and effective core potentials (ECP)[28] were used in all dynamics calculations. These parameters were chosen such that the low-lying excited states characters are in good agreement with higher level theory, for specific, equation-of-motion coupled cluster with single and double excitations (EOM-CCSD)[29] level of theory using the 6-31G** basis set. A further validation of these SA-CASSCF parameters will be discussed in the following text. All dynamical simulations were initiated on the S_1 state of the clusters with their initial positions and momenta randomly sampled from the ground state vibrational Wigner distribution. The ground state minimum geometry is optimized with density functional theory (DFT) with B3LYP functional[30-32] and 6-31G** basis set. The nuclear wavefunction is propagated using an adaptive multiple time step algorithm with classical and quantum degrees of freedom integrated by velocity Verlet and second-order Runge-Kutta shcemes with a 0.5 fs time

step.[25] The widths of the frozen Gaussian basis functions were chosen to be 25.0, 6.0, and 30.0 Bohr^2 for silicon, hydrogen, and oxygen atoms, respectively.

12.1.2 Quantum chemical study of silanol cluster

The accuracy of the SA-CASSCF PES that employed in the AIMS calculations were assessed by optimizing MECI and S_1 minimum (S_1 min) geometries at SA-CASSCF and multi-state complete active space second-order perturbation (MS-CASPT2) level of theory[33] and comparing the energy levels as well as geometrical similarities of these structures optimized with both theoretical levels. Additional single point calculations at multi-reference configuration interaction level of theory with single and double excitations and Davidson correction[34] (MRCI(Q)) were performed at Frank-Condon (FC), MECI and S_1 min geometries. Both MS-CASPT2 and MRCI(Q) employ the same active space and basis set as used in the SA-CASSCF with 6 orbitals besides the active orbitals were correlated. Further analysis of the PES was performed by applying single point calculations on *spawning points* (geometries on where an addition basis is spawned to another state in AISM simulations) at SA-CASSCF and MS-CASPT2 level of theories.

All dynamical calculations were performed with FMSMolPro package[25]. DFT, EOM-CCSD, SA-CASSCF, MS-CASPT2 and MRCI(Q) were performed in MolPro.[35-41] MECI and S_1 min geometries were optimized with MolPro in conjunction with CIOpt.[36,42,43]

12.1.3 Quantum chemical study of SiNC

We have identified a conical intersection in a 1.3-nm SiNC ($Si_{44}H_{45}OH$) following the procedure introduced in our prior work.[44] We first optimized the ground state structure

of the SiNC at the CAM-B3LYP/LANL2DZ level of theory.[28,45] Then, using this structure as a starting point, we optimized the conical intersection at the configuration interaction singles natural orbital complete active space configuration interaction level (CISNOr-CASCI), using the relaxed configuration interaction singles (CIS) density matrix.[46] CISNOr-CASCI is a recently developed alternative to SA-CASSCF with advantages for the treatment of nanoscale systems, i.e. size-intensivity of excitation energies and stability in the presence of orbital near-degeneracy. As above, an active space of two electrons in three orbitals (2/3) is chosen, with the lowest two CIS excited states included in the CISNOr density matrix used to determine the orbitals. Electronic structure calculations were performed using our recent GPU-accelerated implementation of CISNOr-CASCI in the TeraChem software package.[47-50] The optimization was performed in conjunction with the CIOpt software package using parallel numerical gradients.[42,51] Only a subset of atoms were optimized, with the remainder left frozen in their ground state minimum energy positions; specifically a single silaadamantane group including the silanol silicon atom (ten silicon atoms total) was optimized, with all associated hydrogen and oxygen atoms. For comparison, the same protocol was applied to reoptimize the MECI in the $Si_7H_{11}OH$ cluster model described in the previous subsections. For this smaller cluster model all atoms were fully optimized. Absolute energies and geometries for the calculations described in this section are reported in APPENDIX D (see Tables 12-S5 and 12-S6).

12.2 Results and Discussion

12.2.1 Dynamical simulation of the silicon cluster with hydroxide defect

At ground state minimum, the method employed in dynamical simulation, SA3-CASSCF(2/3) level of theory with LANL2DZdp basis set predicts 6.015 eV vertical excitation energy (as compared to EOM-CCSD/6-31G** level of theory: 5.456 eV). The discrepancy between SA-CASSCF and EOM-CCSD is under the error of SA-CASSCF method due to the latter lacks of dynamical correlations. The accuracy of the SA-CASSCF PES will be further validated in the following sections. Dynamical simulations show the S_1-S_0 energy gap drops below 1 eV in 100 fs as shown in Figure 12-1. Individual trajectories hitting zero energy gap suggests the conical intersection may play a role in the non-radiative recombination process. Figure 12-2 shows the ground and excited state population as a function of time. A fast S_1 population decay happens around 100 fs, which is corresponding to the time scale in Figure 12-1 when a swarm of trajectories are hitting small S_1-S_0 energy gaps. The S_1 state population decays 88.6% to ground state in 250 fs. This corresponds to a time scale that is much shorter than the radiative lifetime of the PL of silicon nanoparticles. However, it has to be point out here that this life time corresponds to the non-radiative decay time of the defects only. In real SiNC, the process of excitation energy localization from quantum confined states to defect localized state may be slow.

The population transfer is associated with Si-Si bond breaking. Figure 12-3 shows the averaged Si-Si bond distances increases from 2.37 Å to 4.35 Å with the same time scale.

The red arrow in the insets show the corresponding Si-Si bond that is accounted in Figure 12-3. The rapid bond distance increasing process happens around 100 fs, this corresponds to the same time scale of excited state population decay. The dependence of population decay and Si-Si bond breaking will be investigated further by optimizing the MECI geometry and will be discussed in the next section.

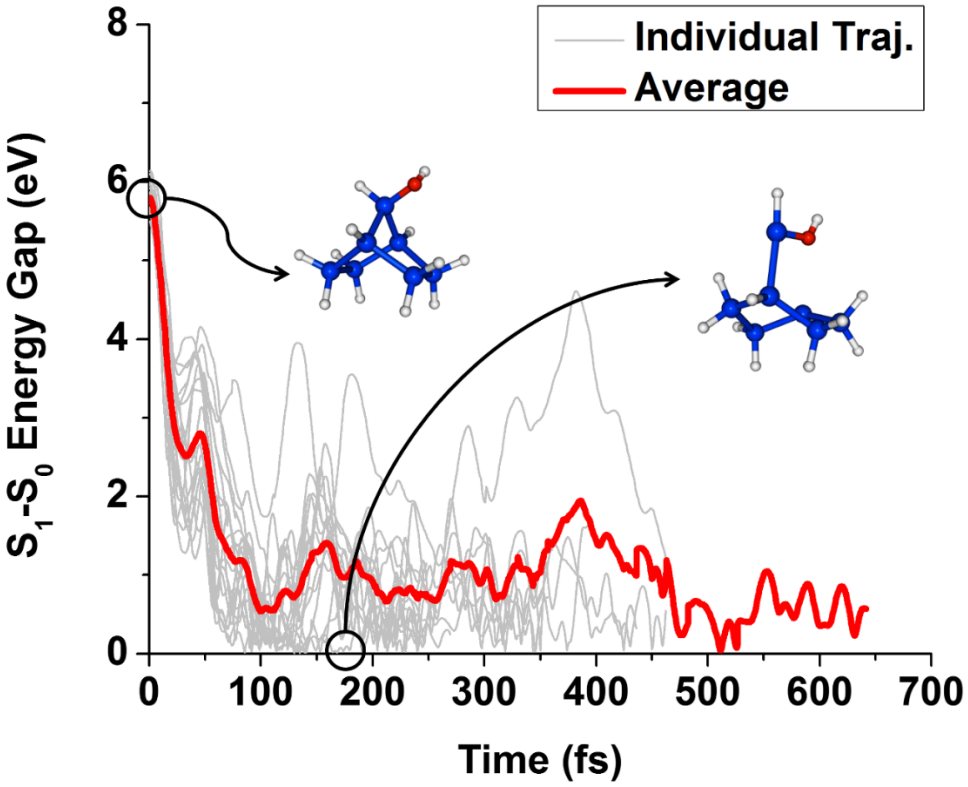


Figure 12-1. Average S_1-S_0 energy gap of silicon cluster as a function of time (bold red) with the energy gaps of 20 initial nuclear basis functions (gray). Snapshots from representative trajectories are shown in the inset, with silicon, hydrogen and oxygen atoms colored blue, white and red, respectively.

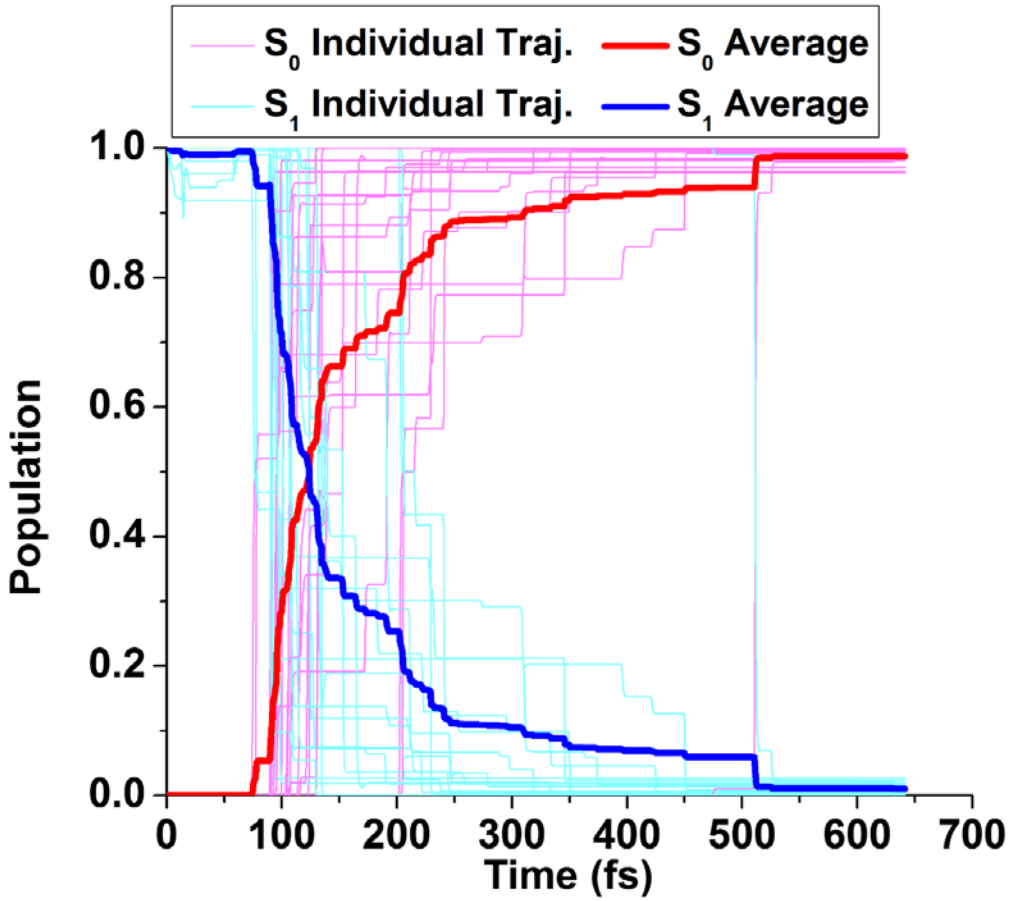


Figure 12-2. Population of S_1 (bold blue) and S_0 (bold red) as a function of time after excitation averaged over all 20 initial nuclear basis functions. Thinner blue and pink lines reflect the population corresponding to the 20 individual basis functions before averaging.

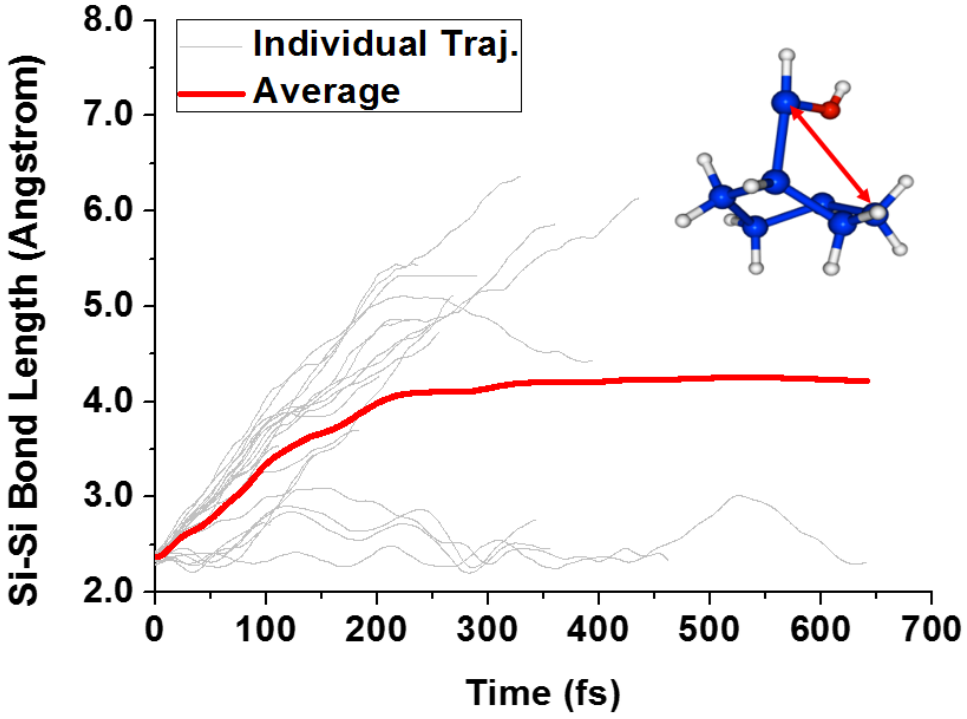


Figure 12-3. Average Si-Si bond distance as a function of time(bold red) with the Si-Si bon distance of 20 initial nuclear basis sets(grey), the inset shows the atoms associated with the atom distance concerned in graph, with red, blue and grey represents oxygen, silicon and hydrogen atoms respectively.

Coherence wagging oscillations of the OH bond is observed during the simulation as shown in Figure 12-4. The H(2)OSiH(1) dihedral angle is defined as the dihedral angle between H(2)OSi plane and OSi(H1) plane. H(2)OSiH(1) dihedral angle varies from -48.8 to 47.5 degrees from individual nuclear trajectories. Despite the averaged wagging coherence is smoothed out at long time scale, individual trajectories show strong

oscillations during the whole time scale of the simulation. Although the S_1-S_0 energy gap has weak dependence on the wagging oscillation, this long lived coherence oscillation may be treated as a signature of hydroxide group defects in the time-dependent spectroscopy. This motion corresponds to 284.8 cm^{-1} frequency calculated at DFT with B3LYP/6-31G** level of theory.

A histogram of S_1-S_0 energy gap integrated over the excited state population is shown in Figure 12-5. A broad peak located around 0.5 eV is observed. The discrepancy between the energy of the peak in histogram and experimental observed 1.5 to 2.1 eV PL peak can be explained by the fact that the experimental observed S-band PL energy corresponds to the quantum confined states. And hence the histogram does not correspond to the PL peak observed in the experiment, as the histogram is calculated by integration over the energy gaps between defect localized state and ground state.

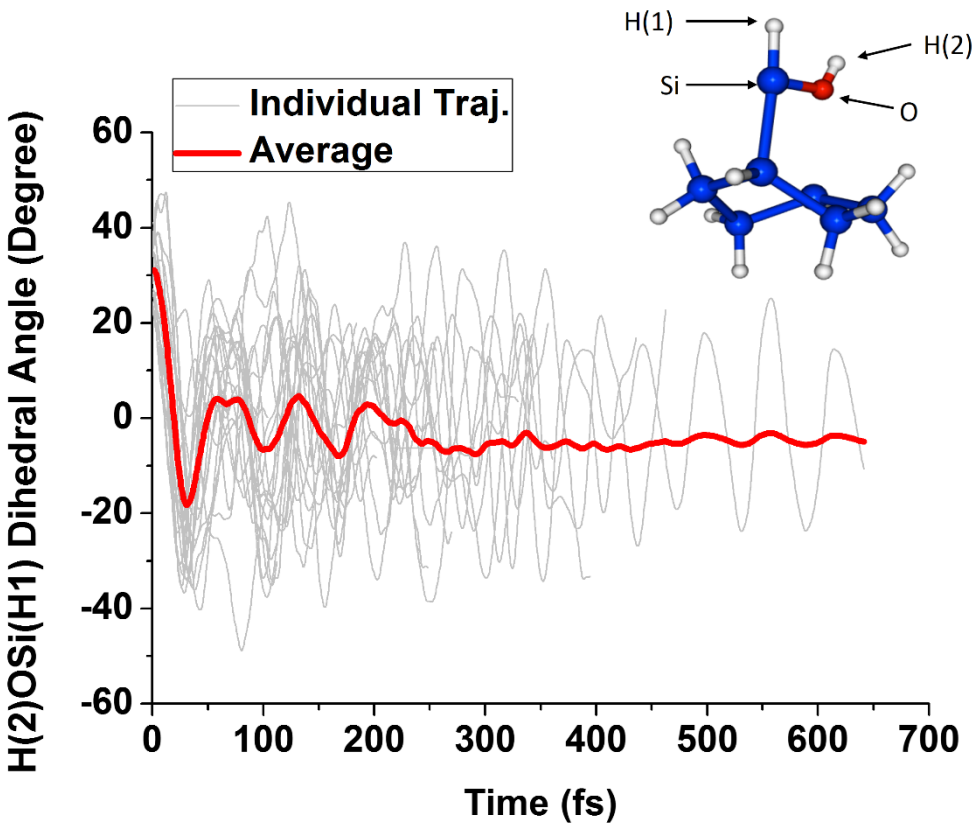


Figure 12-4. Average H(2)OSiH(1) dihedral angle as a function of time(bold red) with the H(2)OSiH(1) dihedral angle of 20 initial nuclear basis sets(grey), the inset shows the atoms associated with the dihedral angle concerned in graph, with red, blue and grey represents oxygen, silicon and hydrogen atoms respectively.

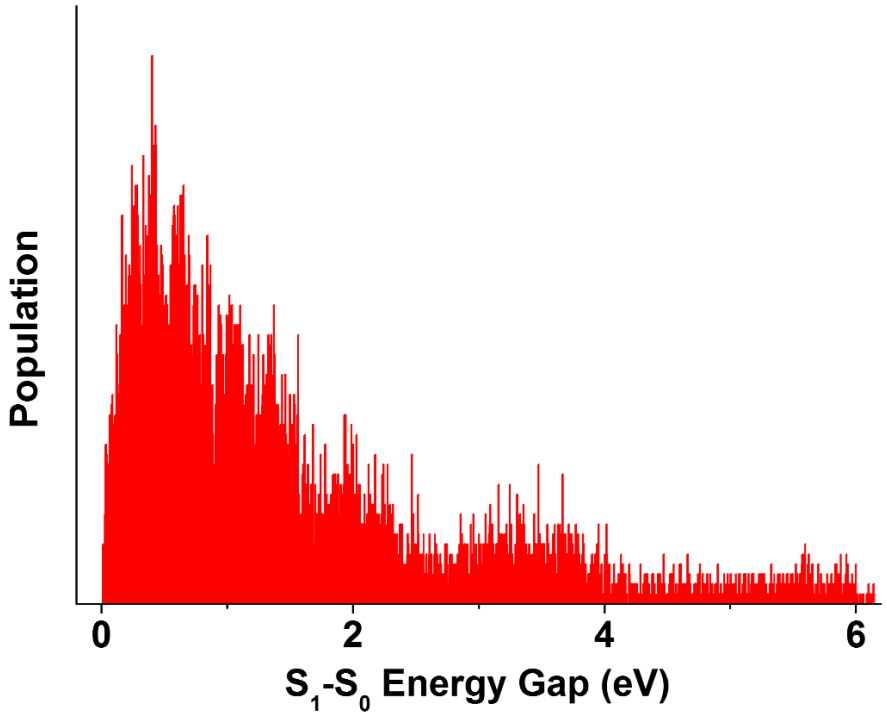


Figure 12-5. Histogram of the S_1-S_0 energy gap of the excited population collected over all 20 initial nuclear basis functions integrated over the duration of the simulations.

12.2.2 PES analysis of silanol cluster

The fact that the time scale of Si-Si bond breaking is the same as population decay and single trajectories' S_1-S_0 energy gap hitting zero suggests the conical intersection is playing a role for the non-radiative decay process and the MECI geometry is reached by breaking a Si-Si bond. To further investigate this possibility, we applied SA-CASSCF and MS-CASPT2 theoretical levels to explore the PESs. The results are reported in Table 12-1, Figure 12-6 and Figure 12-7. The MECI geometries optimized at both theoretical levels

have one broken Si-Si bond as shown in the insets of Figure 12-6 and 12-7. Geometries that are similar to the optimized MECI were observed in the dynamical simulations as well. Single point calculations at more trustworthy MRCISD(Q) level of theory are performed on the optimized MECI geometries for comparison. A stable minimal energy geometry of the S_1 state is found with similar structure and energy of MECI. Both MECI and S_1 min are energetically accessible by vertical excitation of the cluster at Frank-Condon (FC) geometry.

The SA-CASSCF method predicts 4.36 eV above the ground state minimum at MECI geometry as compared to the ground state of the FC geometry. As compared with 4.63 and 4.63 eV excitation energies at SA-CASSCF optimized MECI geometry computed at higher level theories with dynamical correlations MS-CASPT2 and MRCI(Q) methods. The 0.27 eV discrepancy between SA-CASSCF and MS-CASPT2 or MRCI(Q) is very common. MS-CASPT2 and MRCI(Q) also show 0.19 and 0.10 eV S_1-S_0 energy gap at SA-CASSCF optimized MECI geometry. These small gaps suggest the SA-CASSCF PES employed in dynamical simulations is qualitatively correct. At SA-CASSCF optimized S_1 min geometry, SA-CASSCF, MS-CASPT2 and MRCI(Q) predicts the 4.34, 4.49 and 4.42 eV excitation energies, respectively. This corresponds to 0.02, 0.14 and 0.21 eV barriers between MECI and S_1 min geometries computed at SA-CASSCF, MS-CASPT2 and MRCI(Q) level of theories. The small MECI and S_1 min barrier indicates the MECI is nearly a peaked MECI, which is further demonstrated by the fast population decay as shown in Figure 12-2. SA-CASSCF method predicts 0.31 eV S_1-S_0 energy gap at SA-CASSCF optimized S_1 min

geometry, this is closed to the peak energy of the histogram shown in Figure 12-5.

Table 12-1. The S_0 and S_1 energies (in eV) at three important points on the PES of the silicon hydroxide cluster calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory. S_1-S_0 MECI geometries are optimized at both the CASSCF and MS-CASPT2 levels of theory.

Geometry	Optimization Method	CASSCF S_0 Energy	CASSCF S_1 Energy	MS-CASPT2 S_0 Energy	MS-CASPT2 S_1 Energy	MRCI(Q) S_0 Energy	MRCI(Q) S_1 Energy
FC geometry	KS	0	6.02	0	5.91	0	5.96
S_0-S_1 MECI	CASSCF	4.36	4.36	4.44	4.63	4.53	4.63
	MS-CASPT2	4.36	4.52	4.48	4.50	4.34	4.42
S_1 min	CASSCF	4.03	4.34	4.40	4.49	4.27	4.42
	MS-CASPT2	4.28	4.51	4.40	4.49	4.27	4.27

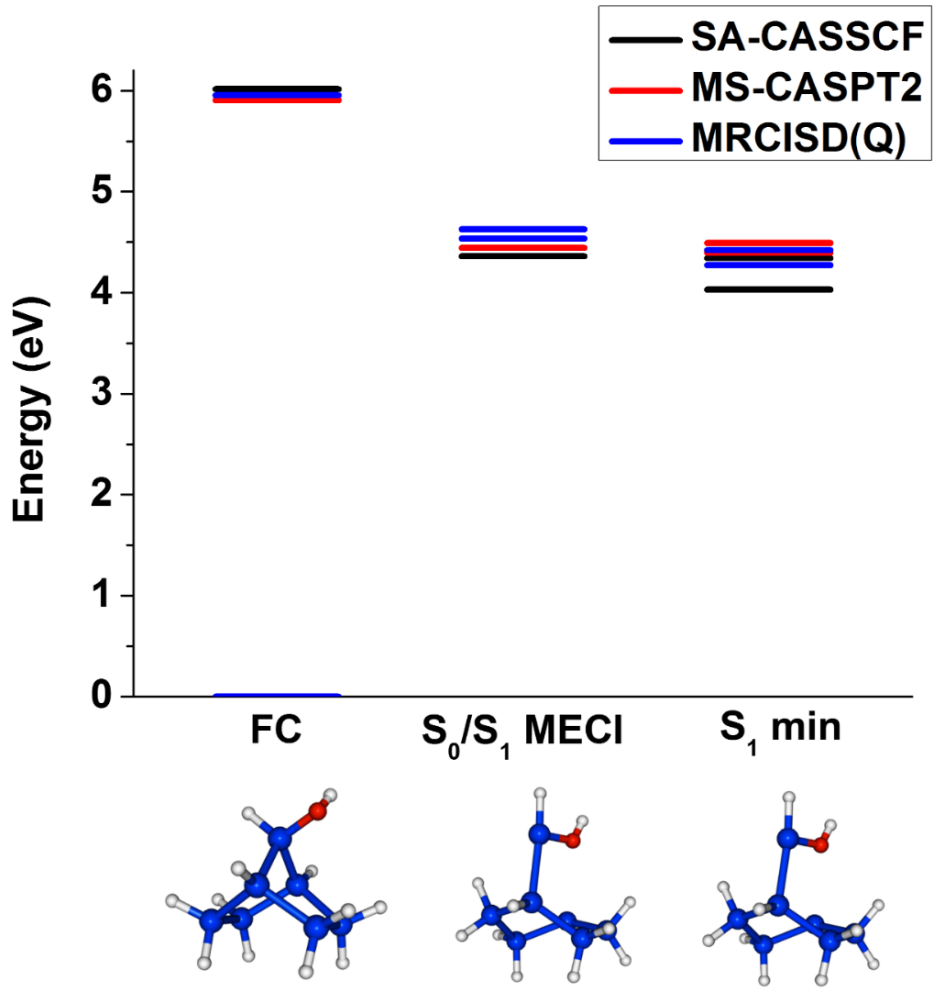


Figure 12-6. The S_0 and S_1 energies of three important points on the PES of the silicon hydroxide cluster: the Franck-Condon (FC) point, the S_0/S_1 MECI, and the S_1 min. Black, red and blue lines correspond to the energies of the SA-CASSCF optimized geometries computed at the SA-CASSCF, MS-CASPT2 and MRCISD(Q) levels of theory, respectively. Energies are shifted so that the S_0 energy of the FC point is zero for all levels of theory. Only one black line is visible for the S_0/S_1 MECI because S_0 and S_1 are degenerate. Optimized geometries are shown as insets.

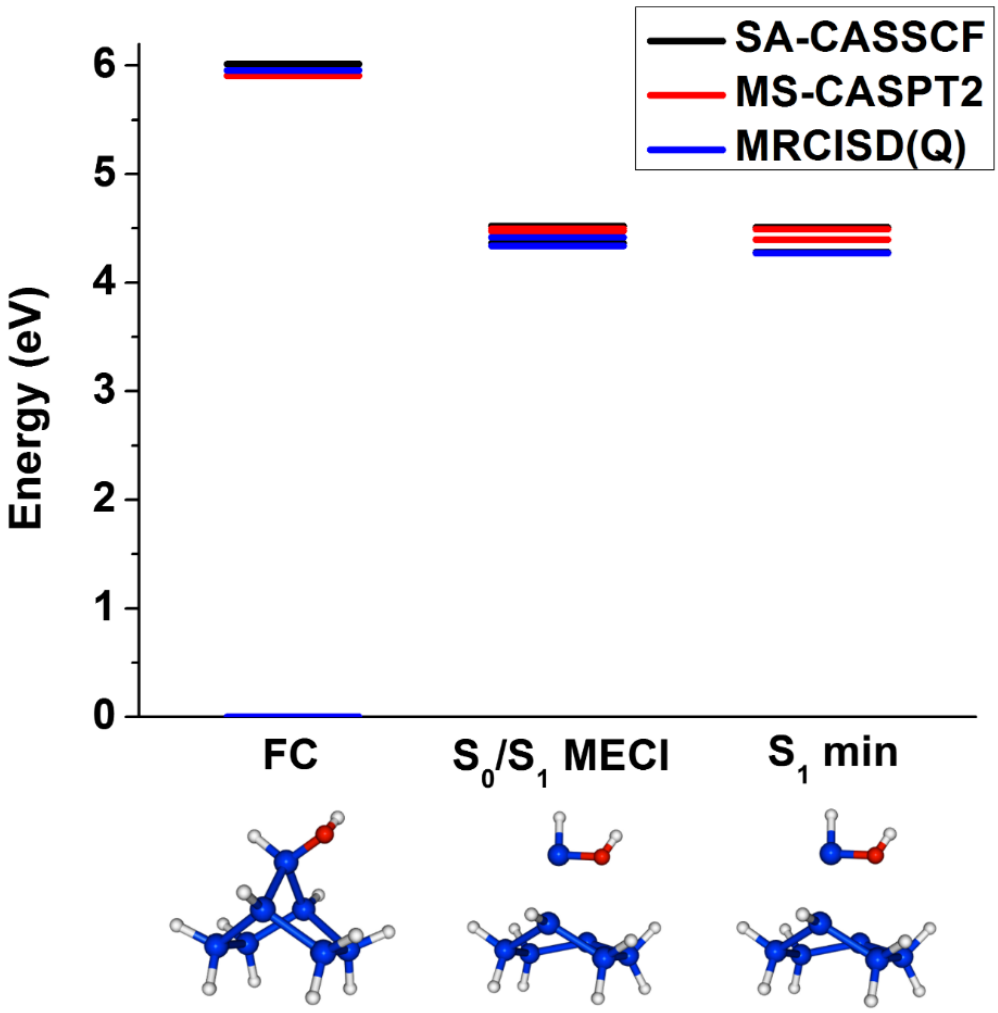


Figure 12-7. The S_0 and S_1 energies of three important points on the PES of the silicon hydroxide cluster: the Franck-Condon (FC) point, the S_0/S_1 MECI, and the S_1 min. Black, red and blue lines correspond to the energies of the MS-CASPT2 optimized geometries computed at the SA-CASSCF, MS-CASPT2 and MRCISD(Q) levels of theory, respectively. Energies are shifted so that the S_0 energy of the FC point is zero for all levels of theory. Only one black line is visible for the S_0/S_1 MECI because S_0 and S_1 are degenerate. Optimized geometries are shown as insets.

To further access the accuracy of SA-CASSCF PES, we re-optimized the MECI and S_1 min geometries at MS-CASPT2 level of theory. The MS-CASPT2 optimized MECI structure is shown as inset in Figure 12-7. MS-CASPT2 predicts the S_1 state as 4.50 eV at MS-CASPT2 optimized MECI geometry. This is closed to 4.36 eV excitation energy calculated with SA-CASSCF method at SA-CASSCF optimized MECI geometry. This small energy level difference suggests our SA3-CASSCF(2/3) PES employed in our dynamical simulations is in qualitative good agreement with higher level of theory. The importance of dynamical correlations is further evaluated by understanding the correlations between SA-CASSCF S_1 and MS-CASPT2 S_1 energies as well as SA-CASSCF S_1-S_0 energy gaps and MS-CASPT2 S_1-S_0 energy gaps at *spawning points* (defined in method section) as shown in Figure 12-8. The roughly linear relationship between SA-CASSCF S_1 and MS-CASPT2 S_1 energies indicates that the SA-CASSCF PES is trustworthy. However, it is shown from Figure 12-8(b) that MS-CASPT2 predicts systematically larger gap at spawning points compared to SA-CASSCF level of theory. The discrepancy suggests that MS-CASPT2 may predict a longer excited state life time.

It is noteworthy that the consistent MECI energies around 4.5 eV predicted at various theories is far beyond the S-band PL energy around 1.5 to 2.1 eV. As has been hypothesized by us in previous chapters, only quantum confined delocalized states with energies higher than MECI energies will be quenched through the localized nuclear distortion to reach the MECI geometry. This indicates the MECI introduced by hydroxide defect at the SiNC surface is less likely to play a significant role in the visible range

photochemistry. The conclusion is also consistent with the experimental observation that by passivating the dangling bonded SiNC surface with hydroxide groups, the PL intensity increases.

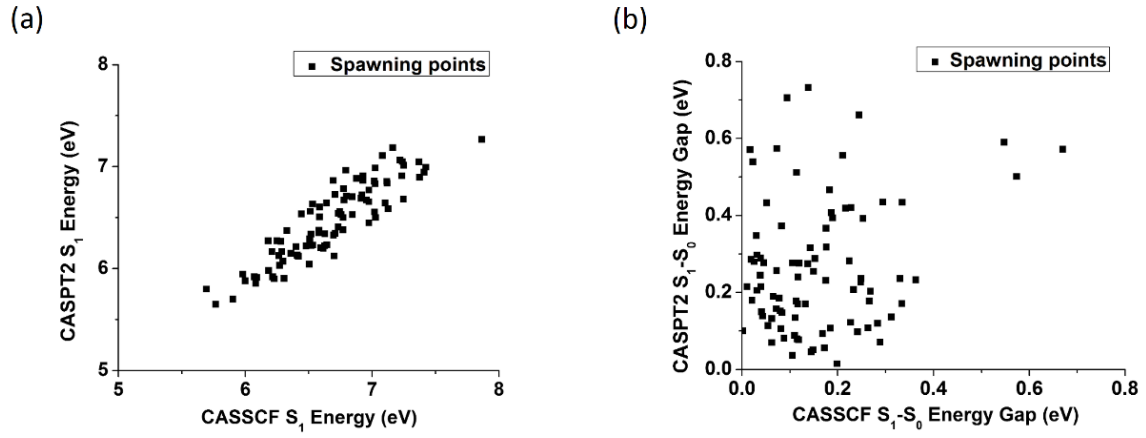


Figure 12-8. (a) The S_1-S_0 energy gap correlation between SA-CASSCF and MS-CASPT2 theoretical levels with 7 orbitals correlated in the MS-CASPT2 calculations. (b) The adiabatic S_1 excitation energy correlation between SA-CASSCF and MS-CASPT2 theoretical levels.

Partially occupied natural orbitals computed on SA-CASSCF level of theory at FC and MECI geometries are shown in Figure 12-9. The natural orbitals are determined by diagonalizing S_1 state density matrix at FC geometry and averaged S_1-S_0 density matrix at MECI geometry. As seen from Figure 12-9, at FC geometry, the transition can be most effectively described as σ Si-Si bond to an anti-bonding character orbital formed by O p orbital and Si p orbitals. At MECI geometry, the excitation character is described by a

non-bonding p orbital localized on one Si atom involved in bond breaking during relaxation to a mixing of Si p orbital and O p orbital with anti-boding character. It is noteworthy here that the excited state character can't be described by one particle picture despite the fact that natural orbital is the most effective way in representing excitation character. This indicates the importance of multi-reference electronic structure theory in studying the photochemistry.

12.2.3 Conical intersection in SiNC

As described above, we identified the constrained S_0/S_1 MECI of a 1.3-nm SiNC at the CISNOr-CASCI/LANL2DZ level of theory. The resulting structure and orbitals are presented in Figure 12-10. This MECI was found to be 4.88 eV above the ground state minimum energy. This is 0.52 eV above the energy of the CASSCF-optimized intersection in the cluster model described in section 12.2.2. Thus, the MECI remains inaccessible at visible energies in the larger SiNC system. The distance between the silanol silicon atom and the remainder of the cluster is also somewhat longer than predicted above; the distances between the silanol silicon atom and the two adjacent atoms are 4.86 Å and 3.65 Å, respectively, compared to 5.13 Å and 2.60 Å in the CASSCF-optimized MECI of the cluster model.

In order to determine whether these differences are the result of increasing the size of the system to the nanoscale or artifacts of the electronic structure method and basis set applied to the larger SiNC, we have identified the MECI of the silanol cluster model at the CISNOr-CASCI/LANL2DZ level. This MECI is only 4.29 eV above the ground

state minimum energy and the Si-Si distances at the silanol silicon atom are 4.13 Å and 2.91 Å. Both the energies and geometry are more similar to those of the CASSCF-optimized MECI of the cluster model described in section 12.2.2 than they are to the CISNOr-CASCI-optimized MECI of the SiNC. This suggests that our cluster model slightly underestimates the MECI energy compared to that of a true nanoscale system, and that SA-CASSCF and CISNOr-CASCI are in good agreement with one another.

The CISNOr highest occupied and lowest unoccupied molecular orbitals (HOMOs and LUMOs) of both the SiNC and the smaller cluster model are presented in Figure 12-10. These orbitals were determined by diagonalization of the state-averaged relaxed CIS density matrix.[46] Note that the chemical character of these orbitals is identical in the larger SiNC and smaller cluster model. The quantitative difference in energy and structure are thus not the result of qualitative differences in the electronic structure or delocalization of the electron or hole. The cause of this quantitative difference in the energy and structure is not immediately apparent, but it may be the result of strain in the silanol cluster or dielectric stabilization of the ground electronic state in the larger SiNC system.

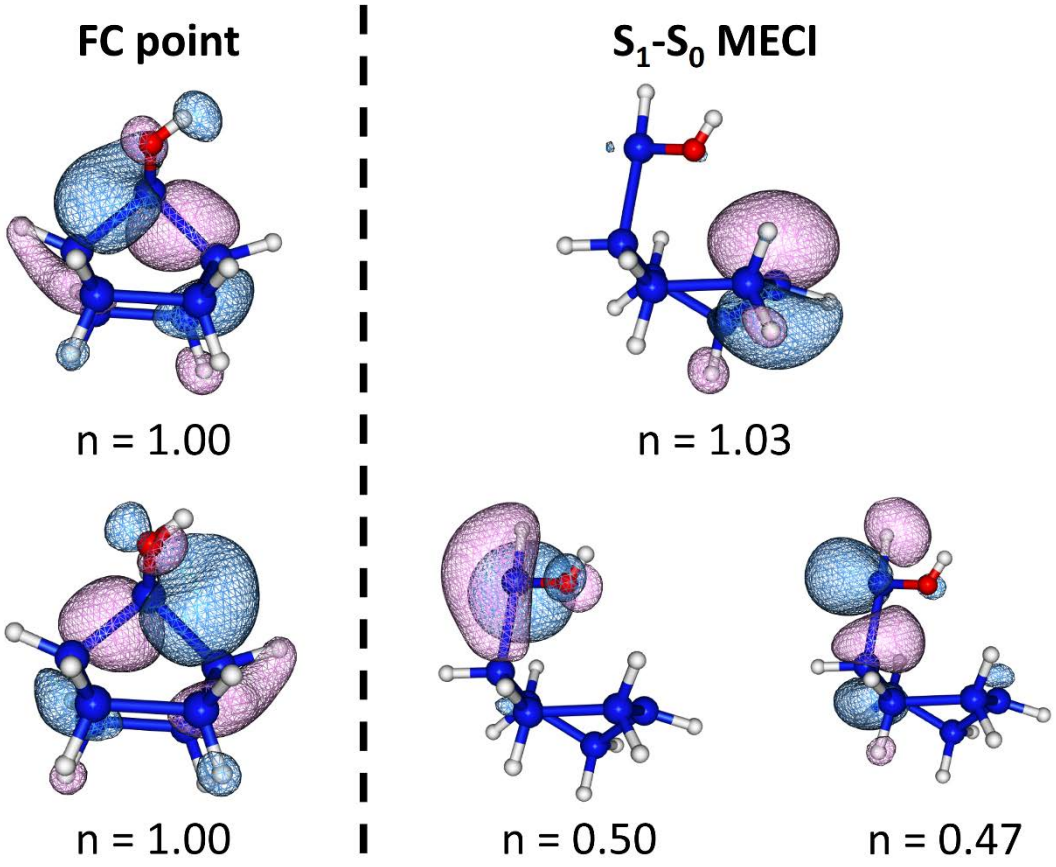


Figure 12-9. Partially occupied SA-CASSCF natural orbitals of silicon hydroxide cluster at FC (left) and MECI (right) geometries are presented with occupation numbers. The FC natural orbitals are computed by diagonalization of the S_1 density matrix, while S_1–S_0 MECI natural orbitals are computed from the state-averaged S_1–S_0 density matrix.

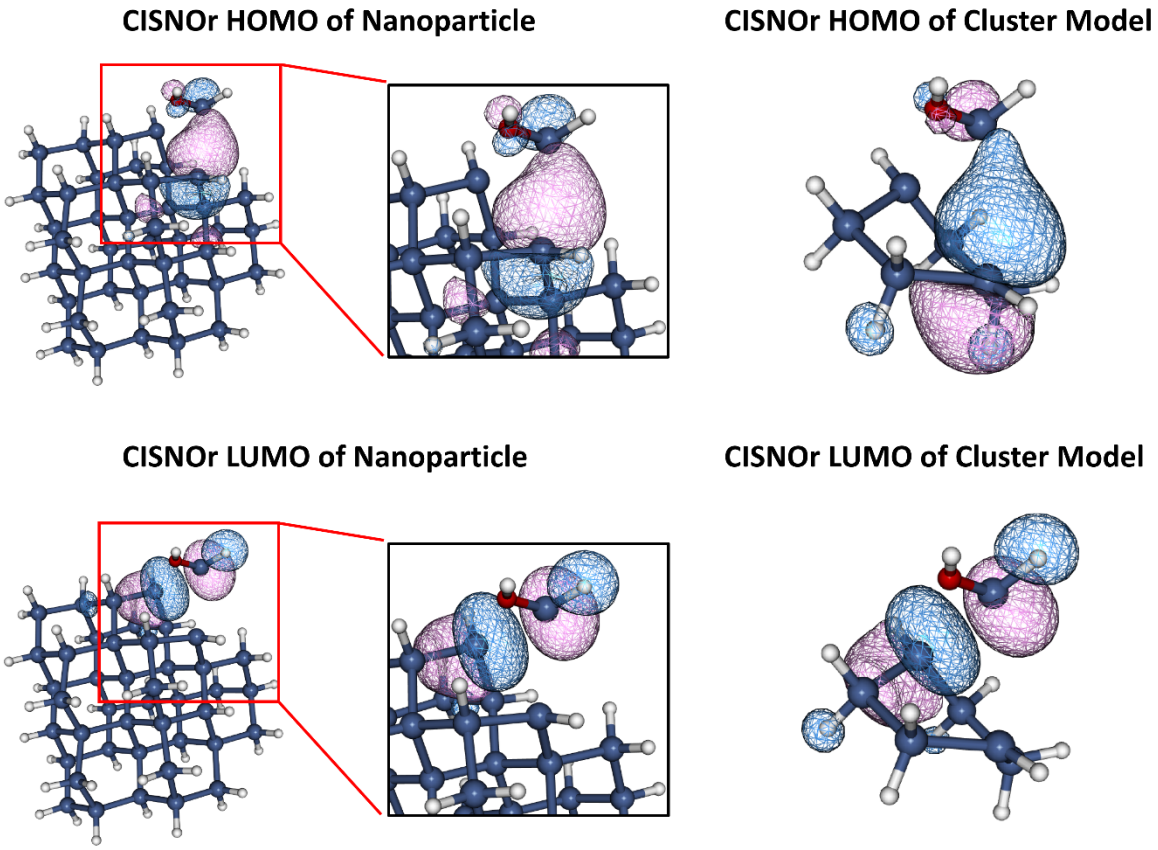


Figure 12-10. The S_0/S_1 MECI structure of the SiNC (left) and cluster model (right) as optimized at the CISNOr-CASCI/LANL2DZ level. The HOMOs (top) and LUMOs (bottom) shown correspond to the CISNOr orbitals determined by diagonalization of the state-averaged relaxed CIS density matrix.[46]

12.3 Conclusions

In this work, we have applied ab initio multiple spawning method to simulate the excited state dynamics of silicon cluster with hydroxide group as defect. Our calculations indicate the MECI geometry is introduced by breaking a Si-Si bond, in which one of the Si

atom is bonded to the hydroxide group at the ground state minimum. The various levels of theory including SA-CASSCF, MS-CASPT2 and MRCI(Q) predict the MECI has adiabatic excitation energy about 4.5 eV. This high excitation energy is in ultraviolet range, hence not relevant with the visible range photochemistry, and we conclude that MECI introduced by hydroxide group does not play a significant role of the S-band PL energy which is located around 1.5 to 2.1 eV. Our prediction also explains that experimental observation that by passivating dangling bond defect SiNC surface with hydroxide group, PL intensity increases as pass ways for non-radiative decay in visible light range is blocked by rising up the MECI energy level. Although the dangling bond defect will not be discussed in this thesis, we have already found the MECI introduced by dangling bond defect is in the visible range and hence plays a significant role in PL.

APPENDIX

APPENDIX

Table 12-S1. The absolute energies of the various optimized geometries of the cluster model at the CASSCF level of theory (with active space and basis sets as described in the text).

CASSCF	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point (B3LYP/6-31G**)	-108.3638494	-108.1427958
CASPT2-optimized S_0/S_1 MECI	-108.2035184	-108.1976480
CASPT2-optimized S_1 min	-108.2065315	-108.1982210
CASSCF-optimized S_0/S_1 MECI	-108.2036976	-108.2035120
CASSCF-optimized S_1 min	-108.2156875	-108.2042386

Table 12-S2. The absolute energies of the various optimized geometries of the cluster model at the MS-CASPT2 level of theory (with active space and basis sets as described in the text).

MS-CASPT2	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point (B3LYP/6-31G**)	-108.5204777	-108.3034123
CASPT2-optimized S_0/S_1 MECI	-108.3571558	-108.3504166
CASPT2-optimized S_1 min	-108.3611753	-108.3568622
CASSCF-optimized S_0/S_1 MECI	-108.3559980	-108.3552157
CASSCF-optimized S_1 min	-108.3588514	-108.3553773

Table 12-S3. The absolute energies of the various optimized geometries of the cluster model at the MRCISD(Q) level of theory (with active space and basis sets as described in the text).

MRCI(Q)	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point (B3LYP/6-31G**)	-108.527430	-108.3065713
CASPT2-optimized S_0/S_1 MECI	-108.3638933	-108.3599961
CASPT2-optimized S_1 min	-108.3703899	-108.3633536
CASSCF-optimized S_0/S_1 MECI	-108.3652961	-108.3623955
CASSCF-optimized S_1 min	-108.3678670	-108.3626624

Table 12-S4. The absolute energies of the Franck-Condon point of the cluster model at the EOM-CCSD level of theory.

EOM-CCSD	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point (B3LYP/6-31G**)	-2105.25375013	-2105.05325249

Table 12-S5. The absolute energies of the various optimized geometries of the cluster model at the CISNOr-CASCI level of theory (with active space and basis sets as described in the text).

CISNOr-CASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point (CAMB3LYP/LANL2DZ)	-108.084696998	-107.834012122
CISNOr-CASCI optimized S_0/S_1 MECI	-107.927166854	-107.926987807

Table 12-S6. The absolute energies of the various optimized geometries of the nanoparticle at the CISNOr-CASCI level of theory (with active space and basis sets as described in the text).

CISNOr-CASCI	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point (CAMB3LYP/LANL2DZ)	-267.00997175	-266.76034526
CISNOr-CASCI optimized S_0/S_1 MECI	-266.83088148	-266.83076768

Optimized Geometries (In Angstrom)

FC cluster geometry (B3LYP/6-31G)**

O	20.91065863	16.26933114	21.13181009
Si	19.75381462	17.30211718	20.48465819
Si	17.75361503	16.25419908	19.76033972
Si	16.36135706	18.12714013	19.36779609
Si	17.04136532	19.68001481	21.03940175
Si	18.74334840	18.53618260	22.22250420
Si	17.69492285	16.80687382	23.45381924
Si	17.02817758	15.25255403	21.77969483
H	14.91531381	17.78565858	19.48548554
H	16.58759158	18.69429453	18.00736878
H	17.56849290	20.92262782	20.40533718
H	15.90640264	20.06578948	21.92515868
H	19.62950900	19.42888576	23.02222262
H	16.52163768	17.30583330	24.22703294
H	18.64310589	16.17542117	24.41382980
H	17.69754145	13.93717660	21.99068566
H	15.55556288	15.02165049	21.80194826
H	17.84625378	15.30923604	18.60850443
H	21.30770017	15.63359810	20.52498370
H	20.34529572	18.16329535	19.41305429

SA-CASSCF-optimized cluster MECI

O	20.68031360	16.52786762	20.93798147
Si	21.07258835	18.05635784	21.36628924
Si	16.72978223	15.95897612	19.62948929
Si	15.92570627	18.14423226	19.65610121
Si	17.53359124	19.53855153	20.63096637
Si	18.70465080	18.46466455	22.35751026
Si	17.55865483	16.65729214	23.30732444
Si	17.13633462	14.99350386	21.71370713
H	14.66871256	18.23404505	20.43549597
H	15.63761873	18.58107586	18.27251658
H	18.51708739	19.90691374	19.57956890
H	16.90477242	20.79557001	21.09375700
H	18.99257120	19.47018262	23.40856989
H	16.26207865	17.06915635	23.89808477
H	18.38193455	16.09399852	24.40218834
H	18.28676407	14.07146517	21.60729470
H	15.96125075	14.19684815	22.13727687
H	15.98669116	15.07259708	18.70497651
H	21.34851200	16.03406942	20.47911952
H	22.50310675	17.96207292	20.76473648

SA-CASSCF-optimized cluster S_1 min

O	20.61362570	16.50662082	20.92769951
Si	21.02115163	18.01715696	21.41655577
Si	16.74889541	15.95689545	19.61455191
Si	15.93369416	18.13753454	19.64104552
Si	17.52266344	19.54949966	20.62053979
Si	18.71849455	18.48990113	22.34275427
Si	17.59980430	16.66408859	23.29378835
Si	17.16126747	15.00154789	21.70245750
H	14.67360948	18.21878111	20.41581886
H	15.64564216	18.57589267	18.25838880
H	18.49466396	19.95299422	19.57380324
H	16.86958106	20.78363957	21.10949939
H	18.99307093	19.49924934	23.39225740
H	16.31192412	17.07656741	23.90208108
H	18.44197946	16.10366336	24.37560693

H	18.30312217	14.06894549	21.59840289
H	15.98176825	14.21735093	22.13732071
H	15.99786275	15.06517881	18.70163453
H	21.28948428	16.01244494	20.48112091
H	22.47041688	17.93148795	20.87762722

MS-CASPT2-optimized cluster MECI

O	20.89390708	16.39968666	21.47770538
Si	20.71478639	17.76370194	20.58777974
Si	16.79107156	16.00131158	19.72011265
Si	16.23606455	18.24049994	19.32075698
Si	17.13006748	19.67581098	20.96253800
Si	18.65739042	18.48834169	22.27687090
Si	17.71772633	16.68873181	23.43445128
Si	17.07688392	15.06878578	21.84797768
H	14.77258356	18.39121749	19.23381870
H	16.79867784	18.58395652	18.00402917
H	17.82152560	20.78380366	20.28462546
H	16.02740751	20.22861405	21.76503296
H	19.36250820	19.40893568	23.19479281
H	16.54402385	17.07283490	24.24335364
H	18.72813294	16.11392218	24.33638937
H	18.12562644	14.04649713	21.72993613
H	15.82720823	14.39415643	22.24689714
H	16.05132013	15.09320293	18.81921635
H	21.61453655	15.83047750	21.24056104
H	21.90127352	17.55495189	19.65610891

MS-CASPT2-optimized cluster S_1 min

O	20.89377784	16.39046392	21.47814762
Si	20.71261865	17.76166138	20.59918184
Si	16.79605596	16.00270690	19.71931077
Si	16.23428021	18.23695170	19.32050851
Si	17.12824987	19.67658088	20.96237907
Si	18.66683847	18.49252467	22.26446542
Si	17.71823133	16.68663062	23.43068289
Si	17.07522089	15.06717800	21.84741183

H	14.77026731	18.39178808	19.23581580
H	16.79866157	18.58554362	18.00531918
H	17.82099128	20.78339551	20.28571209
H	16.02637827	20.22781057	21.76732097
H	19.36520461	19.41145298	23.19074181
H	16.54368706	17.07424628	24.23916000
H	18.72652213	16.11302897	24.33604223
H	18.12464073	14.04536301	21.73221806
H	15.82525775	14.39566077	22.25100840
H	16.05388176	15.09370326	18.81936373
H	21.61636848	15.82584306	21.23980109
H	21.89558794	17.56690656	19.65836290

FC cluster geometry (CAMB3LYP/LANL2DZ)

O	20.79615487	16.17151497	21.17586698
Si	19.73565351	17.30134285	20.46187389
Si	17.73432953	16.28974592	19.73141242
Si	16.34516244	18.15335492	19.38898454
Si	17.04129561	19.68066808	21.05578902
Si	18.75271787	18.53907677	22.19345026
Si	17.71325093	16.81573272	23.40695460
Si	17.05162524	15.27646428	21.74039159
H	14.91249597	17.80117198	19.54630903
H	16.54114449	18.72953033	18.03485312
H	17.53942496	20.93469357	20.43732189
H	15.92080000	20.01964208	21.96726341
H	19.65653207	19.41158659	22.97879413
H	16.52968282	17.33322175	24.13891994
H	18.63889550	16.19172104	24.38175497
H	17.72220748	13.96704917	21.93002067
H	15.58336780	15.06009044	21.76616246
H	17.82183221	15.36417796	18.57525669
H	21.36147773	15.51810656	20.74109583
H	20.41361598	18.11298764	19.41316014

CISNOr-CASCI-optimized cluster MECI

Si	20.73120905	17.76328613	20.56392506
Si	17.07557917	16.07445186	19.65887393
Si	16.21711201	18.23585589	19.36636797
Si	17.04228269	19.68733411	21.03414783
Si	18.58709464	18.55894723	22.36997676
Si	17.70143462	16.65745062	23.42962653
Si	17.07327281	15.08732231	21.78967825
O	21.05489344	16.49863070	21.62748478
H	14.74282557	18.17423403	19.36719898
H	16.64635801	18.69821810	18.03021378
H	17.75176270	20.79271824	20.36192119
H	15.93357325	20.26493432	21.81520771
H	19.18631176	19.49933900	23.33643598
H	16.51245703	16.93919245	24.27806493
H	18.70210556	16.00338791	24.30022132
H	17.99939780	13.94696555	21.77364925
H	15.73431587	14.51712074	22.04677699
H	16.83594270	15.13827328	18.53762564
H	21.66788419	15.77915447	21.59029829
H	21.84530772	17.45948954	19.53425284

FC nanoparticle geometry (CAMB3LYP/LANL2DZ)

O	2.537264995	-1.618487322	-5.528617149
Si	5.675741709	-5.005193789	0.006861701
Si	3.650706049	-4.997914840	1.212297340
Si	4.245025375	-4.964871226	3.487039099
Si	5.342677576	-6.953227990	4.063860779
Si	7.376466153	-6.964792666	2.897443388
Si	6.977211710	-6.879346207	0.583493350
Si	9.006144517	-6.822809849	-0.597450128
Si	8.590209478	-6.781589694	-2.901657363
Si	7.390235248	-8.699225481	-3.507314213
Si	5.336869645	-8.779397189	-2.365681625
Si	4.220140178	-10.7911120	-2.833527410
Si	2.116433262	-10.75874773	-1.802474471
Si	2.494252159	-10.78076881	0.517045446
Si	3.687596280	-8.845644338	1.146512512

Si	2.410790807	-6.927078738	0.619918131
Si	2.007228815	-6.910220495	-1.719652254
Si	0.724232544	-4.992031091	-2.236853922
Si	0.302399720	-4.953665084	-4.559240856
Si	2.335368416	-4.854415336	-5.733240116
Si	3.587624356	-6.780905715	-5.232426051
Si	2.332541173	-8.605267867	-6.001980263
Si	0.395292349	-8.769725446	-4.691533145
Si	-0.907481676	-6.874315936	-5.177853419
Si	-2.954542378	-6.989998052	-4.051962535
Si	-2.502264125	-7.057303845	-1.757524384
Si	-1.334474180	-5.127925304	-1.091838468
Si	-1.094494292	-5.273293167	1.236013750
Si	0.340881144	-7.055732657	1.766715040
Si	0.602853382	-7.221652535	4.098546543
Si	2.037088643	-9.000237172	4.624732213
Si	4.079802888	-8.843404845	3.471597038
Si	5.263571525	-4.936462357	-2.313319314
Si	3.978699032	-3.030486974	-2.837809644
Si	3.536819436	-2.948205893	-5.139858700
Si	5.727250984	-8.767011206	-0.040099041
Si	4.051528178	-6.873426516	-2.918926455
Si	7.327985978	-4.882956062	-3.438969261
Si	0.840329821	-8.862806425	-2.375190655
Si	-1.207764723	-8.942706718	-1.209461936
Si	-0.784275279	-9.019727906	1.101931075
Si	0.476759924	-10.90688968	1.702587855
Si	0.920998560	-10.96826211	4.005814400
Si	1.957798837	-3.073441251	-1.637878751
Si	2.531118553	-3.016156579	0.633752198
H	6.506189935	-9.993995845	0.292759661
H	8.178056093	-9.916868929	-3.180619436
H	7.144558408	-8.694141492	-4.973213511
H	9.880230044	-6.745842481	-3.642937887
H	8.059781057	-3.651995803	-3.040117552
H	7.105932945	-4.822868979	-4.907792398
H	9.819343384	-8.017308650	-0.249298581
H	9.772584616	-5.614599928	-0.195525480
H	-0.365345249	-9.994385372	-5.063805789
H	2.060495649	-4.851002658	-7.194967649
H	-1.130607283	-6.836128768	-6.649147683

H	-0.505691412	-3.743749552	-4.865985856
H	4.745625364	-1.821435354	-2.422447316
H	1.159937549	-1.865999381	-1.986133569
H	3.477797129	-1.888746806	0.844127527
H	3.342255526	-11.961081990	0.844658151
H	1.343950758	-11.975497230	-2.177558696
H	-3.773772545	-7.113350066	-0.985101695
H	-1.933115171	-10.177043330	-1.619234999
H	-2.081788709	-9.036992106	1.833647197
H	-2.426311057	-5.597514252	1.812455644
H	-2.145544320	-3.928814896	-1.441135236
H	-0.635221563	-3.996605614	1.834362846
H	1.353406252	-2.793391831	1.506772589
H	5.017514996	-11.896577590	-2.239859892
H	4.125660720	-11.024706360	-4.294271531
H	4.822898349	-2.881429154	-5.891635384
H	6.423629669	-3.770254652	0.379173437
H	-3.789661108	-5.802396767	-4.365163412
H	-3.692679378	-8.209056696	-4.471010754
H	5.187161508	-3.830637393	3.681744947
H	3.055528020	-4.723156323	4.338806620
H	8.194347468	-5.792690066	3.305324974
H	8.150800084	-8.195736668	3.203136947
H	5.601024134	-6.978823127	5.529484515
H	4.883982876	-10.053911210	3.799879549
H	2.303611704	-9.011911544	6.088851990
H	1.754647241	-12.158432950	4.314036762
H	-0.348467452	-11.076050270	4.770583941
H	-0.740928456	-7.521246282	4.660581923
H	1.066342234	-5.939490903	4.681978245
H	-0.250451442	-12.138704390	1.292313408
H	4.884774561	-6.720446420	-5.962088427
H	3.114978885	-9.863412333	-6.008172250
H	1.914119101	-8.315697077	-7.399059240
H	2.748724808	-0.675649260	-5.474820739

CISNOr-CASCI-optimized MECI

O	2.332984871	-1.777867688	-5.534839024
Si	3.777748286	-2.407793900	-6.114817290
Si	4.154172146	-2.947710435	-2.530001411
Si	1.993531627	-3.028250271	-1.632936296
Si	0.745322948	-4.957316127	-2.261094043
Si	0.348176512	-4.952738092	-4.594842032
Si	2.318806745	-4.930053864	-5.866969104
Si	3.596694869	-6.783124490	-5.245586655
Si	4.041367718	-6.860562968	-2.921579939
Si	5.296500124	-4.957858426	-2.306155359
Si	2.007511377	-6.891457397	-1.715369282
H	1.925065002	-5.209727008	-7.268542544
H	-0.445541487	-3.734084028	-4.849030494
H	4.966046968	-1.837085010	-1.976226278
H	1.326947857	-1.752065113	-1.954712065
H	4.939521702	-6.758472916	-5.888821540
H	4.505390121	-1.051194615	-6.240494788
H	2.085395087	-0.902351904	-5.271530371
Si	5.675741709	-5.005193789	0.006861701
Si	3.650706049	-4.997914840	1.212297340
Si	4.245025375	-4.964871226	3.487039099
Si	5.342677576	-6.953227990	4.063860779
Si	7.376466153	-6.964792666	2.897443388
Si	6.977211710	-6.879346207	0.583493350
Si	9.006144517	-6.822809849	-0.597450128
Si	8.590209478	-6.781589694	-2.901657363
Si	7.390235248	-8.699225481	-3.507314213
Si	5.336869645	-8.779397189	-2.365681625
Si	4.220140178	-10.791112000	-2.833527410
Si	2.116433262	-10.758747730	-1.802474471
Si	2.494252159	-10.780768810	0.517045446
Si	3.687596280	-8.845644338	1.146512512
Si	2.410790807	-6.927078738	0.619918131
Si	2.332541173	-8.605267867	-6.001980263
Si	0.395292349	-8.769725446	-4.691533145
Si	-0.907481676	-6.874315936	-5.177853419
Si	-2.954542378	-6.989998052	-4.051962535
Si	-2.502264125	-7.057303845	-1.757524384
Si	-1.33447418	-5.127925304	-1.091838468

Si	-1.094494292	-5.273293167	1.236013750
Si	0.340881144	-7.055732657	1.766715040
Si	0.602853382	-7.221652535	4.098546543
Si	2.037088643	-9.000237172	4.624732213
Si	4.079802888	-8.843404845	3.471597038
Si	5.727250984	-8.767011206	-0.040099041
Si	7.327985978	-4.882956062	-3.438969261
Si	0.840329821	-8.862806425	-2.375190655
Si	-1.207764723	-8.942706718	-1.209461936
Si	-0.784275279	-9.019727906	1.101931075
Si	0.476759924	-10.906889680	1.702587855
Si	0.920998560	-10.968262110	4.005814400
Si	2.531118553	-3.016156579	0.633752198
H	6.506189935	-9.993995845	0.292759661
H	8.178056093	-9.916868929	-3.180619436
H	7.144558408	-8.694141492	-4.973213511
H	9.880230044	-6.745842481	-3.642937887
H	8.059781057	-3.651995803	-3.040117552
H	7.105932945	-4.822868979	-4.907792398
H	9.819343384	-8.017308650	-0.249298581
H	9.772584616	-5.614599928	-0.195525480
H	-0.365345249	-9.994385372	-5.063805789
H	-1.130607283	-6.836128768	-6.649147683
H	3.477797129	-1.888746806	0.844127527
H	3.342255526	-11.961081990	0.844658151
H	1.343950758	-11.975497230	-2.177558696
H	-3.773772545	-7.113350066	-0.985101695
H	-1.933115171	-10.177043330	-1.619234999
H	-2.081788709	-9.036992106	1.833647197
H	-2.426311057	-5.597514252	1.812455644
H	-2.145544320	-3.928814896	-1.441135236
H	-0.635221563	-3.996605614	1.834362846
H	1.353406252	-2.793391831	1.506772589
H	5.017514996	-11.896577590	-2.239859892
H	4.125660720	-11.024706360	-4.294271531
H	6.423629669	-3.770254652	0.379173437
H	-3.789661108	-5.802396767	-4.365163412
H	-3.692679378	-8.209056696	-4.471010754
H	5.187161508	-3.830637393	3.681744947
H	3.055528020	-4.723156323	4.338806620
H	8.194347468	-5.792690066	3.305324974

H	8.150800084	-8.195736668	3.203136947
H	5.601024134	-6.978823127	5.529484515
H	4.883982876	-10.053911210	3.799879549
H	2.303611704	-9.011911544	6.088851990
H	1.754647241	-12.158432950	4.314036762
H	-0.348467452	-11.076050270	4.770583941
H	-0.740928456	-7.521246282	4.660581923
H	1.066342234	-5.939490903	4.681978245
H	-0.250451442	-12.138704390	1.292313408
H	3.114978885	-9.863412333	-6.008172250
H	1.914119101	-8.315697077	-7.399059240

REFERENCES

REFERENCES

(1) Sun, X. H.; Wang, S. D.; Wong, N. B.; Ma, D. D. D.; Lee, S. T.; Teo, B. K. *Inorg Chem* **2003**, *42*, 2398.

(2) Svrcek, V.; Mariotti, D.; Nagai, T.; Shibata, Y.; Turkevych, I.; Kondo, M. J Phys Chem C 2011, 115, 5084.

(3) Belomoin, G.; Therrien, J.; Nayfeh, M. *Appl Phys Lett* **2000**, *77*, 779.

(4) Dohnalova, K.; Kusova, K.; Pelant, I. *Appl Phys Lett* **2009**, *94*.

(5) Svrcek, V.; Mariotti, D.; Shibata, Y.; Kondo, M. *J Phys D Appl Phys* **2010**, *43*.

(6) Bley, R. A.; Kauzlarich, S. M. *J Am Chem Soc* **1996**, *118*, 12461.

(7) Holmes, J. D.; Johnston, K. P.; Doty, R. C.; Korgel, B. A. *Science* **2000**, *287*, 1471.

(8) Wang, Q.; Ni, H. J.; Pietzsch, A.; Hennies, F.; Bao, Y. P.; Chao, Y. M. J Nanopart Res 2011, 13, 405.

(9) Baldwin, R. K.; Pettigrew, K. A.; Garno, J. C.; Power, P. P.; Liu, G. Y.; Kauzlarich, S. M. J Am Chem Soc 2002, 124, 1150.

(10) Svrcek, V.; Kondo, M.; Kalia, K.; Mariotti, D. *Chem Phys Lett* **2009**, *478*, 224.

(11) Zou, J.; Baldwin, R. K.; Pettigrew, K. A.; Kauzlarich, S. M. *Nano Lett* **2004**, *4*, 1181.

(12) Manhat, B. A.; Brown, A. L.; Black, L. A.; Ross, J. B. A.; Fichter, K.; Vu, T.; Richman, E.; Goforth, A. M. Chem Mater 2011, 23, 2407.

(13) Lin, S. W.; Chen, D. H. *Small* **2009**, *5*, 72.

(14) Goller, B.; Polisski, S.; Wiggers, H.; Kovalev, D. *Appl Phys Lett* **2010**, *96*.

(15) He, Y.; Zhong, Y. L.; Peng, F.; Wei, X. P.; Su, Y. Y.; Lu, Y. M.; Su, S.; Gu, W.; Liao, L. S.; Lee, S. T. J Am Chem Soc 2011, 133, 14192.

(16) Li, Z. F.; Ruckenstein, E. *Nano Lett* **2004**, *4*, 1463.

(17) Warner, J. H.; Hoshino, A.; Yamamoto, K.; Tilley, R. D. Angew Chem Int Edit 2005, 44, 4550.

(18) Svrcek, V.; Sasaki, T.; Shimizu, Y.; Koshizaki, N. *Appl Phys Lett* **2006**, *89*.

(19) Gupta, A.; Swihart, M. T.; Wiggers, H. *Adv Funct Mater* **2009**, *19*, 696.

(20) Froner, E.; Adamo, R.; Gaburro, Z.; Margesin, B.; Pavesi, L.; Rigo, A.; Scarpa, M. J Nanopart Res 2006, 8, 1071.

(21) Park, J. H.; Gu, L.; von Maltzahn, G.; Ruoslahti, E.; Bhatia, S. N.; Sailor, M. J. Nat Mater 2009, 8, 331.

(22) Rastgar, N.; Rowe, D. J.; Anthony, R. J.; Merritt, B. A.; Kortshagen, U. R.; Aydil, E. S. J Phys Chem C 2013, 117, 4211.

(23) Kang, Z. H.; Liu, Y.; Tsang, C. H. A.; Ma, D. D. D.; Fan, X.; Wong, N. B.; Lee, S. T. Adv Mater 2009, 21, 661.

(24) Ben-Nun, M.; Quenneville, J.; Martinez, T. J. *J Phys Chem A* **2000**, *104*, 5161.

(25) Levine, B. G.; Coe, J. D.; Virshup, A. M.; Martinez, T. J. *Chem Phys* **2008**, *347*, 3.

(26) Heller, E. J. *J Chem Phys* **1981**, *75*, 2923.

(27) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. *Chem Phys* **1980**, *48*, 157.

(28) Wadt, W. R.; Hay, P. J. *J Chem Phys* **1985**, *82*, 284.

(29) Stanton, J. F.; Bartlett, R. J. *The Journal of chemical physics* **1993**, *98*, 7029.

(30) Becke, A. D. *J Chem Phys* **1993**, *98*, 1372.

(31) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys Rev B* **1988**, *37*, 785.

(32) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. J Phys Chem-Us 1994, 98, 11623.

(33) Finley, J.; Malmqvist, P. A.; Roos, B. O.; Serrano-Andres, L. Chem Phys Lett 1998, 288, 299.

(34) Langhoff, S. R.; Davidson, E. R. *Int J Quantum Chem* **1974**, *8*, 61.

(35) Knowles, P. J.; Werner, H. J. *Theor Chim Acta* **1992**, *84*, 95.

(36) Shiozaki, T.; Gyorffy, W.; Celani, P.; Werner, H. J. *J Chem Phys* **2011**, *135*.

(37) Hampel, C.; Peterson, K. A.; Werner, H. J. *Chem Phys Lett* **1992**, *190*, 1.

(38) Werner, H. J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schutz, M. Wires Comput Mol Sci 2012, 2, 242.

(39) Celani, P.; Werner, H. J. *J Chem Phys* **2000**, *112*, 5546.

(40) Korona, T.; Werner, H. J. *J Chem Phys* **2003**, *118*, 3006.

(41) Werner, H. J.; Knowles, P. J. *The Journal of chemical physics* **1985**, *82*, 5053.

(42) Levine, B. G.; Coe, J. D.; Martinez, T. J. *J. Phys. Chem. B* **2008**, *112*, 405.

(43) Busch, T.; Degli Esposti, A.; Werner, H. J. *J Chem Phys* **1991**, *94*, 6708.

(44) Shu, Y.; Fales, B. S.; Levine, B. G. *Nano Lett.* **2015**, *15*, 6247.

(45) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51.

(46) Shu, Y.; Hohenstein, E. G.; Levine, B. G. *J. Chem. Phys.* **2015**, *142*.

(47) Fales, B. S.; Levine, B. G. *J. Chem. Theory Comput.* **2015**, *11*, 4708.

(48) Ufimtsev, I. S.; Martinez, T. J. *J. Chem. Theory Comput.* **2009**, *5*, 1004.

(49) Ufimtsev, I. S.; Martinez, T. J. *J. Chem. Theory Comput.* **2008**, *4*, 222.

(50) Isborn, C. M.; Luehr, N.; Ufimtsev, I. S.; Martinez, T. J. J. Chem. Theory Comput. 2011, 7, 1814.

(51) Hansen, J. A.; Piecuch, P.; Levine, B. G. *J. Chem. Phys.* **2013**, *139*.

Chapter 13

Surface Structure and Silicon Nanocrystal Photoluminescence: The Role of Hypervalent Silyl Groups

"Of course, we must avoid postulating a new element for each new phenomenon"
-David J. Bohm (1917-1992)

This chapter is reproduced from paper: Surface Structure and Silicon Nanocrystal Photoluminescence: The Role of Hypervalent Silyl Groups, Y. Shu, U. R. Kortshagen, B. G. Levine, and R. J. Anthony, J. Phys. Chem. C, 119, 26683 (2015).

The efficient and tunable luminescence from quantum-confined silicon nanocrystals (SiNCs) has attracted much attention in recent years. Due to silicon's abundance and low toxicity, SiNCs are appealing for a range of applications, from biological tagging to electroluminescent devices.[1-4] While organic semiconductors can perform quite well in the visible region, SiNCs offer the advantage of efficient luminescence in the infrared range, potentially enabling the use of hybrid LEDs for remote communications. In fact, a recent hybrid organic / SiNC light-emitting device showed peak external quantum efficiency of 8.6% in the near infrared spectral region.[3] In addition, the infrared emission and ultraviolet absorption of SiNCs make them attractive for use in biological applications,

as these wavelengths are compatible with transmission through tissue. [5-8]

Attempts to optimize SiNCs for optoelectronic applications would benefit from a full understanding of the fundamental physical processes affecting the energy and yield of luminescence from the nanocrystals. Among the many issues controlling nanocrystal luminescence efficiency is the role of defects at the nanocrystal surface, which can lead to non-radiative (NR) recombination of excitons.[9-11] By developing a predictive understanding of the effects of surface structure on the emission from SiNCs, these nanocrystals can be engineered to exhibit more efficient luminescence, enabling SiNCs to compete with other light-emitting nanocrystals.

It is understood that passivation of the surface silicon atoms can eliminate dangling bonds and lead to enhanced emission properties.[12-24] Effective methods of passivation include annealing and attachment of functional groups to the SiNC surfaces, but the relationship between passivation and PL yield can be quite complex. Mangolini *et al.* explored the use of a dual-plasma reactor to synthesize SiNCs and then graft alkenes onto the nanocrystal surfaces in-flight.[12] An interesting finding from that work was that while the alkene chains were successfully bound to nanocrystal surfaces in the in-flight functionalization step, a subsequent heating of the functionalized SiNCs was required in order to raise the photoluminescence (PL) quantum yield (QY) of the nanocrystals. Simultaneously, this heating step appeared to cause a reduction in surface dangling bonds as seen in electron paramagnetic spin resonance (EPR) spectroscopy as well as a change in surface hydrogen as seen in Fourier-transform infrared spectroscopy (FTIR)

measurements. The implication from this study is that attachment of ligands is not solely responsible for enhanced QY from these SiNCs, and that the heating of the functionalized nanocrystals plays an important role in determining the defect densities and luminescence behavior of SiNC ensembles.

Thus the assignment of specific surface structures as NR recombination centers is challenging. Recent theoretical work by two of the authors[25-28] has shown that some oxygen containing defects on the oxidized silicon surface offer efficient decay pathways involving conical intersections,[29] points on the potential energy surface (PES) at which two adiabatic electronic states of a molecule or material become degenerate. At conical intersections, the coupling between electronic states is large, leading to a high probability of a NR transition. Such intersections have been established to play a crucial role in facilitating the NR decay of electronically excited molecules, and given that defects are molecule-sized features of a material which localize electronic excitations, it is not surprising that NR recombination involving defects would be facilitated by conical intersections as well. If shown to be a general principle, the theoretical characterization of conical intersections introduced by defects may become a tool for the rational optimization of nanocrystal PL.

In this work, we examine the PL properties of SiNCs produced in a nonthermal plasma reactor with respect to silyl (SiH_3) groups on their surfaces as-produced and after annealing in inert atmosphere. The goal of this work is to pinpoint the effects of heating on the SiNCs from the perspective of defect density and surface hydride species. The

effects of heating on SiNC surfaces is investigated using FTIR measurements, thermal desorption spectroscopy, and EPR spectroscopy, and is correlated to PL QYs of the nanocrystals. These experiments systematically discriminate between dangling bonds and other surface structures capable of quenching PL. We also apply computation to investigate the structural and energetic details of the experimentally observed surface defects and the mechanism of the associated NR recombination. Nonadiabatic molecular dynamics techniques based on a multireference treatment of the electronic structure are employed to investigate the role that conical intersections may play in the experimentally observed PL yields. Together, these results point towards hypervalently bound silyl anions as the defects facilitating non-radiative recombination.

13.1 Methods

13.1.1 Experimental methods

The SiNCs were produced in a low-pressure, nonthermal rf plasma reactor, as described elsewhere.[3,30-33] The precursor was silane (5% in helium) at 13 standard cubic centimeters per minute (sccm), with argon as background gas (20-35 sccm) and a flow of hydrogen (100 sccm) into the effluent of the plasma. The pressure in the reactor was 1.4 Torr and the nominal rf power was 75-90 W, conditions which lead to crystalline nanoparticles with average diameters around 4.5-5 nm as reported in previous work.[32] As SiNCs will readily oxidize upon contact with air, they were collected using a nitrogen-purged glove bag attached to the reactor to provide an inert atmosphere.

For steps involving boiling the SiNCs in solvent, the nanocrystals were immersed in anhydrous, degassed mesitylene using a N_2/vacuum Schlenk line. Dry heating experiments were also performed in a nitrogen environment. FTIR data were taken using a Nicolet Series II Magna-IR System 750 FTIR. Electron paramagnetic spin resonance (EPR) was carried out on a Bruker CW EleXsys E500 EPR spectrometer. Photoluminescence measurements were taken using an Ocean Optics, Inc. USB2000 spectrometer and an integrating sphere. Residual gas analysis was performed using a Stanford Research Systems RGA100.

13.1.2 Computational details

A combination of quantum chemical and dynamical techniques was applied to investigate the bonding of silyl defects to the surface of SiNCs and the role that these defects might play in promoting NR recombination. A small cluster model (described below) was studied to allow the application of highly accurate electronic structure methods. To investigate bonding, bond dissociations curves were computed at the QCISD(T) level of theory[34] with the 6-311++G** basis set. All internal coordinates besides a single Si-Si bond length were frozen at their ground state minimum energy structure, and the counterpoise method was applied to correct for basis set superposition error. The fully dissociated cluster was then optimized to provide a more accurate estimate of the bond dissociation energy.

The dynamics of the cluster after excitation are modeled using the ab initio multiple spawning (AIMS) approach.[35] The electronic wave function in our AIMS simulations is

provided by *ab initio* calculations done on-the-fly, for which we choose the state-averaged CASSCF.[36] AIMS is a hierarchy of approximations that allows the simulation of the full time-dependent molecular wave function following photoexcitation, including all nuclear and electronic coordinates explicitly. A CASSCF active space of two electrons in three orbitals, a state average over two states, and the 6-311+G* basis were chosen. These parameters were chosen because the resulting PES is in good agreement with the more accurate CASPT2 level of theory (as will be shown below). See Supporting Information (http://pubs.acs.org/doi/suppl/ 10.1021/acs.jpcc.5b08578, Figure S1) for images of the active orbitals upon initial excitation. Twenty simulations were run and their results averaged. The positions and momenta of the twenty initial nuclear basis functions were randomly sampled from the ground state vibrational Wigner distribution computed at the QCISD(T)/6-311++G** level of theory and placed on the lowest singlet excited states, which is optically bright. Additional analysis of the PES was performed by optimization of minima and minimal energy conical intersections at the multi-state CASPT2 level of theory,[37] which can more accurately predict the energetics of the process by adding dynamic electron correlation to the CASSCF reference wave function.

The dynamical results reported here were averaged over twenty individual AIMS simulations. The initial average positions and momenta of each simulated nuclear wave packet were randomly drawn from the ground state vibrational Wigner distribution of the model cluster computed at the QCISD(T)/6-311++G** level of theory. All electronic structure calculations were performed with the MolPro software package.[38-42] Dynamical

simulations were done using the FMSMolPro package.[43] Conical intersection optimizations were performed using the CIOpt software package.[44]

13.2 Results and Discussion

13.2.1 Effect of heating on properties of SiNCs

As-synthesized SiNCs, prepared using a nonthermal plasma reactor as described previously,[3,30,31] were heated in neat degassed mesitylene to examine the effects of heating on the photoluminescence of the samples. Heating the SiNCs in mesitylene allows effective heat transfer in a high-boiling-point solvent without otherwise reacting with the SiNC surface. Upon heating to 160°C for 1 hour, the luminescence intensity of the nanocrystals was enhanced significantly, as shown in Figure 13-1(a). The PL QY from the samples heated to 160°C reached as high as 15%, compared to around 1-5% for the unheated samples. This enhancement was achieved in the absence of functionalizing ligands, simply due to the heating of the nanocrystals. By examining the surface bonds of the samples using FTIR spectroscopy, a change can be seen in the relative abundance of the silicon tri-, di-, and monohydrides in the stretch vibration region between 2000 and 2200 cm^{-1}. We assign the peaks near 2140 cm^{-1}, 2110 cm^{-1}, and 2080 cm^{-1} as stretching vibrations from SiH_3, SiH_2, and SiH, respectively.[45,46] Though there is significant evidence to support these assignments, note that they are not uncontroversial.[47] In the heated sample, the SiH_3 peak is diminished when compared to the as-produced SiNC sample (Figure 10-1(b)). The same shift in SiH_x structure (reduction in SiH_3 contribution) is

also seen in the bending/deformation modes between 800-950cm^{-1}, while oxidation is minimal (see Supporting Information Figure S2). Importantly, the low temperature at which these changes in surface structure occur suggests that they arise from a weakly bound species.

To examine the change in the SiNC surface during heating rather than simply before and after, we used a diffuse-reflectance FTIR setup equipped with a heater in a vacuum chamber. We placed dry SiNCs onto gold-coated silicon inside the chamber, pumped it down to vacuum level, and then proceeded to heat the sample as we recorded FTIR spectra periodically. The sample was heated from room temperature to 550°C, and spectra were recorded at intervals of 25°C. The FTIR spectra from the SiH_x stretching region are presented in Figure 13-2. During this heating, the shift in SiH_x surface structure is clear to see, with a gradual reduction in tri-hydrides as heating progresses, accompanied by a shift to di- and mono-hydride coverage.

These experiments demonstrate that heating of the as-produced SiNCs leads to a reduction in SiH_3 groups at the SiNC surface and a simultaneous increase in the intensity of PL from the SiNCs, but do not establish a causal relationship between these two phenomena. To attempt to establish such a relationship, we investigated the possibility that absorbed SiH_3 radicals may be responsible for the decreased PL. Diffusion of hydrogen and SiH_x on silicon surfaces has long been studied.[46,48,49] Marra *et al.* suggested that silicon trihydrides on silicon nanoparticles can be the result of adsorption of SiH_3 on

dangling bond sites, and that these trihydrides can dissociate at temperatures below 250°C.[46] The group of Kessels examined the radicals present near the substrate during thin-film nanocrystalline silicon growth in a low-pressure silane/argon/hydrogen plasma, and showed that the silicon trihydride radical is the most prevalent.[50] If the same is true for the nonthermal plasma reactor used here, the predominance of SiH_3 in the plasma could lead to adsorbed radicals on the nanoparticles' surfaces, creating dangling bonds and limiting the PL intensity from as-produced nanoparticles.

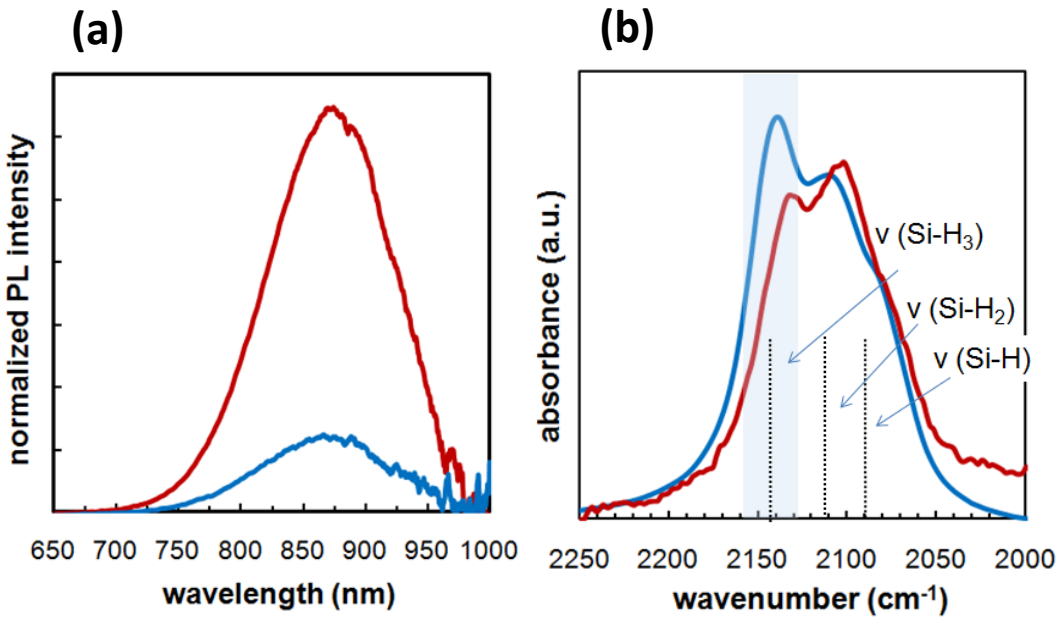


Figure 13-1. PL (a) and FTIR (b) from as-produced SiNCs (blue) and SiNCs refluxed in mesitylene to 160°C (red). The PL spectra are normalized to their respective absorption peaks (not shown), allowing direct comparison of the PL intensities of the two samples.

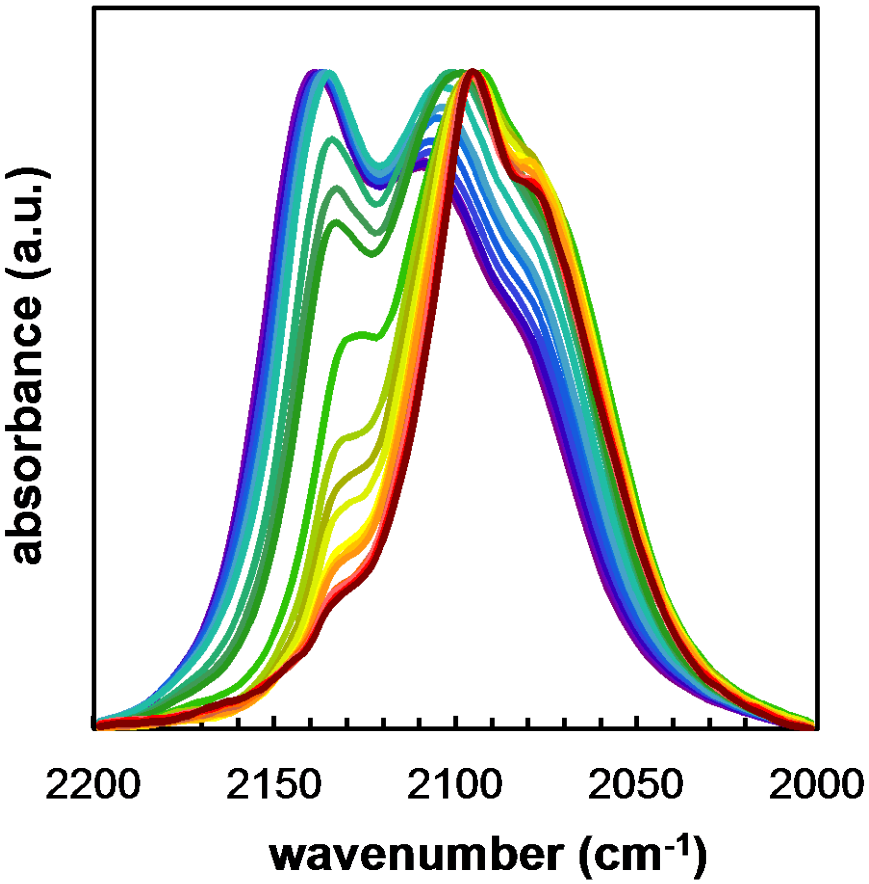


Figure 13-2. *In-situ* FTIR spectra (SiH_x region) of heated SiNCs. Purple shows the as-produced sample, with colors moving towards red as the temperature is increased to 550°C, with ΔT = 50°C. Spectra are normalized to the highest point in the displayed region.

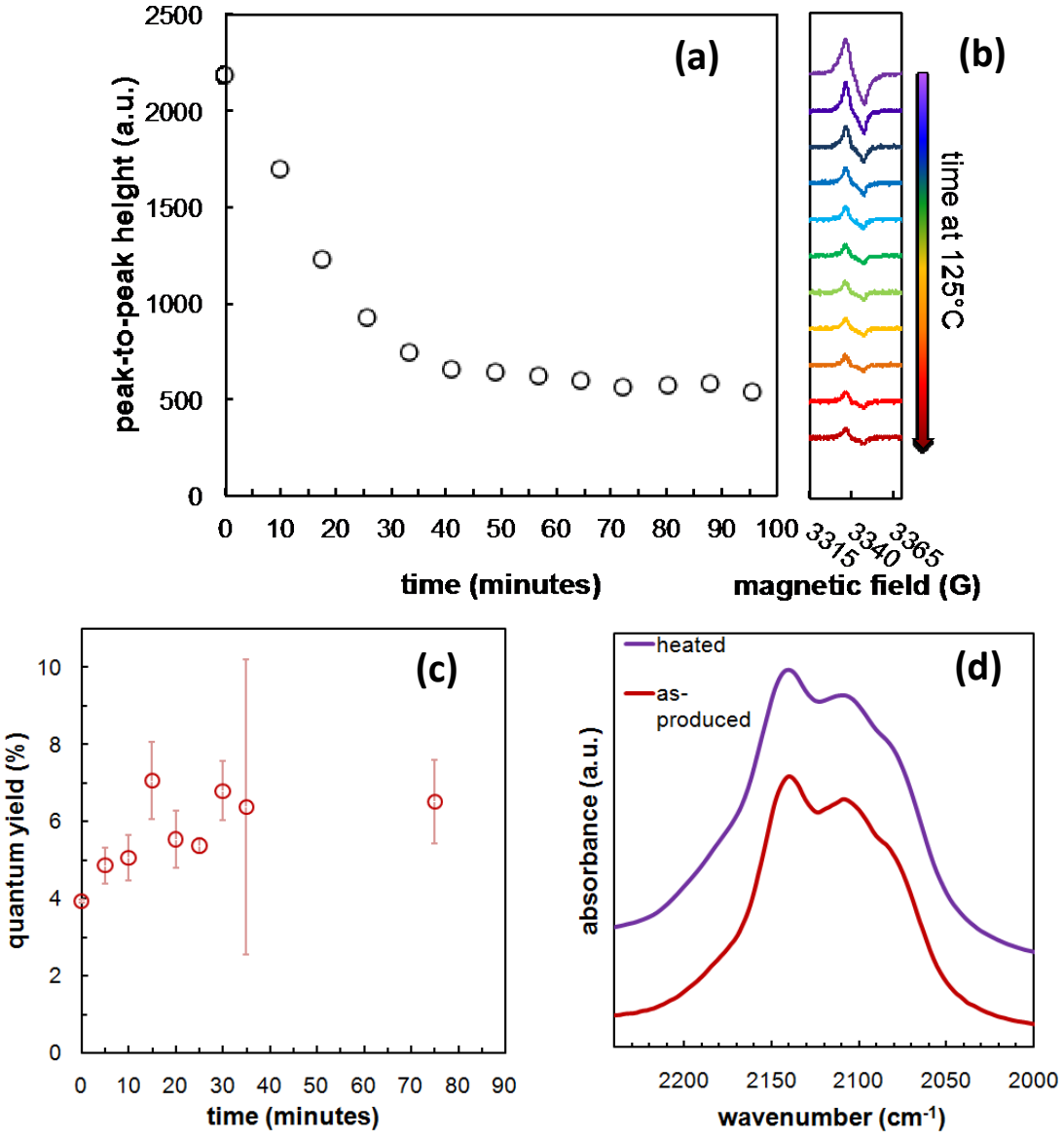


Figure 13-3. EPR signal amplitude (a) and spectral representation (b) from SiNCs as a function of *in-situ* heating time at 125°C. The PL QYs from an SiNC sample heated under the same conditions are shown in (c), and FTIR spectra of an as-produced sample and a sample heated to 125°C is pictured in (d). Error bars in the QY measurement reflect standard deviation.

To investigate the possibility that the heating step may have caused a decrease in dangling bond density via desorption of the SiH_x groups, leading to enhanced luminescence, we performed an EPR study on heated SiNCs in conjunction with PL measurements. EPR spectroscopy is used to characterize free-electron defects in materials. For silicon nanocrystals, the commonly-observed defects are the D-defect, which is associated with disorder in silicon and has a g-value of 2.0052, and the P_b-defect, which is found at Si-SiO_2 interfaces and has parallel and anti-parallel g-values of 2.0019 and 2.0086.[51] These two defects have resonances close enough to one another that the EPR signals are superimposed. The silyl radical $\cdot SiH_3$ has a g-value within spectroscopic range of these, at 2.0036, and would likely have a signal which is superimposed on the D- and P_b defects as well.[52]

We measured PL and dangling bond density both before and after heating the SiNCs to 160°C. The PL of this sample increased significantly, and the dangling bond density was decreased. However, this experiment only showed that both effects occur at 160°C, not that they are causative. To further analyze the relationship between dangling bond density and PL yield, we performed *in-situ* heating/EPR and heating/PL measurements on the nanocrystals at a lower temperature of 125°C. The EPR sample was placed into the instrument cavity and heated to 125°C while recording the EPR signal approximately every 5 minutes. The amplitude of the EPR signal decreased gradually over the course of an hour until it reached a saturation value (Figure 10-3(b)). If dangling bonds (either from D-defects or as silyl radicals) were creating non-radiative exciton decay pathways in the

SiNCs, one would expect to see a gradual increase in PL intensity with the SiNC sample heated and measured under the same conditions. However, our PL experiments showed only a very slight improvement in QY after heating (Figure 10-3(a) and (c)). Furthermore, the SiH_x stretching vibration region for the nanocrystals heated to 125°C does not display a measureable change (Figure 10-3(d)). Analysis of the EPR signal from unheated SiNCs revealed that the as-produced nanocrystals exhibit approximately one dangling bond per 100 nanocrystals—a dangling bond density that is unlikely to be responsible for a reduction in ensemble QY. These experiments and analysis discount the hypothesis that a decrease in dangling bond density (including surface SiH_3 radical density) upon annealing is directly related to the increase in PL QY that we see for these nanocrystals. When SiNCs were heated to 160°C, the dangling bond density was similarly reduced, but at this temperature the SiNC PL increased.

These EPR measurements demonstrate that the rearrangement in surface bonds and the increase in PL associated with heating of the SiNCs are decoupled from any reduction in EPR-active dangling bond density. To investigate the physical processes occurring at the SiNC surfaces during heating, we performed temperature-programmed desorption studies on the SiNCs. Thermal desorption studies have been well-used to examine the release of hydrogen and other species from silicon structures,[53-55] with peaks from effused hydrogen typically occurring between 300-400°C, and sometimes lower temperatures in the case of nanostructures. By using thermal desorption, we intended to study whether hydrogen is removed from the freestanding SiNCs at lower temperatures such as 160°C,

or whether it simply relaxes, forming more stable bond configurations. SiNCs were collected on a stainless steel mesh and then transferred to a glass microscope slide. The sample was loaded into a vacuum chamber equipped with a lamp heater and a residual gas analyzer (RGA). For control purposes, we also performed the experiments using a clean glass slide as the sample. The temperature of the sample, as measured using a thermocouple attached to the sample stage, was ramped from room temperature to 400°C at a rate of 10°C per minute with mass spectra from the RGA recorded every minute. The mass spectra from an SiNC sample and the control are shown in Figure 13-4. The most significant differences between the control spectrum and the sample spectrum came not for hydrogen species (1-2 amu), but near 29-31 amu and 58-62 amu, which derive from silyls and disilyls—SiH_x and Si_2H_{2x}. The signals for silyl groups begin to emerge between 100-200°C. The silyl signal is still present at 400°C; however, by holding the sample at that temperature and recording a spectrum after 1, 5, and 10 minutes (Figure 13-5), we saw the signal fall off. By holding at this temperature, we demonstrated that the silyl groups are indeed effusing from the SiNC surface and evaporating away, as opposed to appearing on the RGA spectrum as an experimental artifact or from a regenerative process involving the SiNC core.

It is likely that the multiple silicon hydride species we see removed from the SiNCs are primarily the result of SiH_3 desorption. In fact, Martín *et al.* observed SiH_x in similar thermal desorption studies on nanostructured porous silicon.[56] As the SiH_3 peak is the first hydride peak that we see diminishing in our FTIR studies on heated SiNCs, a similar silyl

desorption seems to be a plausible explanation for the behavior of plasma-produced SiNCs as well.

Thus, silyl groups appear to effuse from the SiNC surface at low temperature, and this effusion correlates with an increase in the PL yield, but two important questions remain unanswered. First, how are the SiH_3 groups bound to the surface, and second, by what mechanism might these EPR-inactive SiH_3 groups facilitate non-radiative recombination? We will address these questions by application of ab initio calculations.

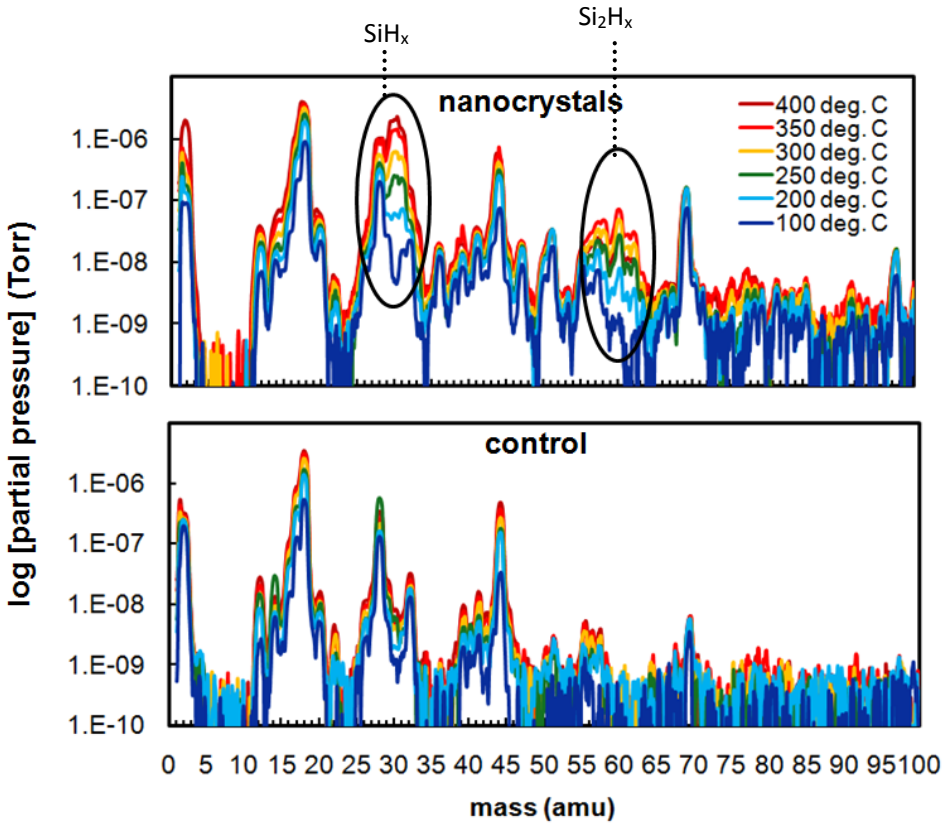


Figure 13-4. RGA mass spectra from SiNCs (top) shows removal of SiH_x groups beginning between 100-200°C. These peaks are not evident in the control mass spectra (bottom). Silyl peaks are circled in black.

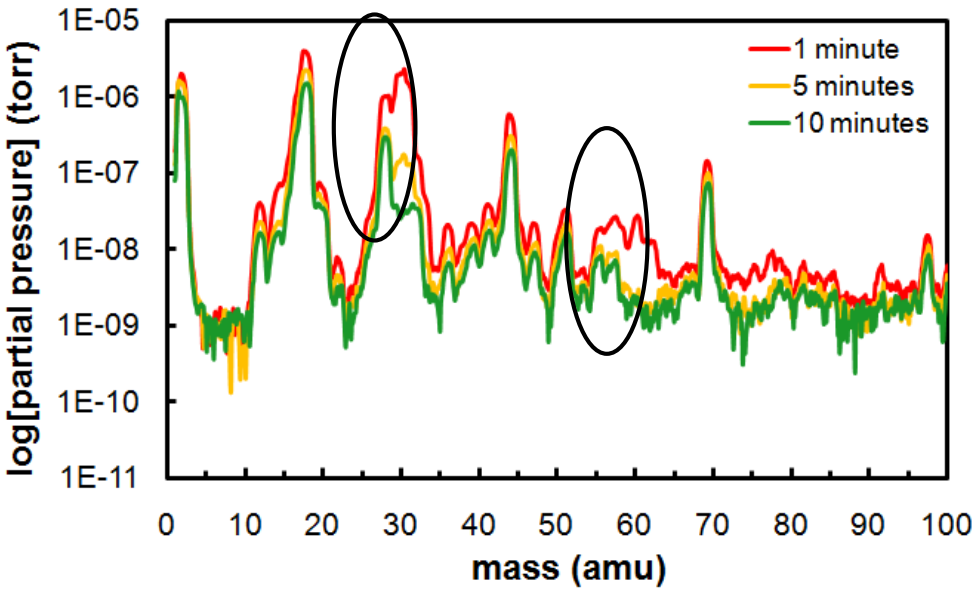


Figure 13-5. RGA mass spectra for sample held at 400°C. While silyl peaks (circled in black) are visible at 1 minute, they are reduced after 5 minutes and not visible above the noise after 10 minutes, demonstrating the removal of these groups during heating.

13.2.2 Bonding and excited state dynamics of silyl defects

The relatively low temperature at which SiH_x effusion is observed suggests that the binding of SiH_3 to the surface is not by standard covalent bonds, which have a large bond energy of ~2.4 eV and thus would be unlikely to dissociated at such a low temperature. Due to the weakly bound, EPR-inactive nature of these SiH_3 defect, we hypothesize an unusual mode of Si-Si bonding. Inspired by reports of hypervalent bonding of solvent molecules to silicon nanocrystals in solution,[57] penta-coordinated ("floating bond") defects in amorphous silicon,[58,59] and penta-coordinate intermediates in the desorption

of adsorbed SiH_3 from the silicon surface,[53] we hypothesize the hypervalent bonding of a silyl anion, SiH_3^-, to the hydrogen-passivated silicon surface. We investigate the anion species specifically because it has a closed shell electronic structure, and is therefore EPR-inactive. It is known that nanoparticles synthesized in nonthermal plasma often take a negative charge,[60-63] consistent with the hypothetical charged defect structure. Upon dissociation of the hypervalently-bound silyl defect, the product is a well-passivated surface which, again, is not detectable by EPR.

To investigate this possibility, we apply the QCISD(T) level of theory, which provides an accurate treatment of dynamic electron correlation and has often been employed in computational thermochemistry,[64] to model the bonding of a SiH_3^- to the central Si atom of a $Si(SiH_3)_3H$ cluster, which mimics the hydrogen-terminated silicon surface. In the resulting $[Si(SiH_3)_4H]^-$ cluster, the central silicon atom forms five bonds, and thus bonding is hypervalent in nature. Note that cleavage of the hypervalent bond is heterolytic and thus does not require a multireference treatment of the electronic structure. The energy of this cluster as a function of Si2-Si5 distance is shown in Figure 13-6. Si2 and Si5 are labeled in the inset, which presents the fully-optimized structure with the various Si-Si bond lengths labeled. Interestingly, the hypervalent Si-Si bond length, 3.24 Å, is considerably longer than the three shorter Si-Si bonds—2.33, 2.33, and 2.39 Å—which are comparable in length to the Si-Si bonds in crystalline silicon (2.35 Å). The dissociation energy of the hypervalently-bound silyl bond—defined as the difference between the energy of this optimized structure and that of the optimal fully dissociated structure—is

0.70 eV. This is considerably weaker than a Si-Si covalent bond in crystalline silicon (~2.4 eV), consistent with the effusion of SiH_3 at relatively low temperatures. The orbital responsible for the bonding, which is dominated by the SiH_3^- lone pair with small contributions from a d orbital on the central Si atom, is shown in the inset of Figure 13-6.

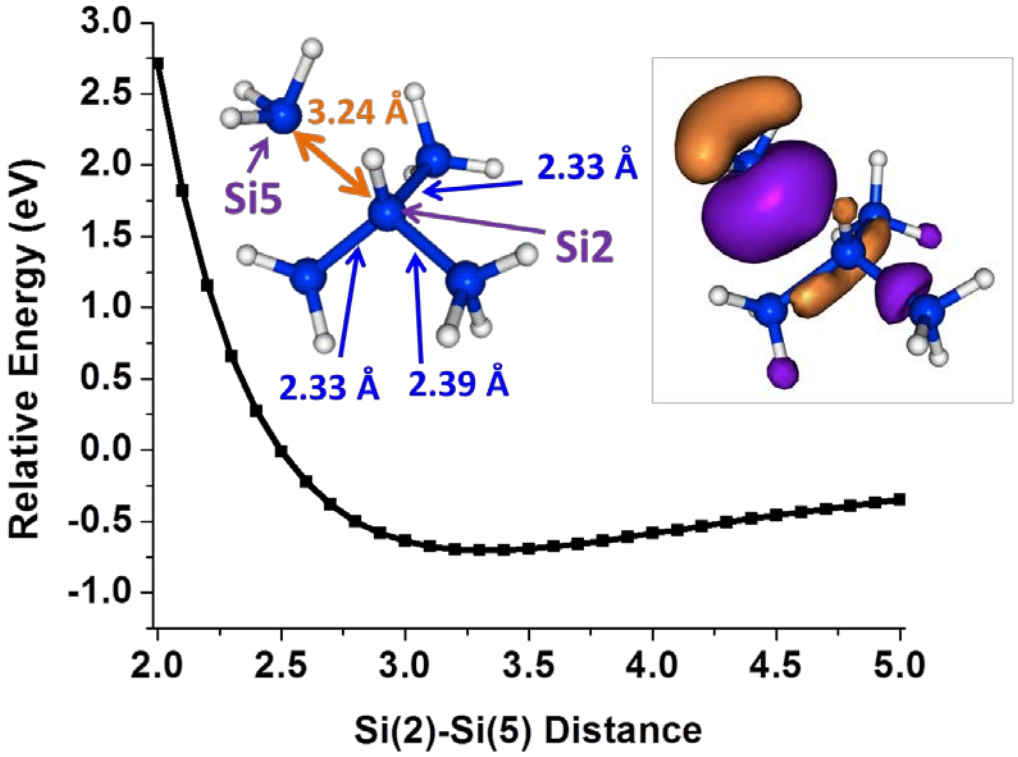


Figure 13-6. Potential energy surface as a function of silyl bond stretching in the $[Si(SiH_3)_4H]^-$ cluster computed at the QCISD(T) level of theory. The energy is shifted such that it is zero at infinite separation. The insets show the atom labels and optimized Si-Si bond lengths (left) and the orbital responsible for the bonding (right).

We now turn our attention to the role these defects play in reducing the PL QY in these SiNCs. To do so, we apply the ab initio multiple spawning (AIMS) method[35] to model the dynamics of the $[Si(SiH_3)_4H]^-$ cluster after electronic excitation. The populations of the ground and first excited states (S_0 and S_1, respectively) as a function of time after excitation are shown in Figure 13-7. Complete relaxation to the ground state is predicted within the first 250 fs after excitation. This relaxation is coincident with the photodissociation of the silyl group, as seen in Figure 13-8, which presents the Si2-Si5 distance as a function of time after excitation. This distance increases roughly linearly upon excitation, leading to the complete dissociation of the hypervalently-bound silyl group from the remainder of the cluster, which now resembles the hydrogen-terminated silicon surface. As can be seen in Figure 13-9, the energy gap between S_0 and S_1 drops dramatically during this bond breaking, and the individual trajectories regularly explore regions of near-zero energy gap. This suggests that conical intersections play a role in this non-radiative recombination process.

We investigate this possibility by application of geometry and conical intersection optimization techniques to explore the PES. The geometries and energies of the optimized ground state minimum (Franck-Condon point; FC), S_1-S_0 minimal energy conical intersection (MECI), and S_1 minimum (S_1 min) as optimized at the complete active space second order perturbation theory (CASPT2) level of theory[37] are presented in Figure 13-10. The energies of these points are relative to the ground state minimum, and thus reflect the electronic excitation energy required to reach these regions of the PES.

Assuming that the energy of this defect-localized excitation is not dependent on the size of the associated nanoparticle, these energies should not be significantly affected by our choice of a relatively small cluster model.

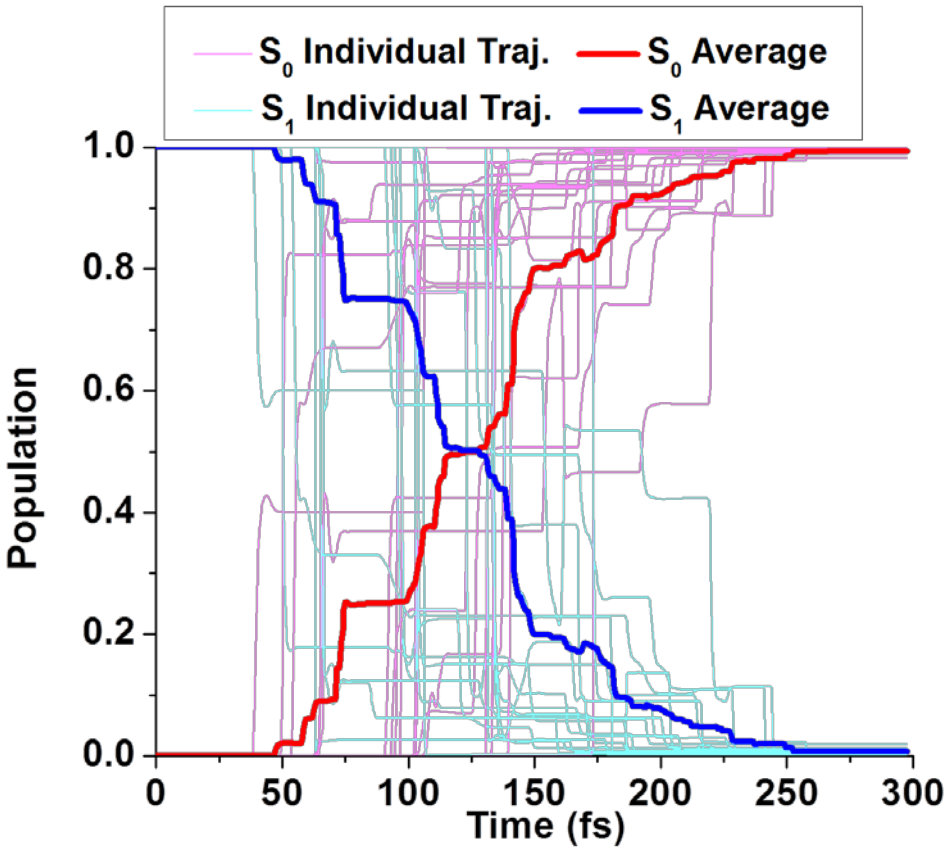


Figure 13-7. The simulated population of S_0 and S_1 of $[Si(SiH_3)_4H]^-$ as a function of time after excitation. The average over all 20 trajectories is shown in bold, with the populations of the individual trajectories shown by thin lines.

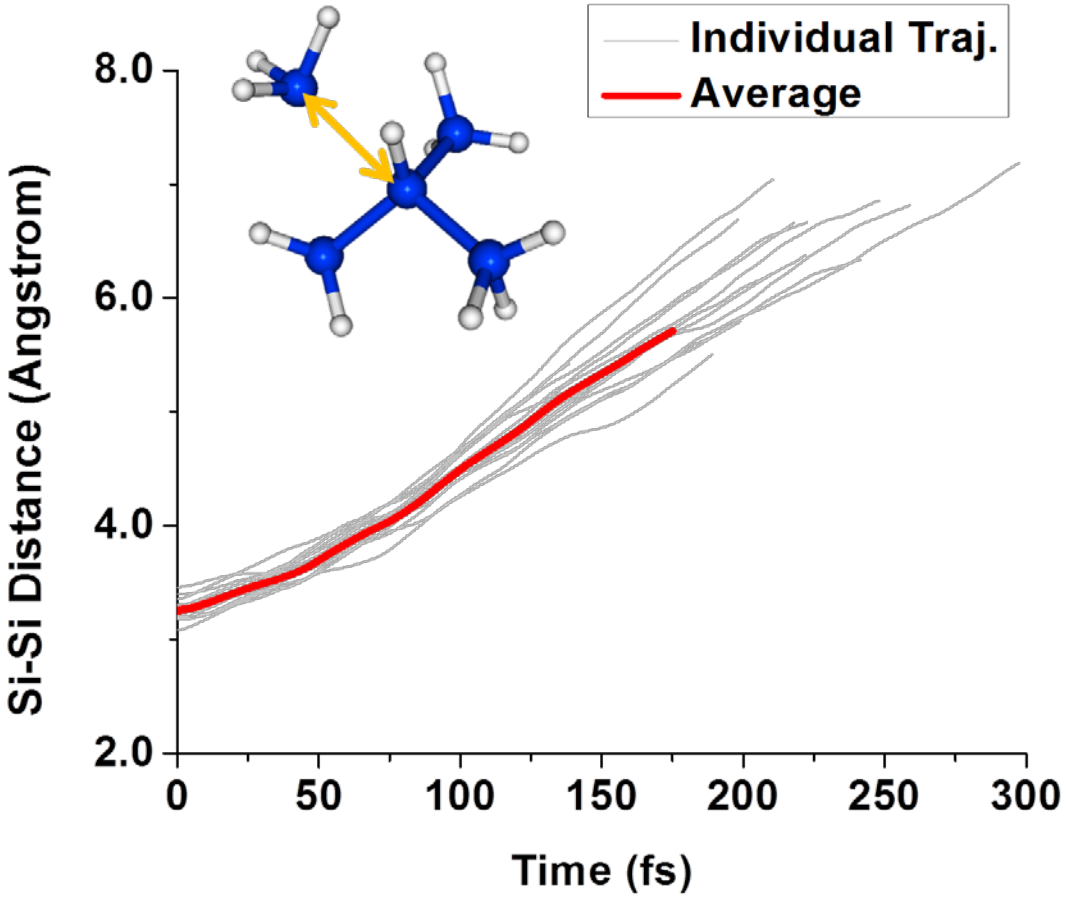


Figure 13-8. The Si2-Si5 distance (illustrated in the inset) of $[Si(SiH_3)_4H]^-$ as a function of time after excitation. The average over all 20 trajectories is shown in bold red, with the values corresponding to the individual trajectories shown by thin gray lines.

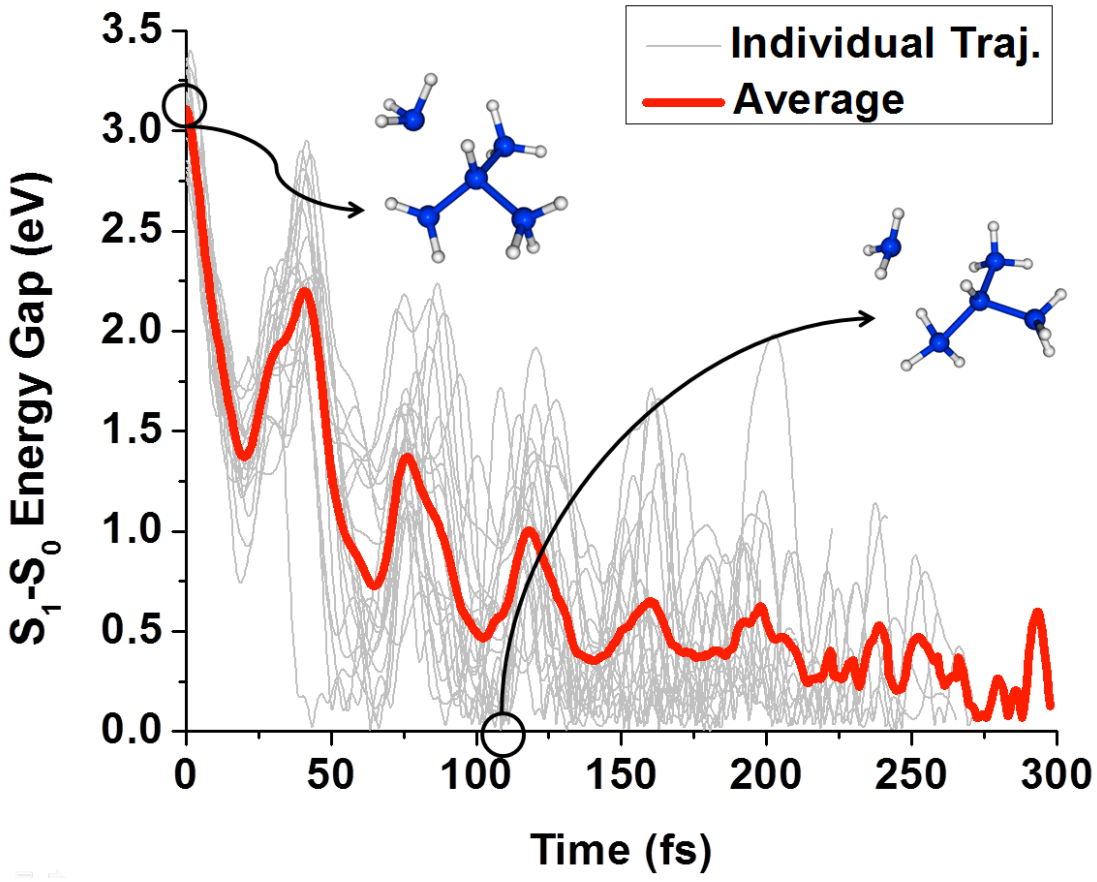


Figure 13-9. The energy gap between the two lowest many-electron states (S_0 and S_1) of $[Si(SiH_3)_4H]^-$ as a function of time after excitation. The average over all 20 trajectories is shown in bold red, with the values corresponding to the individual trajectories shown by thin gray lines. Insets show a representative geometry at the beginning of the simulation and another at a point of near-zero energy gap.

The MECI and S_1 min geometries are very similar, exhibiting a lengthening of the Si2-Si5 distance and the transfer of the terminating hydrogen atom from the $Si(Si_3)_3$ cluster to the SiH_3 unit. Such hydrogen transfer was not observed in the dynamic simulations, however, indicating that the simulated NR decay results from higher energy points on the

conical intersection seam. The highly accurate CASPT2 level of theory predicts that the S_1 min and MECI are 1.17 and 1.38 eV above the ground state minimum respectively. The energy of the MECI is noteworthy, because it indicates the excitation energy at which the intersection becomes accessible, and thus below which quenching of PL would be unlikely to be observed. This energy is low enough that the presence of a silyl defect would likely result in NR recombination via conical intersection in the SiNCs studied here, which exhibit a PL maximum at 1.30-1.42 eV (870-940 nm). The energy difference between the S_1 min and the MECI can be thought of as an activation energy for quenching, assuming thermal equilibration of the nuclear motion of the electronically excited population. The relatively small energy difference seen here (0.21 eV) explains the efficiency of the observed quenching.

Single point calculations at the complete active space self-consistent field (CASSCF) level of theory were performed at the CASPT2-optimized geometries to assess the accuracy of the CASSCF PES employed in our AIMS dynamical simulations (Figure 13-10). Qualitatively, the CASPT2 and CASSCF levels of theory agree well, with CASSCF predicting a very small energy gap at the CASPT2-optimized MECI geometry. CASSCF predicts that the MECI and S_1 min geometries are 0.72 and 0.73 eV higher in energy relative to the ground state minimum than at the more accurate CASPT2 level of theory. Such errors are not unusual with CASSCF and are not expected to have a dramatic effect on the simulated dynamics; more important is the fact that the energy difference between the MECI and S_1 min geometries is similar at the CASSCF and CASPT2 levels of theory (0.20

and 0.21 eV, respectively), suggesting that the efficient decay of the defect-localized excited state predicted by the AIMS simulations is accurate.

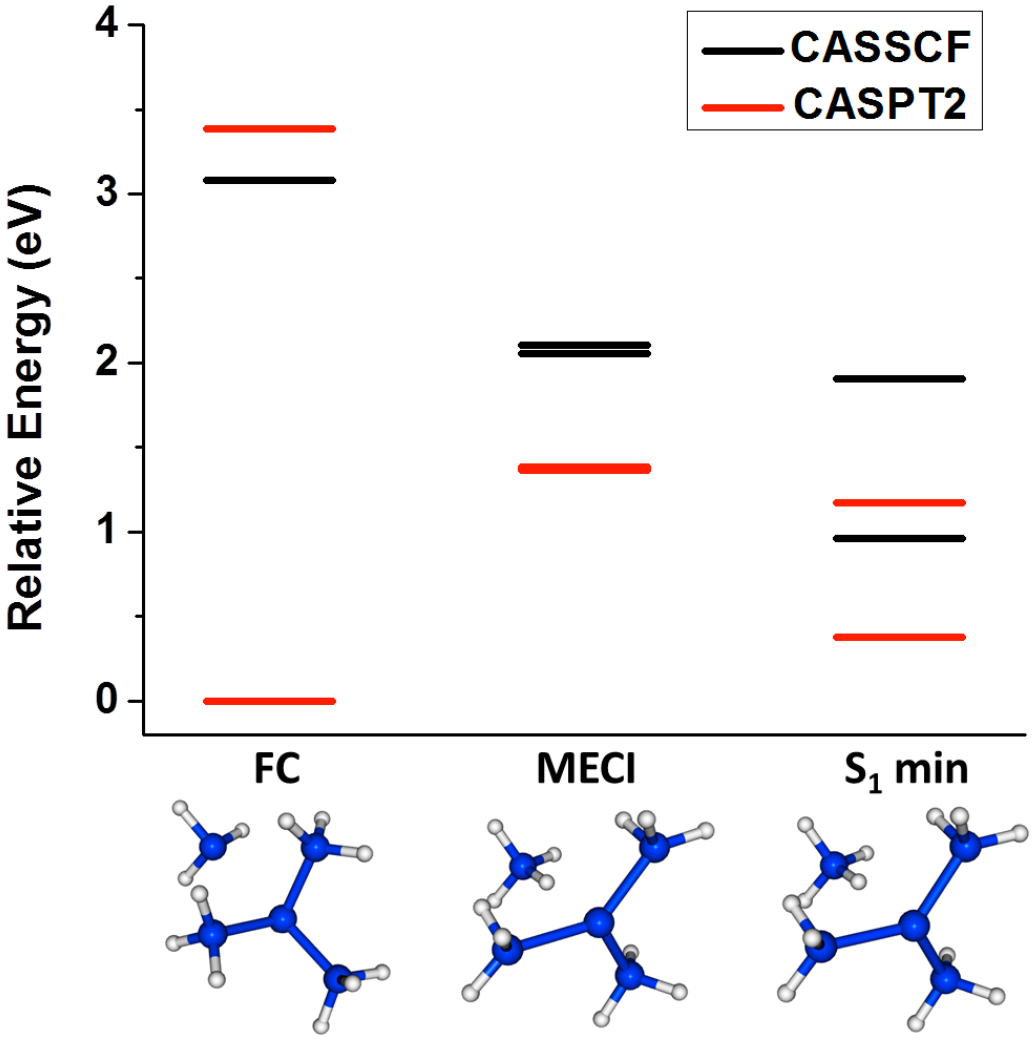


Figure 13-10. Energies and geometries of three important points on the PES of $[Si(SiH_3)_4H]^-$. S_0 and S_1 energies are shown for all geometries. All structures are optimized at the highly accurate CASPT2 level of theory. CASPT2 energies are presented in red. Single point calculations at the CASSCF level of theory used in our AIMS simulations are presented in black for comparison.

To rule out the possibility that standard, covalently bound SiH_3 groups on the SiNC surface facilitate NR recombination, we conducted a similar CASSCF/CASPT2 study of the tetra-coordinated $Si(SiH_3)_4$ cluster, which has no hypervalent bond. In this cluster, no conical intersection was observed below 5.3 eV above the ground state minimum, suggesting no clear pathway for NR recombination of excitons in the near infrared. In fact, despite significant effort no points on the excited state PES of this cluster lower than the CI at 5.3 eV could be identified. Given that covalently bound silyl introduces no features of the PES accessible in the NIR or visible ranges, the hypervalent nature of the surface silyl appears to be essential to its ability to quench NIR excitations. Details of this study are presented in Figure S3.

13.3 Conclusions

This work has shown that heating the plasma-produced SiNCs leads to a change in SiH_x (specifically SiH_3) species at the nanocrystal surface through effusion, and that this is concurrent with an improvement in PL intensity. The increased PL intensity does not, on the other hand, closely correlate with measured dangling-bond density. It is hypothesized that these features are the result of hypervalently bound SiH_3^- groups on the SiNC surface. Because of the dynamic gas environment in the nonthermal plasma reactor, it is reasonable to expect that these metastable species could be deposited at the SiNC surface. That these defects effuse at a temperature of 160º is consistent with the relative weakness of the hypervalent bonds, and the assignment of a -1 charge to this

defect leads to a closed shell electronic structure which is consistent with the absence of an associated EPR signal. First principles molecular dynamics calculations on a cluster model containing such a hypervalently bound SiH_3^- defect predict fast NR recombination via a conical intersection. This intersection is accessible at near infrared energies comparable to the PL energy of the SiNCs studied here, and thus this intersection provides a viable explanation for the correlation between the decreasing SiH_3 population and increasing PL yield.

REFERENCES

REFERENCES

(1) Erogbogbo, F.; Yong, K. T.; Roy, I.; Xu, G. X.; Prasad, P. N.; Swihart, M. T. *Acs Nano* **2008**, *2*, 873.

(2) Biaggi-Labiosa, A.; Sola, F.; Resto, O.; Fonseca, L. F.; Gonzalez-Berrios, A.; De Jesus, J.; Morell, G. Nanotechnology 2008, 19.

(3) Cheng, K. Y.; Anthony, R.; Kortshagen, U. R.; Holmes, R. J. *Nano Lett* **2011**, *11*, 1952.

(4) Puzzo, D. P.; Henderson, E. J.; Helander, M. G.; Wang, Z. B.; Ozin, G. A.; Lu, Z. H. Nano Lett 2011, 11, 1585.

(5) Weissleder, R. *Nat Biotechnol* **2001**, *19*, 316.

(6) Gao, X. H.; Cui, Y. Y.; Levenson, R. M.; Chung, L. W. K.; Nie, S. M. Nat Biotechnol 2004, 22, 969.

(7) Yu, Y. X.; Hessel, C. M.; Bogart, T. D.; Panthani, M. G.; Rasch, M. R.; Korgel, B. A. Langmuir 2013, 29, 1533.

(8) Nishimura, H.; Ritchie, K.; Kasai, R. S.; Morone, N.; Sugimura, H.; Tanaka, K.; Sase, I.; Yoshimura, A.; Nakano, Y.; Fujiwara, T. K.; Kusumi, A. J Cell Biol 2013, 202, 967.

(9) Lannoo, M.; Delerue, C.; Allan, G. *J Lumin* **1993**, *57*, 243.

(10) Dovrat, M.; Goshen, Y.; Jedrzejewski, J.; Balberg, I.; Sa'ar, A. *Phys Rev B* **2004**, *69*.

(11) Lee, B. G.; Hiller, D.; Luo, J. W.; Semonin, O. E.; Beard, M. C.; Zacharias, M.; Stradins, P. Adv Funct Mater 2012, 22, 3223.

(12) Mangolini, L.; Kortshagen, U. *Adv Mater* **2007**, *19*, 2513.

(13) Cheylan, S.; Elliman, R. G. *Appl Phys Lett* **2001**, *78*, 1912.

(14) Godefroo, S.; Hayne, M.; Jivanescu, M.; Stesmans, A.; Zacharias, M.; Lebedev, O. I.; Van Tendeloo, G.; Moshchalkov, V. V. Nat Nanotechnol 2008, 3, 174.

(15) Mangolini, L.; Jurbergs, D.; Rogojina, E.; Kortshagen, U. *J Lumin* **2006**, *121*, 327.

(16) Hua, F. J.; Swihart, M. T.; Ruckenstein, E. *Langmuir* **2005**, *21*, 6054.

(17) Gupta, A.; Wiggers, H. *Nanotechnology* **2011**, *22*.

(18) Chen, C. H.; Chen, Y. F.; Shih, A.; Lee, S. C. *Phys Rev B* **2002**, *65*.

(19) Cho, K. S.; Park, N. M.; Kim, T. Y.; Kim, K. H.; Sung, G. Y.; Shin, J. H. Appl Phys Lett 2005, 86.

(20) Fernandez, B. G.; Lopez, M.; Garcia, C.; Perez-Rodriguez, A.; Morante, J. R.; Bonafos, C.; Carrada, M.; Claverie, A. J Appl Phys 2002, 91, 798.

(21) Comedi, D.; Zalloum, O. H. Y.; Mascher, P. *Appl Phys Lett* **2005**, *87*.

(22) Jung, Y. J.; Yoon, J. H.; Elliman, R. G.; Wilkinson, A. R. *J Appl Phys* **2008**, *104*.

(23) Seraphin, A. A.; Ngiam, S. T.; Kolenbrander, K. D. *J Appl Phys* **1996**, *80*, 6429.

(24) Brawand, N. P.; Voros, M.; Galli, G. *Nanoscale* **2015**, *7*, 3737.

(25) Shu, Y. A.; Levine, B. G. *J Phys Chem C* **2014**, *118*, 7669.

(26) Shu, Y. A.; Levine, B. G. *J Chem Phys* **2013**, *139*.

(27) Shu, Y. N.; Levine, B. G. *J Phys Chem C* **2015**, *119*, 1737.

(28) Shu, Y. N.; Fales, B. S.; Levine, B. G. *Nano Lett* **2015**, *15*, 6247.

(29) Domcke, W.; Yarkony, D. R. *Annual review of physical chemistry* **2012**, *63*, 325.

(30) Mangolini, L.; Thimsen, E.; Kortshagen, U. *Nano Lett* **2005**, *5*, 655.

(31) Anthony, R. J.; Rowe, D. J.; Stein, M.; Yang, J. H.; Kortshagen, U. Adv Funct Mater 2011, 21, 4042.

(32) Anthony, R.; Kortshagen, U. *Phys Rev B* **2009**, *80*.

(33) Kramer, N. J.; Anthony, R. J.; Mamunuru, M.; Aydil, E. S.; Kortshagen, U. R. J Phys D Appl Phys 2014, 47.

(34) Pople, J. A.; Headgordon, M.; Raghavachari, K. *J Chem Phys* **1987**, *87*, 5968.

(35) Ben-Nun, M.; Quenneville, J.; Martinez, T. J. *J Phys Chem A* **2000**, *104*, 5161.

(36) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. *Chem Phys* **1980**, *48*, 157.

(37) Finley, J.; Malmqvist, P. A.; Roos, B. O.; Serrano-Andres, L. Chem Phys Lett 1998, 288, 299.

(38)Werner, H.-J.; Knowles, P.; Knizia, G.; Manby, F.; Schütz, M.; Celani, P.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G. See http://www. molpro. net 2010.

(39) Knowles, P. J.; Werner, H. J. *Chem Phys Lett* **1985**, *115*, 259.

(40) Werner, H. J.; Knowles, P. J. *J Chem Phys* **1985**, *82*, 5053.

(41) Celani, P.; Werner, H. J. *J Chem Phys* **2000**, *112*, 5546.

(42) Werner, H.-J. *Molecular Physics* **1996**, *89*, 645.

(43) Levine, B. G.; Coe, J. D.; Virshup, A. M.; Martinez, T. J. *Chem Phys* **2008**, *347*, 3.

(44) Levine, B. G.; Coe, J. D.; Martinez, T. J. *J Phys Chem B* **2008**, *112*, 405.

(45) Jariwala, B. N.; Kramer, N. J.; Petcu, M. C.; Bobela, D. C.; van de Sanden, M. C. M.; Stradins, P.; Ciobanu, C. V.; Agarwal, S. J Phys Chem C 2011, 115, 20375.

(46) Marra, D. C.; Edelberg, E. A.; Naone, R. L.; Aydil, E. S. J Vac Sci Technol A 1998, 16, 3199.

(47) DeBenedetti, W. J. I.; Chabal, Y. J. *Journal of Vacuum Science & Technology A* **2013**, *31*.

(48) Bakos, T.; Valipa, M. S.; Maroudas, D. *J Chem Phys* **2006**, *125*.

(49) Gates, S.; Greenlief, C.; Beach, D. *The Journal of chemical physics* **1990**, *93*, 7493.

(50) Kessels, W.; Nadir, K.; Van De Sanden, M. *J Appl Phys* **2006**, *99*.

(51) Pereira, R. N.; Rowe, D. J.; Anthony, R. J.; Kortshagen, U. *Phys Rev B* **2011**, *83*.

(52) Chatgilialoglu, C. *Chem Rev* **1995**, *95*, 1229.

(53) Ceriotti, M.; Bernasconi, M. *Phys Rev B* **2007**, *76*, 245309.

(54) Gupta, P.; Colvin, V.; George, S. *Phys Rev B* **1988**, *37*, 8234.

(55) Rivolo, P.; Geobaldo, F.; Rocchia, M.; Amato, G.; Rossi, A. M.; Garrone, E. Phys Status Solidi A 2003, 197, 217.

(56) Martin, P.; Fernandez, J. F.; Sanchez, C. *Mat Sci Eng B-Solid* **2004**, *108*, 166.

(57) Wheeler, L. M.; Neale, N. R.; Chen, T.; Kortshagen, U. R. *Nat Commun* **2013**, *4*.

(58) Pantelides, S. T. *Phys Rev Lett* **1986**, *57*, 2979.

(59) von Keudell, A.; Abelson, J. R. *Phys Rev B* **1999**, *59*, 5791.

(60) Goree, J. *Plasma Sources Sci T* **1994**, *3*, 400.

(61) Khrapak, S. A.; Morfill, G. E. *Phys Plasmas* **2008**, *15*.

(62) Gatti, M.; Kortshagen, U. *Phys Rev E* **2008**, *78*.

(63) Galli, F.; Kortshagen, U. R. *Ieee T Plasma Sci* **2010**, *38*, 803.

(64) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. J Chem Phys 1998, 109, 7764.

Chapter 14

Recombination Mechanism of Tantalum Nitride Ta_3N_5: Why Is a Sacrificial Reagent Required?

"There is no quantum world. There is only an abstract quantum physical description. It is wrong to think that the task of physics is to find out how nature is. Physics concerns what we can say about nature"

-Niels H.D. Bohr (1885-1962)

Beside the global consumption of energy increases every year, short comings of fossil fuels as well as environmental problems associated with it, developing efficient ways to utilize solar energy is an active field of research in the recent decades. The total solar energy that irradiates on the surface of the Earth is 1.3 X 10^5 TW which is four orders of magnitude larger than the global energy consumption (17.7 TW as of year 2015). Photocatalyst to produce clean chemical fuels like water splitting into hydrogen and oxygen,[1-3] is one of the most promising ways in utilizing the solar energy. The first solar water splitting process was discovered in 1972 by utilizing TiO_2 electrode.[4] Due to the 3.0-3.2 eV wide band gap of the TiO_2 photocatalyst, only 4% of the incoming solar energy was utilized.[5] Hence developing photocatalyst with visible range band gap is of great importance[5-7] with numerous attempts.[8-30] The overall solar water splitting process can be generally divided into three steps including light absorption, charge separation and

oxidation reduction of water molecule (see Figure 14-1). Figure 14-1 introduces both static (a) and dynamical (b) ways of thinking the solar water splitting process. Static picture does not consider nuclear motion of the materials and hence rendering the static band structures. The holes and particles are generated upon excitation and migrates towards the surfaces of the material if they can be separated. While in the dynamical picture, nuclear motion are being considered and the many electron state is a function of nuclear geometry which changes as time propagates. By simply considering the non-charge transfer states only, holes and particles are being generated after excitation as a pair locates on the same state. The charge separation step describes the process of the hole-particle pair successfully localizes to the surfaces state without quenching by non-radiative recombination. Hence, one of the key difference of the two ways of looking at the solar water splitting process lies at the way of describing the charge separation process. Static picture provides a visualization of the hole particle separation as shown in Figure 14-1(a) that the hole particle are spacially separated. While dynamical picture provides a visualization of the non-radiative recombination process which hinders the hole particle separation. Figure 14-1(b) shows that the existence of conical intersection provides a possible path way of hole particle annihilation.

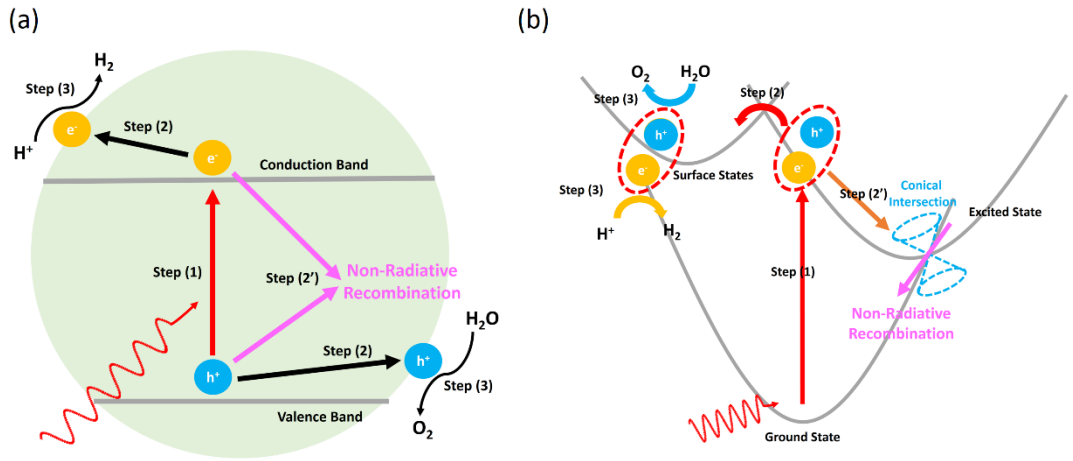


Figure 14-1. Processes involved in photocatalytic overall water splitting on a heterogeneous photocatalyst in a (a) quasi-particle picture, (b) many electron wavefunction picture.

As a thermodynamical unfavorable process, water splitting requires 238 kJ/mol of Gibbs free energy. In order to overcome the oxidation reduction potential of water, the band gap of the water splitting photocatalyst must be above 1.23 eV. Researches in recent years[31-39] have been showing that the suitable photocatalyst for overall water splitting has metal center on its highest oxidation state with d^0 (Ti^{4+}, Zr^{4+}, Nb^{5+}, Ta^{5+} and W^{6+}) or d^{10} (Ga^{3+}, In^{3+}, Ge^{4+}, Sn^{4+} and Sb^{5+}) electronic configuration and associated light absorption process involves ligands p orbitals to transition metal empty d orbital.[2,3,6,40]

Tantalum Nitride (Ta_3N_5) has been applied as a photocatalyst for water splitting with visible light range band gap. It was first discovered by Hitoki and Domen et. al.[41] Since then, Ta_3N_5 has been extensively studied,[42-58] partially due to its relatively small band gap (2.1 eV)[59] such that 45% of the solar spectrum energy can be utilized,[58] as well as its

stability (although nitrides are less stable compared with oxides in water, they are stable enough to be used as photocatalyst).[58,60-65] The reduction of the band gap of Ta_3N_5 compared with tantalum oxide is due to the weaker electronegativity of N as compared with that of O, which shifts the valence band to higher potential.[3,59]

However, due to the fast recombination process, the application of Ta_3N_5 as photocatalyst for water splitting requires the presence of sacrificial reagents.[41,61-63] The sacrificial reagent removes the hole or excited electron from photocatalyst upon excitation before the non-radiative recombination of the photocatalyst happens. It either plays a role as an electron donor which recombines the hole of the photocatalyst and leaves the excited electron in photocatalyst to reduce water molecule to generate hydrogen; or the sacrificial reagent plays a role as an electron acceptor such that the photocatalyst oxidizes water to generate oxygen. Hence when sacrificial reagent being used, only half reaction can happen which seriously hinders the practical application of the photocatalyst in solar water splitting scheme. Equation (14-1) and (14-2) show the overall reaction of water splitting by utilizing the two different types of sacrificial reagents CH_3OH and Ag^+,

$$CH_3OH+H_2O \rightarrow CO_2+3H_2 \quad (14\text{-}1)$$

$$4Ag^+ +2H_2O \rightarrow O_2+4Ag^0+4H^+ \quad (14\text{-}2)$$

Previous theoretical and experimental studies have focused on the geometrical and electronic structures of Ta_3N_5.[40,59,66-72] Density functional theory (DFT) calculations indicate that Ta has distorted octahedral coordination, with calculated band gap varies

from 1.1 to 3.1 eV depend on the theoretical levels.[40,66,70-72] Giulia Galli et al. reported that Ta_3N_5 is an indirect band-gap material with 0.2-0.3 eV lower than the direct gap.[40] These studies have provided us detailed information of the ground state stable structure of the materials as well as the electronic structures upon excitation from a theoretical perspective. However, the recombination mechanism and excited state dynamics has not been investigated yet to the best of our knowledge. As one of the most important efficiency limiting processes, non-radiative recombination plays a significant role in charge separation step. Successful charge separation is the key in developing photocatalyst. Since it competes with the non-radiative recombination process, a comprehensive understanding of the non-radiative recombination mechanism will help us develop better photocatalyst in the future.

In this chapter, we apply *ab initio* molecular dynamics and multi-reference electronic structure theories to investigate the excited state and ground state properties as well as the dynamics of the Ta_3N_5 with our model cluster (the model cluster will be introduced into detail in the following sections). We will address three questions here: 1) What types of nuclear and electronic motions are responsible for the non-radiative recombination process? 2) What are the energy levels of those important geometries on potential energy surfaces (PESs)? 3) How the energy level of the minimal energy conical intersection (MECI) (a CI geometry is where two or more potential energy surfaces are degenerate) introduced by nuclear deformation is related to the fact that sacrificial reagent is required?

14.1 Methods

14.1.1 Nonadiabatic molecular dynamics simulations

Ab initio multiple spawning (AIMS) algorithm[73,74] was employed to simulate the excited state dynamics of our model Ta_3N_5 cluster ($Ta_2N_{10}H_{20}$). AIMS is a hierarchy of approximations that solves full molecular time-dependent Schrödinger equation. The nuclear wavefunction in AIMS is expanded as a linear combination of time-dependent basis of frozen Gaussian functions[75] and the time-dependent coefficients of the basis functions are obtained by integrating the time-dependent Schrödinger equation. The average positions and momenta of the frozen Gaussian basis functions evolve on the adiabatic PES and are computed on-the-fly at the multi-reference *ab initio* electronic structure level of theory. In our simulation, the state averaged complete active space self-consistent field (SA-CASSCF) theory[76] was employed to obtain the electronic wavefunctions. An active space of four electrons in six orbitals and state average two were used. The LANL2DZ basis set and effective core potentials (ECP)[77] were used for Ta and 6-31G** basis set was used for N and H atoms (We will denote this combination of basis set and ECP as non-polarized basis set in the following text). These parameters were chosen such that low-lying excited state characters are in good agreement with equation-of-motion coupled cluster with single and double excitations (EOM-CCSD)[78] level of theory which is calculated with the LANL2TZ and f polarization basis set and ECP for Ta atom and 6-31G** basis set for N and H atoms (We will denote this combination of basis set and ECP as polarized basis set in the following text). All dynamical simulations were

initiated on the first excite state (S_1) at the ground state minimum geometry optimized by DFT with B3LYP functional[79-81] employing non-polarized basis set. Twenty initial nuclear basis functions were placed on S_1 state of the clusters with their initial positions and momenta randomly sampled from the ground state vibrational Wigner distribution. The nuclear wavefunction is propagated using an adaptive multiple time step algorithm with classical and quantum degrees of freedom integrated by velocity Verlet and second-order Runge-Kutta shcemes with a 0.5 fs time step.[74] The widths of the frozen Gaussian basis functions were chosen to be 20.0, 6.0, and 30.0 $Bohr^2$ for tantalum, hydrogen, and nitrogen atoms, respectively. The frozen Gaussian width for Ta is estimated by maximizing the overlap between the frozen Gaussian wave function and the exact wave function (it is worth noting that the dynamics are not strongly affected by the chosen of Gaussian width).[82] All dynamical calculations were performed with FMSMolPro package.[74]

14.1.2 Ab-Initio calculations

Due to the fact that SA-CASSCF is known lacking of dynamical correlations, we evaluate the accuracy of SA-CASSCF PES employed in the dynamical simulation by comparing the energies of important geometries on PESs with higher level of theory. We optimize the MECI and S_1 minimum(S_1 min) geometries on SA-CASSCF level of theory with the non-polarized basis set as employed in AIMS simulations and multi-state complete active space second-order perturbation (MS-CASPT2) level of theory[83] with polarized basis set. Additional single point calculations at SA-CASSCF with non-polarized basis set and

MS-CASPT2, multi-reference configuration interaction with single and double excitations and Davidson correction[84] (MRCI(Q)) level of theories with polarized basis set were performed on these important geometries. Both MS-CASPT2 and MRCI(Q) level of theories employ the same active space as used in the dynamical simulations with 6 occupied orbitals besides active orbitals being correlated.

All *ab initio* quantum chemistry calculations including DFT, EOM-CCSD, SA-CASSCF, MS-CASPT2 and MRCI(Q) were performed in MolPro.[85-91] MECI and S_1 min geometries were optimized with MolPro in conjunction with CIOpt.[86,92,93]

14.2 Results and Discussions

14.2.1 Model cluster and ground state minimum

The model cluster $Ta_2N_{10}H_{20}$ was chosen due to a good balance between computational cost and a correct description of the excited state character obtained by this representative cluster. As has been introduced in section 14.1 that the excitation character of the Ta_3N_5 material is described by the N p orbital to Ta d orbital. The surface of the cluster is capped with H atoms to remove the dangling bond defect. As has been mentioned in the previous section, the ground state minimum geometry was optimized by DFT employing B3LYP functional and LANL2DZ basis set with ECPs. The ground state minimum geometry is shown in Figure 14-2(a) with corresponding Ta-N bond lengths. The model cluster has a center of symmetry which locates at the center of Ta-N-Ta-N plane (the dihedral angle of TaNTaN is 0.8 degree) as indicated by the yellow dotted arrow in

Figure 14-2(a). This center of symmetry is also reflected by the equivalent Ta-N bond length associated with the N atoms at symmetrical positions as shown by the red numbers in Figure 14-2(a). The symmetrical positions here are defined that the N atom, the center of symmetry and another N atom forms nearly 180 degree angle. Figure 14-2(a) yellow arrows are pointing to an example pair of N atoms with symmetrical positions to each other. Hence the molecule has C_i symmetry. The twisted octahedral structure described above, has N-Ta-N varies from 75.9 to 112.5 degrees and 150.8 to 161.1 degrees as compared with that of perfect octahedral geometry, 90 and 180 degrees. This twisted octahedral structure is consistent with that has been previous reported by DFT studies of periodic system.[66]

The EOM-CCSD calculation shows the S_1 state is the first bright state with an excitation energy of 4.42 eV. This is much higher excitation energy compared to the band gap of bulk Ta3N5 (~2.1 eV), the discrepancy is caused by the limited size of the cluster. Figure 14-2(b) shows the HF orbitals with large EOM-CCSD amplitudes associated with the S_1 state. The S_1 state is described by the excitation from N p orbitals to Ta d orbitals which is consistent with that has been observed from the previous studies.[2,3,6,40,65] The consistency of the electronic structures with the previously reported theoretical and experimental results supports that our chosen of the model cluster is a reasonable representative of the bulk material.

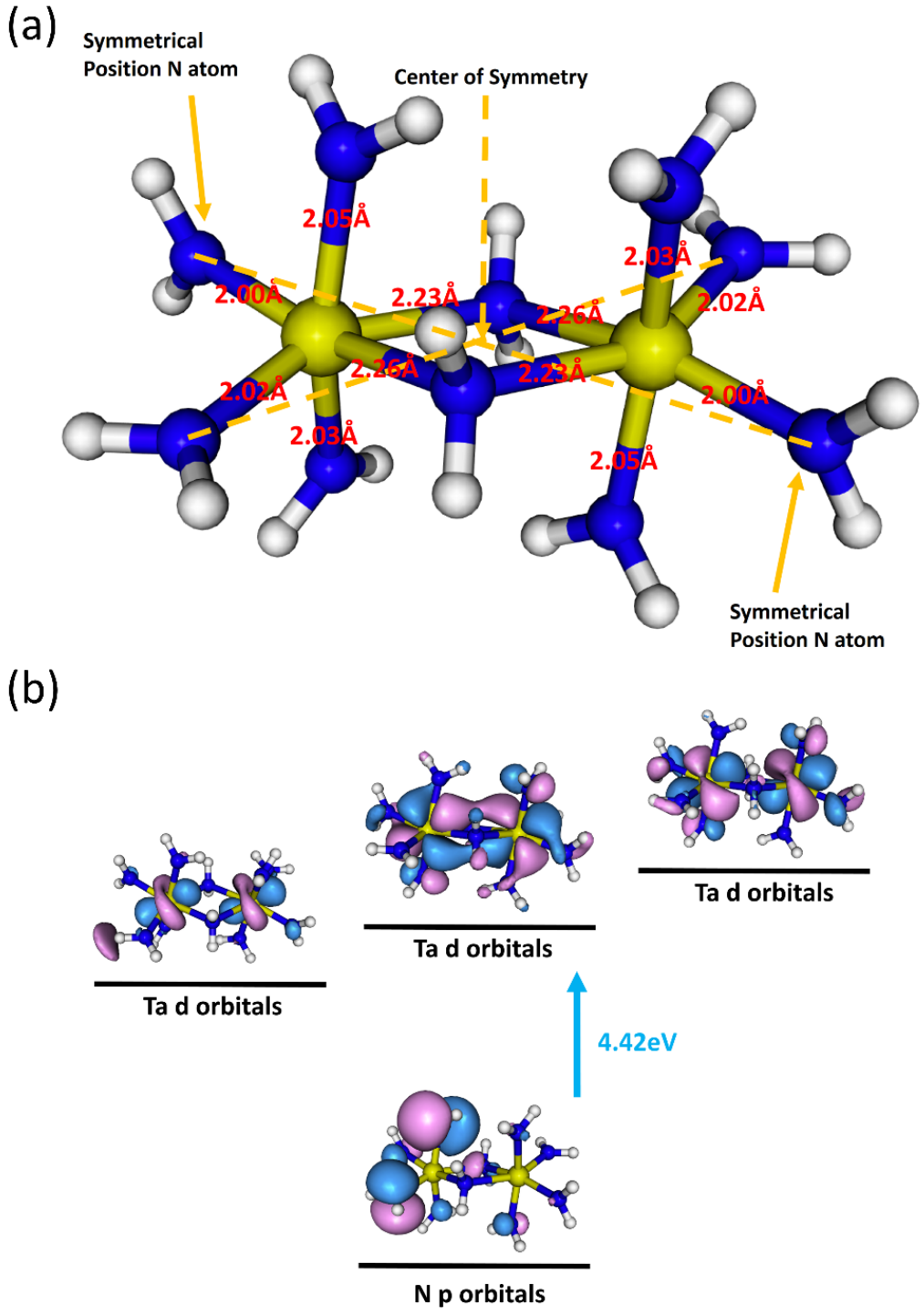


Figure 14-2. (a) Ground state minimum geometry of the model cluster $Ta_2N_{10}H_{20}$, with Ta, N and H represented by yellow, blue and grey colors respectively. The yellow dotted arrow indicates the center of symmetry, the yellow solid arrow shows a sample pair of N

Figure 14-2 (cont'd). atoms with symmetric positions. The red numbers show the Ta-N bond length. (b) The S_1 state has an excitation energy of 4.28eV computed on EOM-CCSD level, the HF orbitals indicate the S_1 character is mostly described by excitations from N p orbitals to Ta d orbitals.

14.2.2 Dynamical simulation of the model cluster

The electronic structure theory employed in dynamical simulation, SA2-CASSCF(4/6) with non-polarized basis set predicts the vertical excitation energy of S_1 state as 3.91 eV at the ground state minimum geometry. This indicates 0.51 eV discrepancy as compared to EOM-CCSD result, which is common at SA-CASSCF theoretical level due to the lack of dynamical correlation. We will investigate more important features of the adiabatic PES employed in dynamical simulations to evaluate the accuracy in the following sections. AIMS simulation shows that S_1 state population begins to decay at 25 fs. 99% of the excited state population transfers to the ground state in 150 fs as shown in Figure 14-3. This population transfer is accompanied by the S_1-S_0 energy gap drop. Figure 14-4 shows the S_1-S_0 energy gap as a function of time. The averaged S_1-S_0 energy gap drops below 0.5 eV in 25 fs and keeps fluctuating around 0.25 eV in the rest of the simulation. The 25 fs time scale is consistent with the time scale when the averaged excited state population starts decaying. As indicated in Figure 14-4 that individual trajectories were hitting the zero S_1-S_0 energy gap, which suggests that conical intersection may play a role in the population transfer. The representative snapshots from the dynamical simulations

show that one of the NH_2 ligand has a broken Ta-N bond at 25 fs and dissociates at longer simulation time. This also suggests that the conical intersection geometry may be reached by Ta-N bond breaking process.

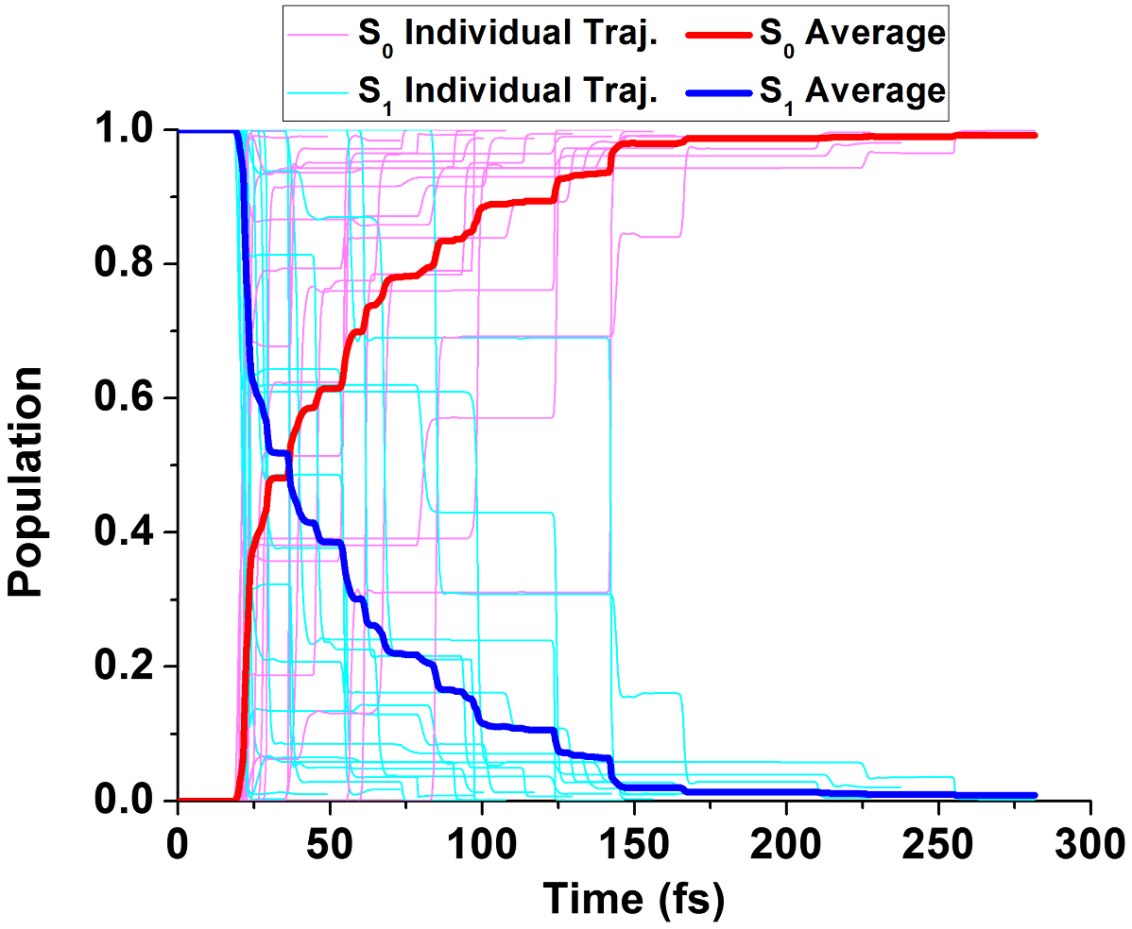


Figure 14-3. Population of S_1 (bold blue) and S_0 (bold red) as a function of time after excitation averaged over all 20 initial nuclear basis functions. Thinner blue and pink lines reflect the population corresponding to the 20 individual basis functions before averaging.

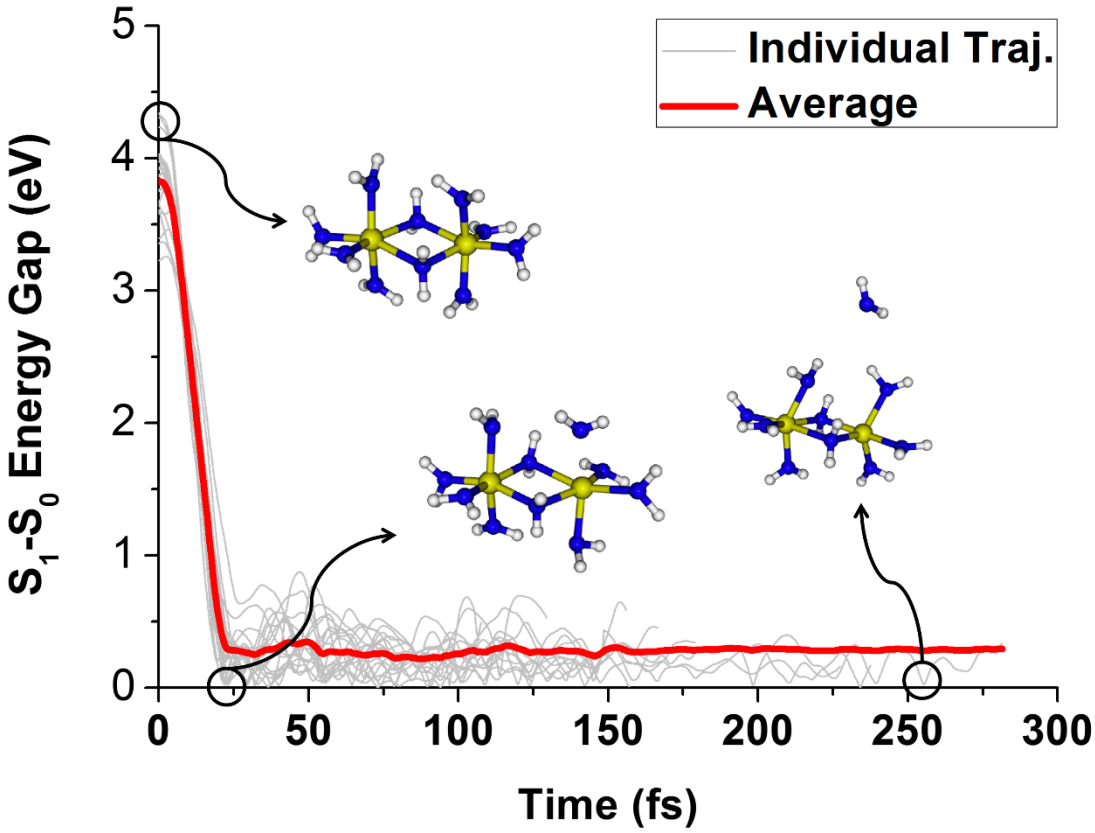


Figure 14-4. Average S_1-S_0 energy gap of the model Ta_3N_5 cluster as a function of time (bold red) with the energy gaps of 20 initial nuclear basis functions (gray). Snapshots from representative trajectories are shown in the inset, with tantalum, nitrogen and hydrogen atoms colored yellow, blue and grey, respectively.

The decay of the excited state population occurs after Ta-N bond breaks. Figure 14-5 shows that one of the Ta-N bond lengths has significantly lengthened during the simulation. It is clear from the dynamical simulation snapshot that only one of the Ta-N bond breaks from Figure 14-4 and this is also a general observation for all trajectories. This is a direct evidence that relaxation on the excited state only involves changing of local

electronic structure and nuclear motion, despite the fact that there is no defect in the cluster. This is our second time observing this phenomena, previous studies related to pristine silicon nanocrystal capped with hydrogen atoms also support this conclusion. The localized nuclear deformation upon relaxation on excited state is important in the validation of our model cluster. The AIMS simulation shows that the averaged Ta-N bond length increases from 2 Å to 3.7 Å in 150 fs with fast increment speed around 25 fs. This time scale corresponds to that of both S_1-S_0 energy gaps drop below 0.5 eV and the beginning of the excited state population decay. The Ta-N bond breaking or lengthening is a typical nuclear motion that happened in all the individual trajectories. It is also suggested from the individual trajectory in Figure 14-5 and snapshots from dynamics that surface NH_2 ligand is releasing upon excitation of the material. This indirectly supports the experimental observation that at the beginning of the water splitting reaction, low level of N_2 evolution was observed.[3]

The Ta-N bond rotation is also observed from the dynamical simulation. Figure 14-6(a) shows two N-TaNN pyramidalization angles as a function of time with the insets show the associated atoms for each pyramidalization angle. The N-TaNN pyramidalization angle is defined as

$$\theta = \left| \arcsin \frac{\vec{C} \bullet \vec{R}_{TaN(1or2)}}{\left\| \vec{C} \right\| \left\| \vec{R}_{TaN(1or2)} \right\|} \right| \qquad (14\text{-}3)$$

where $\vec{C} = \vec{R}_{TaN(3)} \times \vec{R}_{TaN(4)}$, $\vec{R}_{AB}$ is a vector from atom A to atom B, with $\bullet$ and $\times$ indicating the scalar and vector products of the two vectors respectively. Nitrogen atoms

with label N(1), N(2), N(3) and N(4) are defined in Figure 14-6(b). Both N-TaNN angles initiate at 10 degrees because the ground state minimum geometry is twisted from perfect octahedral coordination. The averaged N(1)-TaN(3)N(4) pyramidalization angle changes from -10.6 degree to +13.9 degree in 150 fs. This is possibly due to the reduction of steric effect as one of the neighboring Ta-N bond breaks during the simulation. This N-TaNN pyramidalization motion corresponds to 323.1cm^{-1} frequency computed at B3LYP/LAN2LDZ level of theory and can be treated as a signature in the transient spectroscopy.

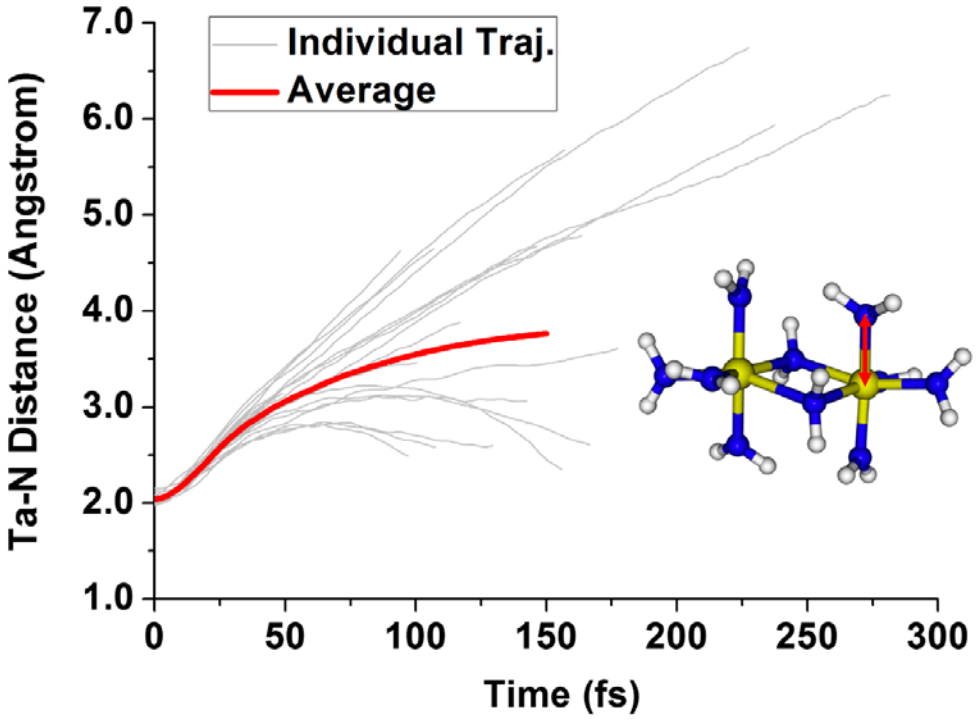


Figure 14-5. Averaged Ta-N bond length as a function of time (bold red) with bond length of 20 initial nuclear basis functions (gray).

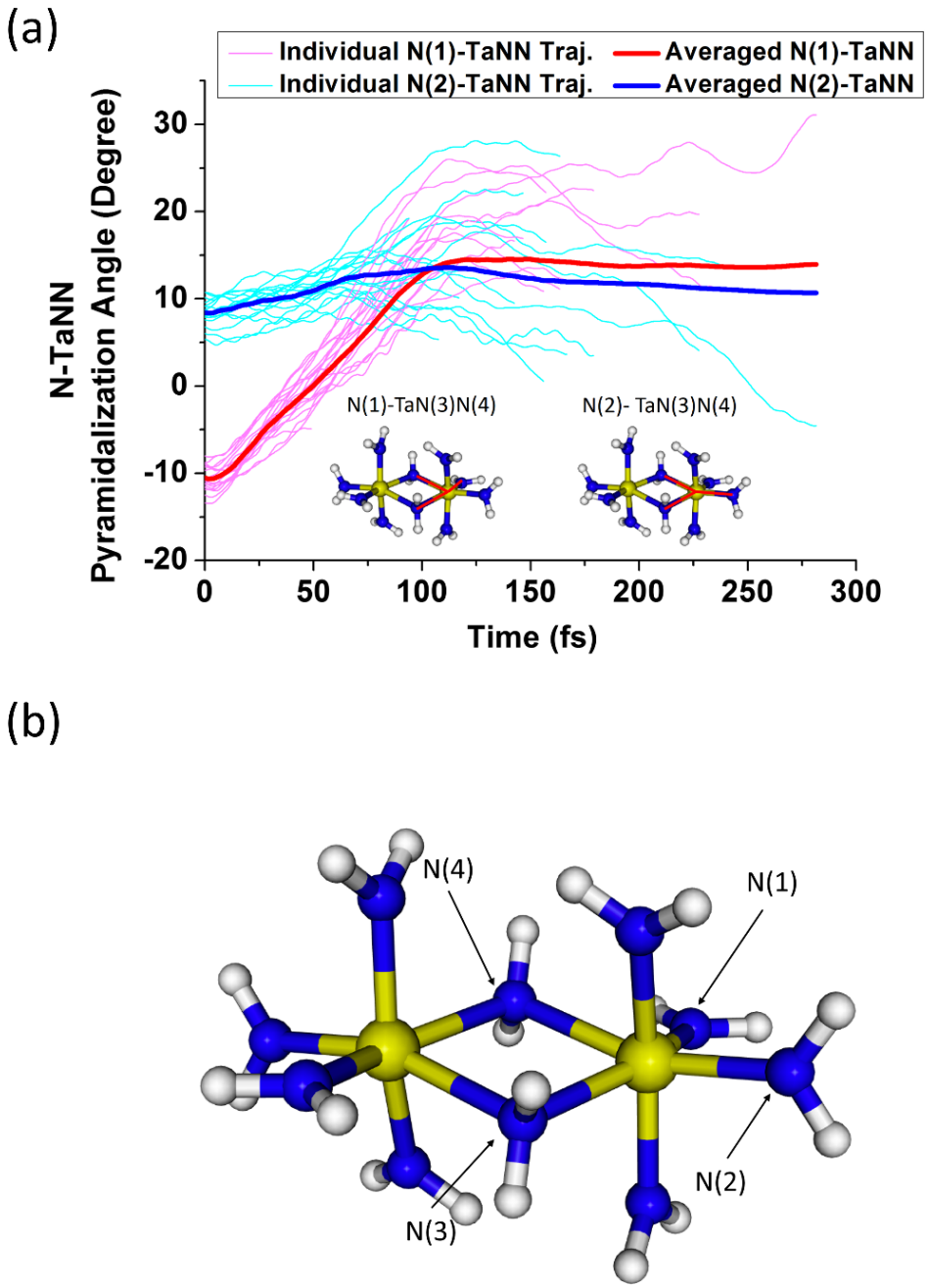


Figure 14-6. (a) Averaged two different N-TaNN pyramidalization angles as a function of time (bold red and blue) with bond length of 20 initial nuclear basis functions (thin red and blue). The corresponding pyramidalization angles are shown in the insets. (b) The atoms defined in the definition of pyramidalization angle.

14.2.3 Potential energy surface analysis

As has been pointed out in the previous section that conical intersection may play a role in the non-radiative recombination process due to the observation of AIMS single trajectories' S_1-S_0 energy gaps hitting zero. To validate our hypothesis, we applied SA-CASSCF and MS-CASPT2 theoretical levels to explore the S_1 state PES. We optimize both the excited state minima and S_1-S_0 MECIs at both theoretical levels. We found the MECIs with similar structures that have been observed in AIMS simulation at the time scale close to population decay. However, despite various effort, we were not being able to locate a stable excited state minimum. This is possible when MECI is a peaked MECI, which is consistent with the fact that efficient population transfer was observed (99% S_1 state population decays to ground state in 150 fs). At the MECI geometries optimized at both SA-CASSCF and MS-CASPT2 theoretical levels, single point calculations at the SA-CASSCF, MS-CASPT2 and MRCI(Q) levels are performed. Note that we employ non-polarized basis set for SA-CASSCF calculations and polarized basis set for MS-CASPT2 and MRCI(Q) calculations. The results are shown in Table 14-1 and Figure 14-7.

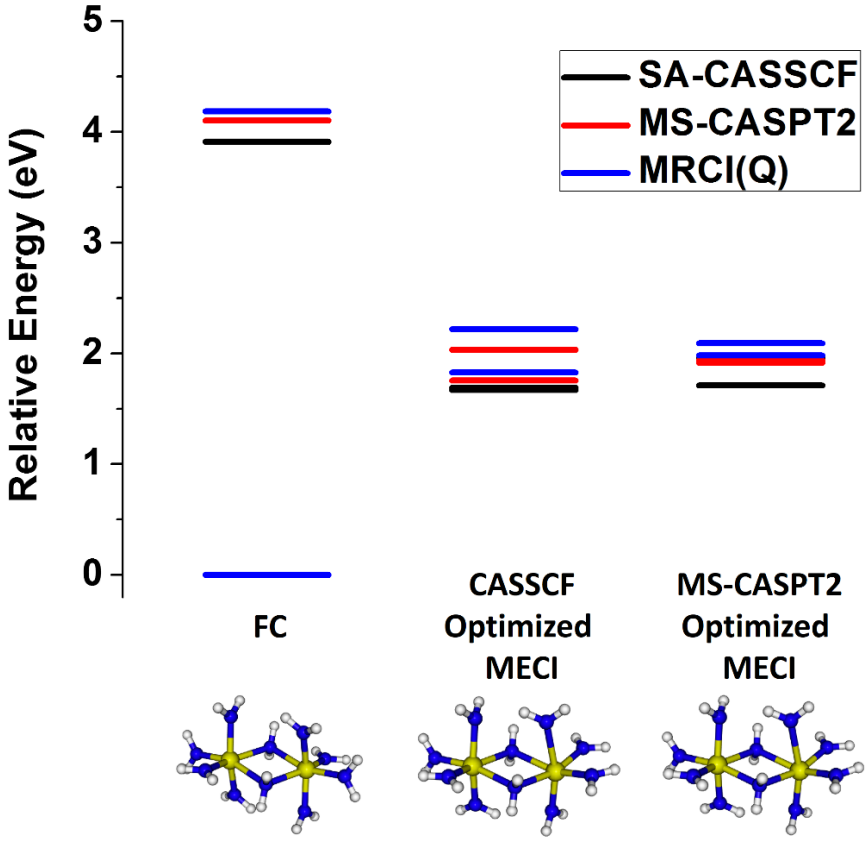


Figure 14-7. The S_0 and S_1 energies of three important points on the PES of the tantalum nitride model cluster: the Franck-Condon (FC) point, S_0/S_1 MECI and S_1 min optimized on SA-CASSCF theoretical level. Black, red and blue lines correspond to the energies of the single point calculation computed at the SA-CASSCF, MS-CASPT2 and MRCISD(Q) levels of theory, respectively. Energies are shifted so that the S_0 energy of the FC point is zero for all levels of theory. Optimized geometries are shown as insets.

Table 14-1. The S_0 and S_1 energies (in eV) at two important points on the PES of the model tantalum nitride cluster calculated at the CASSCF, MS-CASPT2, and MRCI(Q) levels of theory. S_1-S_0 MECI geometries are optimized at both the CASSCF and MS-CASPT2 levels of theory.

Geometry	Optimization Method	SA-CASSCF S_0 Energy	SA-CASSCF S_1 Energy	MS-CASPT2 S_0 Energy	MS-CASPT2 S_1 Energy	MRCI(Q) S_0 Energy	MRCI(Q) S_1 Energy
FC geometry	KS/B3LYP	0	3.91	0	4.10	0	4.19
S_0-S_1 MECI	SA-CASSCF	1.67	1.69	1.76	2.03	1.83	2.22
	MS-CASPT2	1.71	1.96	1.92	1.94	1.98	2.09

The SA-CASSCF optimized MECI is found to be 1.69, 2.03 and 2.22 eV above the ground state minimum at the SA-CASSCF, MS-CASPT2 and MRCI(Q) theoretical levels respectively. It is noteworthy that these computed MECI energy level is in the visible range and close or under the 2.1 eV band gap of bulk Ta_3N_5 material. Combined with our hypothesis that MECI is insensitive to the size of the cluster,[94] it is reasonable to assume that the MECI is energetically accessible by the excitons in the bulk material. Hence efficient non-radiative recombination process is expected. This again, suggests why a sacrificial reagent is required. The existing of the sacrificial reagent removes either the excited electron or the hole after excitation, thus prevents the non-radiative

recombination. The computed SA-CASSCF S_1 energy at SA-CASSCF optimized MECI has 0.34 and 0.53 eV error as compared to MS-CASPT2 and MRCI(Q) energies. This error is common due to the lack of dynamical correlation at the SA-CASSCF level. MS-CASPT2 and MRCI(Q) predicts small S_1-S_0 energy gaps at SA-CASSCF optimized MECI, suggesting that SA-CASSCF predicts qualitatively correct MECI nuclear geometry. These results validate our PES employed in dynamical simulations.

To further assess the accuracy of the SA-CASSCF PES, we re-optimized the MECI geometry at MS-CASPT2 theoretical level with polarized basis set. The MS-CASPT2 optimized MECI is shown in Figure 14-7 as inset. MS-CASPT2 predicts 1.94 eV S_1 excitation energy at MS-CASPT2 optimized MECI which is close to 1.69 eV predicted at SA-CASSCF theoretical level. Single point calculations at SA-CASSCF and MRCI(Q) theoretical levels predict 1.96 and 2.09 eV S_1 excitation energies at MS-CASPT2 optimized MECI geometry. The small discrepancies among various theoretical levels again indicate the PES employed in AIMS simulation is qualitatively accurate.

The insets in Figure 14-7 shows the geometry of our model cluster at FC and MECI geometries. It is not clear from the figure that Ta-N bond is broken or lengthened at MECI geometries. However, we do find one relatively long Ta-N bond with 2.46 and 2.51 Å bond length at MECI geometries optimized with SA-CASSCF and MS-CASPT2 respectively (as compare with the 2.02 to 2.05 Å optimal Ta-N bond lengths at ground state minimum geometry). This bond length is close to the averaged Ta-N bond length has been observed in AIMS dynamical simulations at 25 fs which is about 2.5 Å as shown in Figure

14-5. Note that 25 fs corresponds to the first time that averaged S_1-S_0 energy gaps for individual trajectories hitting zero.

Partially occupied natural orbitals (PONOs) are shown in Figure 14-8 with two different perspectives at FC and SA-CASSCF optimized MECI geometries. The PONOs at FC geometry is computed by diagonalizing of S_1 density matrix at SA-CASSCF theoretical level. While the PONOs at MECI geometry is computed by the diagonalization of the averaged S_1-S_0 density matrix. At FC geometry, the PONOs are mixture of N p orbital and Ta d orbital (Figure 14-8 left) with a weak anti-bonding interaction. The relatively degenerate occupation numbers indicates the exact mixture of these orbitals meaningless. In fact, the excitation is characterized by N p orbital to Ta d orbital as shown in Figure 14-2(b). The existing of anti-bonding interaction between N p orbital and Ta d orbital explains the NH_2 releasing dynamical behavior observed in AIMS simulation. The PONOs at MECI geometry indicates that the population transfer facilitated by the MECI is accompanied with the electron from the Ta d orbital to the N p and Ta d bonding orbital.

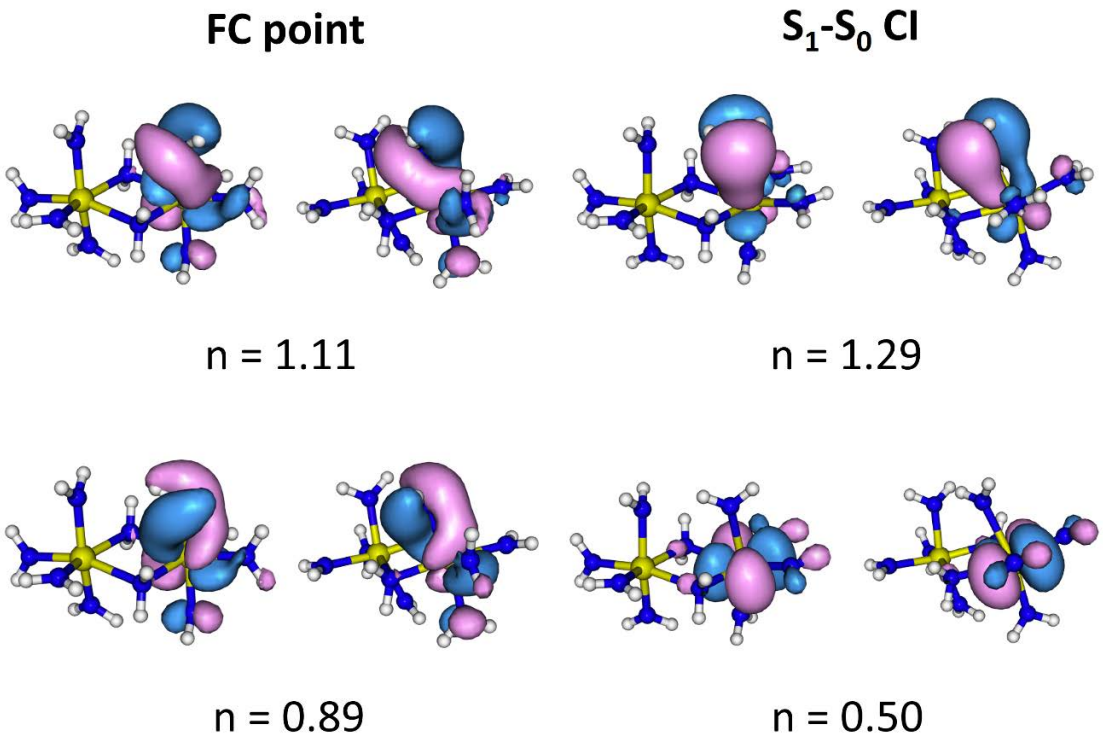


Figure 14-8. Partially occupied SA-CASSCF natural orbitals of silicon hydroxide cluster at FC (left) and MECI (right) geometries are presented with occupation numbers. The FC natural orbitals are computed by diagonalization of the S_1 density matrix, while S_1–S_0 MECI natural orbitals are computed from the state-averaged S_1–S_0 density matrix.

14.2.4 Spin-orbit coupling effect

To investigate the spin-orbit coupling effect especially for heavy transition metal Ta, we applied single point calculations at both FC and SA-CASSCF optimized MECI geometries. Single point calculations are performed at SA-CASSCF theoretical level with double-zeta basis set and Stuttgart-Cologne pseudopotentials.[95] ECP60MDF psudopotential for Ta[95]

were empolyed to account for the non-scalar relativistic effect. The same active space and number of states averaged in SA-CASSCF were employed as that of AIMS simulation. At the FC geometry, the excitation energy of spin-orbit coupling corrected S_1 state is 4.00 eV (compared to 3.91 eV predicted by SA2-CASSCF(4/6) calculation without non-scalar relativistic effect). At the MECI geometry, spin-orbit coupling effect shift the S_1 state to 1.91 eV with 0.18 eV S_1-S_0 energy gap. This suggests our SA-CASSCF PES is qualitatively good although only scalar relativistic effect is included.

14.3 Efficient Ta_3N_5 Based Water Splitting Photocatalyst Design

Our simulation has illustrated the non-radiative recombination mechanism. Combined with our previous work,[94,96-98] it is reasonable to believe cluster size effect is weak in determining the MECI energies. Hence we summarize the following requirements for efficient Ta_3N_5 based photocatalyst according to the previous study and current theoretical investigation: 1) The band gap of photocatalyst should be in the visible range for efficient light absorption, 2) The band gap should be large enough to overcome the oxidation reduction potential of water, 3) The conical intersection energy level should be above the band gap of the bulk material absorption energy to block the efficient non-radiative recombination pathway and hence a sacrificial reagent is not required.

The reduction of the band gap of bulk tantalum nitride compared to tantalum oxide is due to the rise of the potential of the p orbitals from oxygen to nitrogen.[3] It is hence promising to utilize computational studies to screening other ligands to obtain possible

candidates for water splitting photocatalyst that fulfills the above three requirements.

14.4 Conclusions

In this work, we have applied AIMS to the model cluster to investigate the excited state dynamics of the Ta_3N_5 photocatalyst. We demonstrated that conical intersection plays a role in non-radiative recombination process. The conical intersection is accessible by the exciton in bulk material. MS-CASPT2 predicts 1.94 eV excitation energy of MECI above the ground state minimum, while the band gap of bulk Ta_3N_5 is around 2.1 eV. This explains the necessity of sacrificial reagent in order to prevent the non-radiative decay through conical intersection and photosplit water. AIMS dynamics show that the model cluster has efficient population transfer upon excitation, this is accompanied by NH_2 ligand releasing process. The non-radiative recombination mechanism provided by this work illustrates the importance of MECI energy level in the charge separation step. Efficient photocatalyst requires high energy level MECI above ground state minimum to prevent charge recombination.

APPENDIX

APPENDIX

Table 14-S1. The absolute energies of the various optimized geometries of the Ta_3N_5 cluster model at the SA-CASSCF level of theory (with active space and basis sets as described in the text).

CASSCF	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point	-670.9160627	-670.7723372
SA-CASSCF optimized S_0/S_1 MECI	-670.8546370	-670.8539081
MS-CASPT2 optimized S_0/S_1 MECI	-670.8532352	-670.8439781

Table 14-S2. The absolute energies of the various optimized geometries of the Ta_3N_5 cluster model at the MS-CASPT2 level of theory (with active space and basis sets as described in the text).

MS-CASPT2	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point	-671.1043376	-670.9535653
SA-CASSCF optimized S_0/S_1 MECI	-671.0398402	-671.0296588
MS-CASPT2 optimized S_0/S_1 MECI	-671.0338840	-671.0331291

Table 14-S3. The absolute energies of the various optimized geometries of the Ta_3N_5 cluster model at the MRCI(Q) level of theory (with active space and basis sets as described in the text).

MRCI(Q)	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point	-671.1457107	-670.9918698
SA-CASSCF optimized S_0/S_1 MECI	-671.0785710	-671.0641625
MS-CASPT2 optimized S_0/S_1 MECI	-671.0729281	-671.0688888

Table 14-S4. The absolute energies of the FC geometry of the Ta_3N_5 cluster model at the EOM-CCSD level of theory.

EOM-CCSD	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point	-673.1387129	-672.9762072

Optimized Geometries (In Angstrom)

FC cluster geometry

N	0.145861685	0.384279231	2.021534623
Ta	-0.296823050	0.262337888	0.054258577
N	-0.522296625	-1.689801291	-0.518724490
N	1.882811937	-0.318633837	-0.030878317
N	-2.277217851	0.123263800	0.325120661
N	0.255339607	0.541464912	-2.086404144
N	-0.428103689	2.282928822	-0.095996097
H	-1.409801431	-2.122176848	-0.772792971
H	0.259334952	-2.267603982	-0.833772421
H	-0.280817975	-0.136053300	-2.637127676
H	0.125870799	1.474310810	-2.48676364
H	-2.789469953	-0.743249491	0.487421144
H	-2.884689810	0.938815547	0.373512902
H	1.066628774	0.322821262	2.450518611
H	-0.588170682	0.440253979	2.726436122
H	-0.669978386	2.850816764	-0.908591960
H	-0.387359462	2.856582294	0.747338348
H	1.996443721	-1.260904640	0.352541627
H	2.431393509	0.338267533	0.532641691
Ta	2.441807389	-0.009787846	-2.166224918
N	2.653826186	1.935004441	-1.564480368
N	2.559957408	-2.033103378	-2.051142749
N	2.019524795	-0.092666419	-4.140092918
N	4.424638161	0.133229479	-2.418222412
H	2.521916499	-2.590878536	-2.905373017
H	2.786331354	-2.619225623	-1.246571941
H	4.941726249	1.000499010	-2.560362035
H	5.030550044	-0.683032821	-2.476367557
H	1.100592107	-0.034980968	-4.573662038
H	2.759609761	-0.118810508	-4.840319569
H	1.866639152	2.502483394	-1.243461632
H	3.538111824	2.375354320	-1.313277438

SA-CASSCF Optimized S_0/S_1 MECI

N	0.187114300	0.561695050	2.107136960
Ta	-0.246915237	0.333301567	0.138865438
N	-0.473773833	-1.666538185	-0.217293933
N	1.955517881	-0.137400937	0.012120725
N	-2.246342080	0.330303984	0.376765076
N	0.173964224	0.459161410	-2.058036678
N	-0.252393332	2.395017727	-0.122129280
H	-1.339934953	-2.159212833	-0.163426956
H	0.154926394	-2.168921022	-0.811958632
H	-0.332964313	-0.277771264	-2.513078237
H	-0.091318933	1.305559172	-2.524504143
H	-2.836821262	-0.469905623	0.453936641
H	-2.771984176	1.166432178	0.518306783
H	1.082607942	0.472094444	2.536759705
H	-0.525331266	0.660512074	2.799841472
H	-0.683284334	2.873781959	-0.887587153
H	-0.308620305	2.966942637	0.696391733
H	2.101593077	-1.049066409	0.408408479
H	2.535103742	0.476834018	0.557425395
Ta	2.426780982	-0.053257753	-2.256774718
N	2.351971311	2.333044637	-1.655872214
N	2.144846254	-2.064467517	-2.403827318
N	2.525173325	0.465739610	-4.257276007
N	4.438494699	0.310654334	-1.993443221
H	2.142481139	-2.545238634	-3.279883548
H	2.362064096	-2.715171486	-1.675721018
H	4.959566200	0.051166490	-1.179935773
H	5.074115095	0.487709983	-2.745068886
H	1.738743551	0.634328886	-4.852011066
H	3.359216279	0.537592999	-4.803392575
H	1.662004481	2.767540197	-1.054418719
H	3.106416343	2.952409545	-1.895594159

MS-CASPT2 Optimized S_0/S_1 MECI

N	0.178632816	0.520327390	2.072258029
Ta	-0.223941617	0.333489535	0.142409835
N	-0.470496389	-1.646879713	-0.209568606
N	1.956043596	-0.124352746	0.001173160
N	-2.239689308	0.343351843	0.355109493
N	0.179553932	0.450556023	-2.047921931
N	-0.245383348	2.377752268	-0.114436106
H	-1.335861813	-2.141048182	-0.155366865
H	0.163350097	-2.151905503	-0.797838060
H	-0.340671317	-0.277159656	-2.502030820
H	-0.090954661	1.298047106	-2.510180810
H	-2.823397693	-0.460323761	0.459487444
H	-2.766128803	1.172592359	0.532239040
H	1.073640309	0.472726071	2.508414535
H	-0.530946267	0.667743454	2.759580589
H	-0.666464873	2.849844490	-0.889790471
H	-0.315430778	2.953390545	0.700567631
H	2.097257390	-1.035436737	0.401018598
H	2.537467411	0.482701167	0.552374450
Ta	2.417318092	-0.087735794	-2.260669405
N	2.360973682	2.348109620	-1.646961144
N	2.129488738	-2.073318742	-2.415730250
N	2.529920587	0.434617375	-4.229835383
N	4.418975380	0.264515040	-2.009193147
H	2.137746967	-2.551260483	-3.292787406
H	2.376667915	-2.714530408	-1.688696625
H	4.942312495	0.069468837	-1.179414184
H	5.060837460	0.510269833	-2.737728353
H	1.744231917	0.629020627	-4.819438968
H	3.364714095	0.550527343	-4.770876317
H	1.664699926	2.777268990	-1.044920892
H	3.088551357	2.992503061	-1.906522713

REFERENCES

REFERENCES

(1) Walter, M. G.; Warren, E. L.; McKone, J. R.; Boettcher, S. W.; Mi, Q. X.; Santori, E. A.; Lewis, N. S. *Chem Rev* **2010**, *110*, 6446.

(2) Kudo, A.; Miseki, Y. *Chem Soc Rev* **2009**, *38*, 253.

(3) Maeda, K.; Domen, K. *J Phys Chem C* **2007**, *111*, 7851.

(4) Fujishima, A.; Honda, K. *Nature* **1972**, *238*, 37.

(5) Li, Z.; Luo, W.; Zhang, M.; Feng, J.; Zou, Z. *Energ Environ Sci* **2013**, *6*, 347.

(6) Hisatomi, T.; Kubota, J.; Domen, K. *Chem Soc Rev* **2014**, *43*, 7520.

(7) Moniz, S. J.; Shevlin, S. A.; Martin, D. J.; Guo, Z.-X.; Tang, J. Energ Environ Sci 2015, 8, 731.

(8) Yoshimura, J.; Ebina, Y.; Kondo, J.; Domen, K.; Tanaka, A. J Phys Chem-Us 1993, 97, 1970.

(9) Kudo, A.; Mikami, I. *Chem Lett* **1998**, 1027.

(10) Hwang, D. W.; Kim, H. G.; Kim, J.; Cha, K. Y.; Kim, Y. G.; Lee, J. S. *J Catal* **2000**, *193*, 40.

(11) Sayama, K.; Mukasa, K.; Abe, R.; Abe, Y.; Arakawa, H. *Chem Commun* **2001**, 2416.

(12) Zou, Z. G.; Ye, J. H.; Sayama, K.; Arakawa, H. *Nature* **2001**, *414*, 625.

(13) Kato, H.; Kudo, A. *J Phys Chem B* **2002**, *106*, 5029.

(14) Kim, H. G.; Hwang, D. W.; Lee, J. S. *J Am Chem Soc* **2004**, *126*, 8912.

(15) Hwang, D. W.; Kirn, H. G.; Lee, J. S.; Kim, J.; Li, W.; Oh, S. H. J Phys Chem B 2005, 109, 2093.

(16) Kim, H. G.; Borse, P. H.; Choi, W. Y.; Lee, J. S. *Angew Chem Int Edit* **2005**, *44*, 4585.

(17) Abe, R.; Sayama, K.; Sugihara, H. *J Phys Chem B* **2005**, *109*, 16052.

(18) Maeda, K.; Takata, T.; Hara, M.; Saito, N.; Inoue, Y.; Kobayashi, H.; Domen, K. J Am Chem Soc 2005, 127, 8286.

(19) Liu, Z. W.; Hou, W. B.; Pavaskar, P.; Aykol, M.; Cronin, S. B. *Nano Lett* **2011**, *11*, 1111.

(20) Primo, A.; Marino, T.; Corma, A.; Molinari, R.; Garcia, H. J Am Chem Soc 2011, 133, 6930.

(21) Higashi, M.; Domen, K.; Abe, R. *J Am Chem Soc* **2012**, *134*, 6968.

(22) Ohno, T.; Bai, L.; Hisatomi, T.; Maeda, K.; Domen, K. *J Am Chem Soc* **2012**, *134*, 8254.

(23) Osterloh, F. E. *Chem Soc Rev* **2013**, *42*, 2294.

(24) Liao, L. B.; Zhang, Q. H.; Su, Z. H.; Zhao, Z. Z.; Wang, Y. N.; Li, Y.; Lu, X. X.; Wei, D. G.; Feng, G. Y.; Yu, Q. K.; Cai, X. J.; Zhao, J. M.; Ren, Z. F.; Fang, H.; Robles-Hernandez, F.; Baldelli, S.; Bao, J. M. Nat Nanotechnol 2014, 9, 69.

(25) Yeh, T. F.; Teng, C. Y.; Chen, S. J.; Teng, H. S. *Adv Mater* **2014**, *26*, 3297.

(26) Martin, D. J.; Reardon, P. J. T.; Moniz, S. J. A.; Tang, J. W. J Am Chem Soc 2014, 136, 12568.

(27) Liu, J.; Liu, Y.; Liu, N. Y.; Han, Y. Z.; Zhang, X.; Huang, H.; Lifshitz, Y.; Lee, S. T.; Zhong, J.; Kang, Z. H. Science 2015, 347, 970.

(28) Xu, J. S.; Pan, C. S.; Takata, T.; Domen, K. *Chem Commun* **2015**, *51*, 7191.

(29) Pan, C. S.; Takata, T.; Nakabayashi, M.; Matsumoto, T.; Shibata, N.; Ikuhara, Y.; Domen, K. Angew Chem Int Edit 2015, 54, 2955.

(30) Han, Q.; Hu, C. G.; Zhao, F.; Zhang, Z. P.; Chen, N.; Qu, L. T. J Mater Chem A 2015, 3, 4612.

(31) Kato, H.; Kudo, A. *The Journal of Physical Chemistry B* **2001**, *105*, 4285.

(32) Shimizu, K.-i.; Tsuji, Y.; Kawakami, M.; Toda, K.; Kodama, T.; Sato, M.; Kitayama, Y. Chem Lett 2002, 1158.

(33) Ikeda, S.; Itani, T.; Nango, K.; Matsumura, M. *Catalysis letters* **2004**, *98*, 229.

(34) Sato, J.; Saito, N.; Nishiyama, H.; Inoue, Y. The Journal of Physical Chemistry B 2001, 105, 6061.

(35) Ikarashi, K.; Sato, J.; Kobayashi, H.; Saito, N.; Nishiyama, H.; Inoue, Y. The Journal of Physical Chemistry B 2002, 106, 9048.

(36) Sato, J.; Saito, N.; Nishiyama, H.; Inoue, Y. Journal of Photochemistry and Photobiology A: Chemistry 2002, 148, 85.

(37) Sato, J.; Kobayashi, H.; Saito, N.; Nishiyama, H.; Inoue, Y. Journal of Photochemistry and Photobiology A: Chemistry 2003, 158, 139.

(38) Sato, J.; Saito, N.; Nishiyama, H.; Inoue, Y. The Journal of Physical Chemistry B 2003, 107, 7965.

(39) Sato, J.; Kobayashi, H.; Ikarashi, K.; Saito, N.; Nishiyama, H.; Inoue, Y. The Journal of Physical Chemistry B 2004, 108, 4369.

(40) Morbec, J. M.; Narkeviciute, I.; Jaramillo, T. F.; Galli, G. *Phys Rev B* **2014**, *90*, 155204.

(41) Hitoki, G.; Ishikawa, A.; Takata, T.; Kondo, J. N.; Hara, M.; Domen, K. Chemistry Letters 2002, 736.

(42) Zhang, Q. H.; Gao, L. *Langmuir* **2004**, *20*, 9821.

(43) Li, Y. B.; Takata, T.; Cha, D.; Takanabe, K.; Minegishi, T.; Kubota, J.; Domen, K. Advanced Materials 2013, 25, 125.

(44) Ma, S. S. K.; Hisatomi, T.; Maeda, K.; Moriya, Y.; Domen, K. Journal of the American Chemical Society 2012, 134, 19993.

(45) Cong, Y. Q.; Park, H. S.; Wang, S. J.; Dang, H. X.; Fan, F. R. F.; Mullins, C. B.; Bard, A. J. Journal of Physical Chemistry C 2012, 116, 14541.

(46) Dabirian, A.; van de Krol, R. *Applied Physics Letters* **2013**, *102*.

(47) Ma, S. S. K.; Maeda, K.; Hisatomi, T.; Tabata, M.; Kudo, A.; Domen, K. Chemistry-a European Journal 2013, 19, 7480.

(48) Dang, H. X.; Hahn, N. T.; Park, H. S.; Bard, A. J.; Mullins, C. B. Journal of Physical Chemistry C 2012, 116, 19225.

(49) Nurlaela, E.; Ould-Chikh, S.; Harb, M.; del Gobbo, S.; Aouine, M.; Puzenat, E.; Sautet, P.; Domen, K.; Basset, J. M.; Takanabe, K. Chemistry of Materials 2014, 26, 4812.

(50) Hara, M.; Hitoki, G.; Takata, T.; Kondo, J. N.; Kobayashi, H.; Domen, K. Catalysis Today 2003, 78, 555.

(51) Liu, G.; Shi, J.; Zhang, F.; Chen, Z.; Han, J.; Ding, C.; Chen, S.; Wang, Z.; Han, H.; Li, C. Angewandte Chemie International Edition 2014, 53, 7295.

(52) Yuliati, L.; Yang, J.-H.; Wang, X.; Maeda, K.; Takata, T.; Antonietti, M.; Domen, K. Journal of Materials Chemistry 2010, 20, 4295.

(53) Cong, Y.; Park, H. S.; Dang, H. X.; Fan, F.-R. F.; Bard, A. J.; Mullins, C. B. Chemistry of Materials 2012, 24, 579.

(54) Ho, C. T.; Low, K. B.; Klie, R. F.; Maeda, K.; Domen, K.; Meyer, R. J.; Snee, P. T. Journal of Physical Chemistry C 2011, 115, 647.

(55) Ishikawa, A.; Takata, T.; Kondo, J. N.; Hara, M.; Domen, K. Journal of Physical Chemistry B 2004, 108, 11049.

(56) Tabata, M.; Maeda, K.; Higashi, M.; Lu, D.; Takata, T.; Abe, R.; Domen, K. Langmuir 2010, 26, 9161.

(57) Li, Y.; Zhang, L.; Torres-Pardo, A.; González-Calbet, J. M.; Ma, Y.; Oleynikov, P.; Terasaki, O.; Asahina, S.; Shima, M.; Cha, D. Nat Commun 2013, 4.

(58) Feng, X.; LaTempa, T. J.; Basham, J. I.; Mor, G. K.; Varghese, O. K.; Grimes, C. A. Nano Lett 2010, 10, 948.

(59) Chun, W. J.; Ishikawa, A.; Fujisawa, H.; Takata, T.; Kondo, J. N.; Hara, M.; Kawai, M.; Matsumoto, Y.; Domen, K. Journal of Physical Chemistry B 2003, 107, 1798.

(60) Hitoki, G.; Ishikawa, A.; Takata, T.; Kondo, J. N.; Hara, M.; Domen, K. Chem Lett 2002, 736.

(61) Yokoyama, D.; Hashiguchi, H.; Maeda, K.; Minegishi, T.; Takata, T.; Abe, R.; Kubota, J.; Domen, K. Thin Solid Films 2011, 519, 2087.

(62) Higashi, M.; Domen, K.; Abe, R. *Energy & Environmental Science* **2011**, *4*, 4138.

(63) Li, Y.; Takata, T.; Cha, D.; Takanabe, K.; Minegishi, T.; Kubota, J.; Domen, K. *Advanced Materials* **2013**, *25*, 125.

(64) Hara, M.; Chiba, E.; Ishikawa, A.; Takata, T.; Kondo, J. N.; Domen, K. The Journal of Physical Chemistry B 2003, 107, 13441.

(65) Chun, W.-J.; Ishikawa, A.; Fujisawa, H.; Takata, T.; Kondo, J. N.; Hara, M.; Kawai, M.; Matsumoto, Y.; Domen, K. The Journal of Physical Chemistry B 2003, 107, 1798.

(66) Fang, C. M.; Orhan, E.; de Wijs, G. A.; Hintzen, H. T.; de Groot, R. A.; Marchand, R.; Saillard, J. Y.; de With, G. J Mater Chem 2001, 11, 1248.

(67) Stampfl, C.; Freeman, A. J. *Phys Rev B* **2005**, *71*.

(68) Wu, Y.; Ceder, G. *The Journal of Physical Chemistry C* **2013**, *117*, 24710.

(69) Wang, J.; Luo, W.; Feng, J.; Zhang, L.; Li, Z.; Zou, Z. *Phys Chem Chem Phys* **2013**, *15*, 16054.

(70) Wang, J.; Fang, T.; Zhang, L.; Feng, J.; Li, Z.; Zou, Z. *Journal of Catalysis* **2014**, *309*, 291.

(71) Reshak, A. *Comp Mater Sci* **2014**, *89*, 45.

(72) Wang, J.; Feng, J.; Zhang, L.; Li, Z.; Zou, Z. *Phys Chem Chem Phys* **2014**, *16*, 15375.

(73) Ben-Nun, M.; Quenneville, J.; Martinez, T. J. Journal of Physical Chemistry A 2000, 104, 5161.

(74) Levine, B. G.; Coe, J. D.; Virshup, A. M.; Martinez, T. J. *Chem Phys* **2008**, *347*, 3.

(75) Heller, E. J. *J Chem Phys* **1981**, *75*, 2923.

(76) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. *Chem Phys* **1980**, *48*, 157.

(77) Wadt, W. R.; Hay, P. J. *J Chem Phys* **1985**, *82*, 284.

(78) Stanton, J. F.; Bartlett, R. J. *The Journal of chemical physics* **1993**, *98*, 7029.

(79) Becke, A. D. *Journal of Chemical Physics* **1993**, *98*, 1372.

(80) Lee, C. T.; Yang, W. T.; Parr, R. G. *Physical Review B* **1988**, *37*, 785.

(81) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Journal of Physical Chemistry 1994, 98, 11623.

(82) Thompson, A. L.; Punwong, C.; Martinez, T. J. *Chem Phys* **2010**, *370*, 70.

(83) Finley, J.; Malmqvist, P. A.; Roos, B. O.; Serrano-Andres, L. Chem Phys Lett 1998, 288, 299.

(84) Langhoff, S. R.; Davidson, E. R. *Int J Quantum Chem* **1974**, *8*, 61.

(85) Knowles, P. J.; Werner, H. J. *Theor Chim Acta* **1992**, *84*, 95.

(86) Shiozaki, T.; Gyorffy, W.; Celani, P.; Werner, H. J. *J Chem Phys* **2011**, *135*.

(87) Hampel, C.; Peterson, K. A.; Werner, H. J. *Chemical Physics Letters* **1992**, *190*, 1.

(88) Werner, H. J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schutz, M. Wiley Interdisciplinary Reviews-Computational Molecular Science 2012, 2, 242.

(89) Celani, P.; Werner, H. J. *J Chem Phys* **2000**, *112*, 5546.

(90) Korona, T.; Werner, H. J. *J Chem Phys* **2003**, *118*, 3006.

(91) Werner, H. J.; Knowles, P. J. *The Journal of chemical physics* **1985**, *82*, 5053.

(92) Levine, B. G.; Coe, J. D.; Martinez, T. J. *J Phys Chem B* **2008**, *112*, 405.

(93) Busch, T.; Degli Esposti, A.; Werner, H. J. *J Chem Phys* **1991**, *94*, 6708.

(94) Shu, Y. N.; Fales, B. S.; Levine, B. G. *Nano Letters* **2015**, *15*, 6247.

(95) Figgen, D.; Peterson, K. A.; Dolg, M.; Stoll, H. *J Chem Phys* **2009**, *130*.

(96) Shu, Y. N.; Levine, B. G. *J Chem Phys* **2013**, *139*.

(97) Shu, Y. N.; Levine, B. G. *J Phys Chem C* **2015**, *119*, 1737.

(98) Shu, Y. N.; Levine, B. G. *J Phys Chem C* **2014**, *118*, 7669.

Chapter 15

Non-radiative Recombination Mechanism of Lead Halide Perovskite Material

"A method is more important than a discovery, since the right method will lead to new and even more important discoveries"

-Lev D. Landau (1908-1968)

In recent years, lead halide perovskite (LHP) material has emerged as a new option for solar cell photosensitizer.[1] LHP composed of earth-abundant elements with relatively simple synthesis procedure[2-4] combined with its optical property tunability, it is attracting a tremendous amount of research efforts. The optical properties of LHP were investigated decades ago[5-8]. However, it was until 2009 the first LHP based solar cell was made with 4% power conversion efficiency.[9] A burst of research started from 2012 when successful fabrication method pushed the efficiency up to 9.7%[10] and 10.9%.[11] Within four years, the conversion efficiency reached around 20% by the end of 2015[12-14] and the current world record is 22.1%.[15] The LHP crystal structure is characterized by formula AMX_3, where A is a small organic cation with relatively large radius, among which, methylammonium ($CH_3NH_3^+$)[16-18] and formamidinium ($H_2NCHNH_2^+$)[19-21] are the two representative ones. M is the metal cation with octahedral coordination of X, typically

being lead or tin. And X is the halide anion. Among various LHP configurations, $CH_3NH_3PbI_3$ and $H_2NCHNH_2PbI_3$ are acting as leading types.[22-29] However, due to the toxicity of lead and instability when exposed to moisture environment[13,30,31], research based on lead-free LHP is becoming an active field of research.[32-34] One possible candidate in reducing the toxicity is Sn, but the stability under humidity is even worse.[32]

LHP is believed to be direct band gap semiconductor,[27,35-41] with band gap can be made in visible range[10,42-46] and close to the 1.1-1.5 eV optimal band gap value based on Shockley-Queisser model.[47] The metal cation and halide anion have a direct effect on the band gap as involved in the electronic structures. The valence band is characterized by a mixture of metal cation s orbital and halide anion p orbitals with anti-bonding interactions, while the conduction band involves metal cation p orbital character.[22,48,49] By changing the halide anion type or the ratio of mixed halides, it is possible to tune the band gap of LHP.[19,43,45,49-51] The organic cation will affect the band gap by changing the ground state stable geometries of the crystal structure.[22,52,53] Generally, the tin LHP has near-infrared band gap[6,32,42,54], lead LHP with I, Br and Cl fall into red, green and blue visible range band gap.[55-58]

Single crystal LHP has low recombination rate with nearly unity internal quantum efficiency[59-63] and high charge carrier mobility.[64-66] Based on our previous theoretical work,[67-70] the low combination rete is suggesting a high energy level of conical intersection (CI) beyond the band gap. CIs are nuclear geometries at which two or more potential energy surfaces (PESs) are degenerate.[71] CI provides efficient pathway for non-radiative

recombination process.[72] Recombination mechanism has been investigated previously mostly with the perovskite solar cell device.[44,56,65,73,74] It is interesting and important to unravel the possible non-radiative recombination pathway of the single crystal and understand why it has near unity internal quantum efficiency. In the present work, we apply ab initio molecular dynamics simulation to investigate the excited state dynamics of LHP with our model cluster and to address the questions we proposed above.

15.1 Methods

15.1.1 Nonadiabatic molecular dynamics simulations

To investigate the possible non-radiative decay mechanism, we have applied ab initio multiple spawning (AIMS) to simulate the excited state dynamics of our LHP model cluster which represents the crystal structure $CsPbCl_3$ (CP cluster). AIMS is a full molecular (nuclear and electronic) wavefunction ansatz that solves the full time-dependent Schrodinger equation.[75,76] The AIMS nuclear wavefunction is expanded as a linear combination of the frozen Gaussian basis,[77] with the average position and momentum of each basis function evolve on an adiabatic PES approximated at the state averaged complete active space self-consistent field (SA-CASSCF) level of theory.[78] In this sense, the average position and momentum of each basis function can be treated as the classical counterpart of the nucleus. A SA-CASSCF active space of two electrons in four orbitals with state average five was used for CP cluster. The small split double zeta basis set (SBKJC_VDZ) and relativistic compact effective potentials were used.[79-81] The

excitation energies and state characters of the low-lying electronic states of the CP cluster computed at this level of theory are in good agreement with those predicted at the trustworthy equation-of-motion coupled cluster with single and double excitations[82] (EOM-CCSD) level of theory using the SBKJC_VDZ(p,2d) polarized basis set and effective core potentials (ECPs).[83,84] All dynamical simulations were initiated on the lowest singlet excited state (S_1). Twenty initial nuclear basis functions were chosen with their initial average positions and momenta randomly sampled from the ground state vibrational Wigner distribution calculated with the harmonic approximation by density functional theory (DFT) with B3LYP functional[85-87] and SBKJC_VDZ(p,2d) polarized basis set and effective core potentials (ECPs). The nuclear wavefunction is propagated at a time step of 20 a.u. (~0.5 fs) with NPI scheme in evaluating time derivative coupling.[88] Simulation data was collected for 1 ps after excitation or when the threshold for population decay is reached. The widths of the frozen Gaussian basis functions were chosen to be 20.0, 25.0, 20.0, 30.0 and 6.0 $Bohr^2$ for the lead, chlorine, caesium, nitrogen and hydrogen atoms, respectively. The frozen Gaussian widths are approximated by maximizing the frozen Gaussian wave function and the exact wave function overlap.[89] All dynamical calculations were performed using the FMSMolPro package.[76]

15.1.2 Ab-Initio calculations

To further analyze the accuracy of SA-CASSCF PES employed in the AIMS simulations, additional single point calculations and geometry optimizations were performed at the multi-state complete active space second-order perturbation (MS-

CASPT2) level of theory[90] to consider the dynamic correlation to the SA-CASSCF reference. All MS-CASPT2 calculations were performed with SBKJC_VDZ(p,2d) polarized basis set and effective core potentials (ECPs).[83,84] The same active space is employed as the SA-CASSCF level of theory with 8 orbitals beside active orbitals were correlated in MS-CASPT2 calculations. Conical intersections were optimized using a multiplier-penalty approach[91] which does not require the calculation of nonadiabatic coupling vectors. Ground state minimum geometry is optimized at DFT/B3LYP level of theory with SBKJC_VDZ(p,2d) polarized basis set and effective core potentials (ECPs) using MolPro[92,93] software package. Single point calculations including SA-CASSCF, MS-CASPT2, and EOM-CCSD were performed with MolPro[94-101] software package. Excited state minimum and CI geometries optimization at the SA-CASSCF and MS-CASPT2 levels of theory were calculated by MolPro in conjunction with CIOpt.[91]

15.2 Results and Discussion

15.2.1 Model cluster and ground state minimum

Due to the existing and interchanging of multiple phases of perovskite material when external stimuli are applied like temperature,[22,29,42,102] the thermal stability is of great importance to crystalize the material into the photoactive perovskite phase. It has been suggested from the previous studies that the thermal stability is improved by adopting smaller size inorganic cations to formed mixed cations perovskite.[29,103-107] Hence, it is important to investigate both organic and inorganic A position cations behaviors. In

this work, we will present the model cluster contains inorganic cations. The CP cluster contains inorganic cation with chemical formula $PbCl_6Cs_4$.

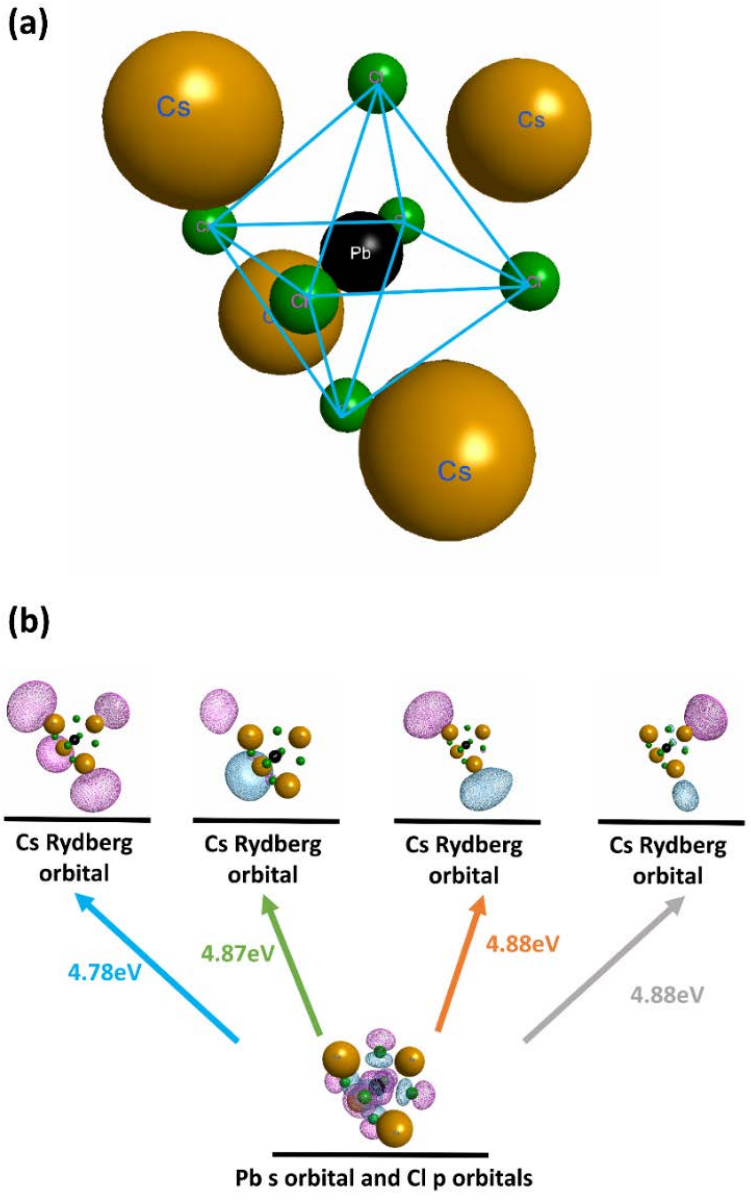


Figure 15-1. (a) Ground state minimum geometry of the CP model cluster $PbCl_6Cs_4$, with Pb, Cl and Cs represented by black, green and brown colors respectively. The blue thin line linked together to show the octahedral structure of $[PbCl_6]^{4-}$. (b) The first four

Figure 15-1(cont'd). excited states are nearly degenerate with the associated excitation energies and characters calculated at EOM-CCSD theoretical level shown in the graph. Single arrow with the same color indicates the same excited state. The S_1, S_2, S_3, and S_4 are represented with blue, green, orange and gray colors respectively.

The ground state minimum geometry of CP cluster is shown in Figure 15-1(a). The six Pb-Cl bond lengths vary from 2.89 to 2.90 Å with 179.8, 179.9 and 179.8° of Cl-Pb-Cl angles for opposite position Cl atoms (treating Pb as the center and in perfect octahedral coordination, there exist three Cl-Pb-Cl angles equal 180°, here we call these Cl atoms are opposite positions with each other). EOM-CCSD calculation predicts the first four excited states being nearly degenerate, with S_1 to S_4 has excitation energies of 4.78, 4.87, 4.88 and 4.88 eV respectively. The corresponding excited state characters are shown in Figure 15-1(b). Each arrow indicates a type of excitation character (singly excited determinant). The arrows with the same color represent the same excited state. From Figure 15-1(b), we found that the nearly degenerate S_1 to S_4 states have similar excitation character described by an excitation from mixed Pb s and Cl p orbital to mixed Cs s and p orbitals. The highest occupied molecular orbital (HOMO) has the similar character as the valence band minimum (VBM) of that has been calculated from density functional theory (DFT) for periodic systems previously.[49] However, the virtual orbitals associated with the low-lying excited states are different from what has been found from DFT calculations of cesium halide perovskites, which claimed the conduction band minimum

(CBM) is composed of metal cation p orbital.[108,109] This is related to the size effect of the model cluster, and we will explain the causing of the difference in the following text.

It is noteworthy the huge difference of the excitation energy of our CP cluster as compared to the visible range band gaps of perovskite materials. The discrepancy comes from the size effect, which known in the community.[110] However, investigating the ground state stable structure and make property prediction is not the goal of the current work. Previous studies from our group indicate the CI introduced non-radiative decay pathway has week size dependence.[70] The localized deformation of the nuclear geometry is the key to reaching the CI geometry and facilitate efficient non-radiative recombination process, which rationalizes our choice of small model cluster despite the fact that the ground state stable CP cluster's excited states electronic structures disagree with that have been predicted by DFT on periodic systems.

The discrepancy between the low-lying excited states excitation character of CP cluster and CBM from DFT calculation is due to the size effect as well. This suggests the energy levels associated with the excitation character described by Pb s and Cl p mixed orbital to Pb p orbital is higher than to Cs s and p mixed orbitals at the model cluster geometry. However, it is still interesting to investigate the possible non-radiative decay pathway associated this type of states as it provides a possibility of population decay.

15.2.2 Dynamical simulation of the CP model cluster

We employed SA-CASSCF theoretical level in AIMS dynamical simulations, the detailed parameters chosen is shown in Table 15-1. SA-CASSCF predicts 5.35 eV vertical

excitation energy of S_1 state at Frank-Condon (FC) geometry (as compared to 4.78 eV calculated at EOM-CCSD theoretical level). The 0.57 eV discrepancy is common at SA-CASSCF theoretical level as it lacks dynamical correlations. The accuracy of the SA-CASSCF PES will be further assessed in the following sections.

In 1 ps simulation, there is no population decay observed. The S_1-S_0 energy gaps show strong coherence oscillation during the dynamical simulation as shown in Figure 15-2. All twenty individual nuclear basis functions behave synchronous oscillation. The initial energy gap oscillation spans roughly 4.5 eV with the time period of around 140 fs. The coherence oscillation is damped in 1 ps. By analyzing the nuclear motions of individual trajectories, we found that this strong coherence energy gap oscillation is accompanied by Pb-Cl bond length change.

Table 15-1. Size of the Active Space (Abbreviated: Number of Active Electrons/Number of Active Orbitals) and Number of States to Be Averaged in Our SA-CASSCF Calculations for CP Cluster. Also given are the numbers of occupied orbitals to be correlated in the MS-CASPT2 calculations. Note that more orbitals are correlated in the MS-CASPT2.

Model Cluster	Electronic Structure Methods	Active Space	No. of States To Be Averaged	Total No. of Correlated Orbitals
CP Cluster	SA-CASSCF	2/4	5	N/A
	MS-CASPT2			9

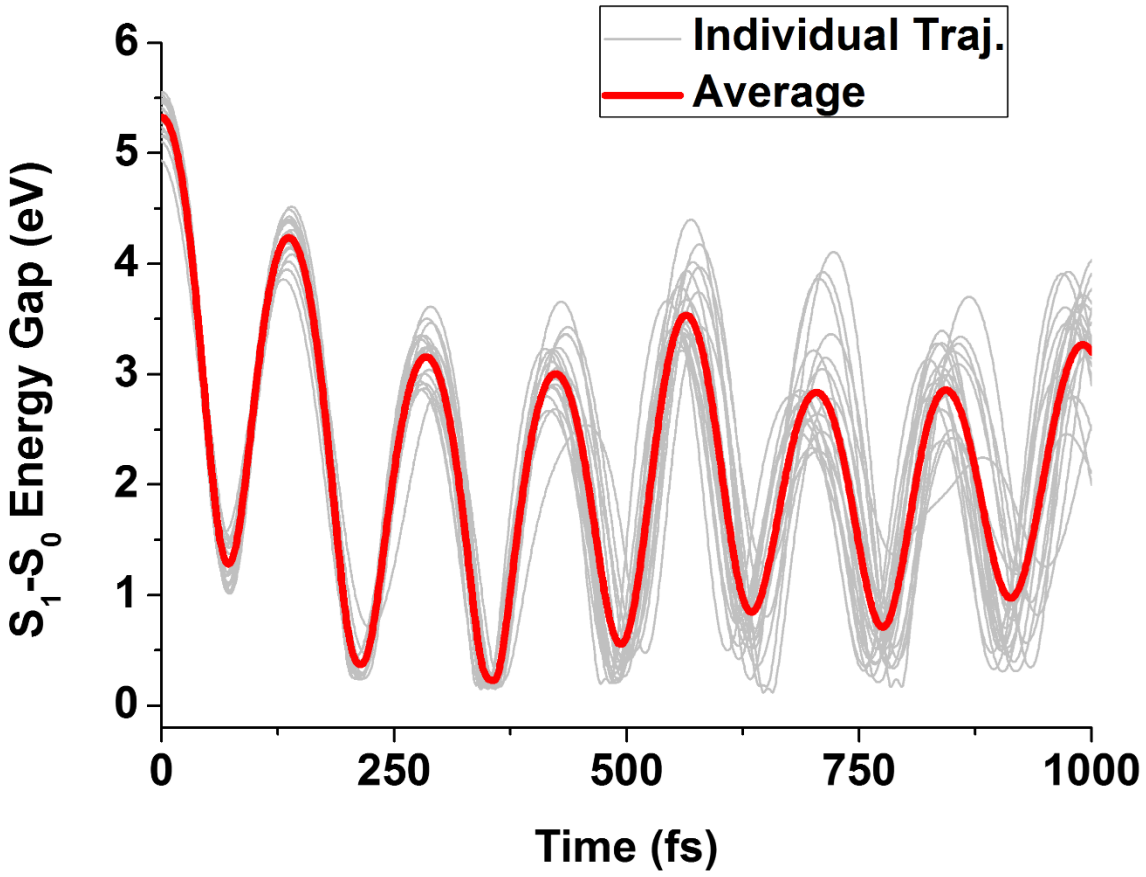


Figure 15-2. Average S_1-S_0 energy gap of the model CP cluster as a function of time (bold red) with the energy gaps of 20 initial nuclear basis functions (gray).

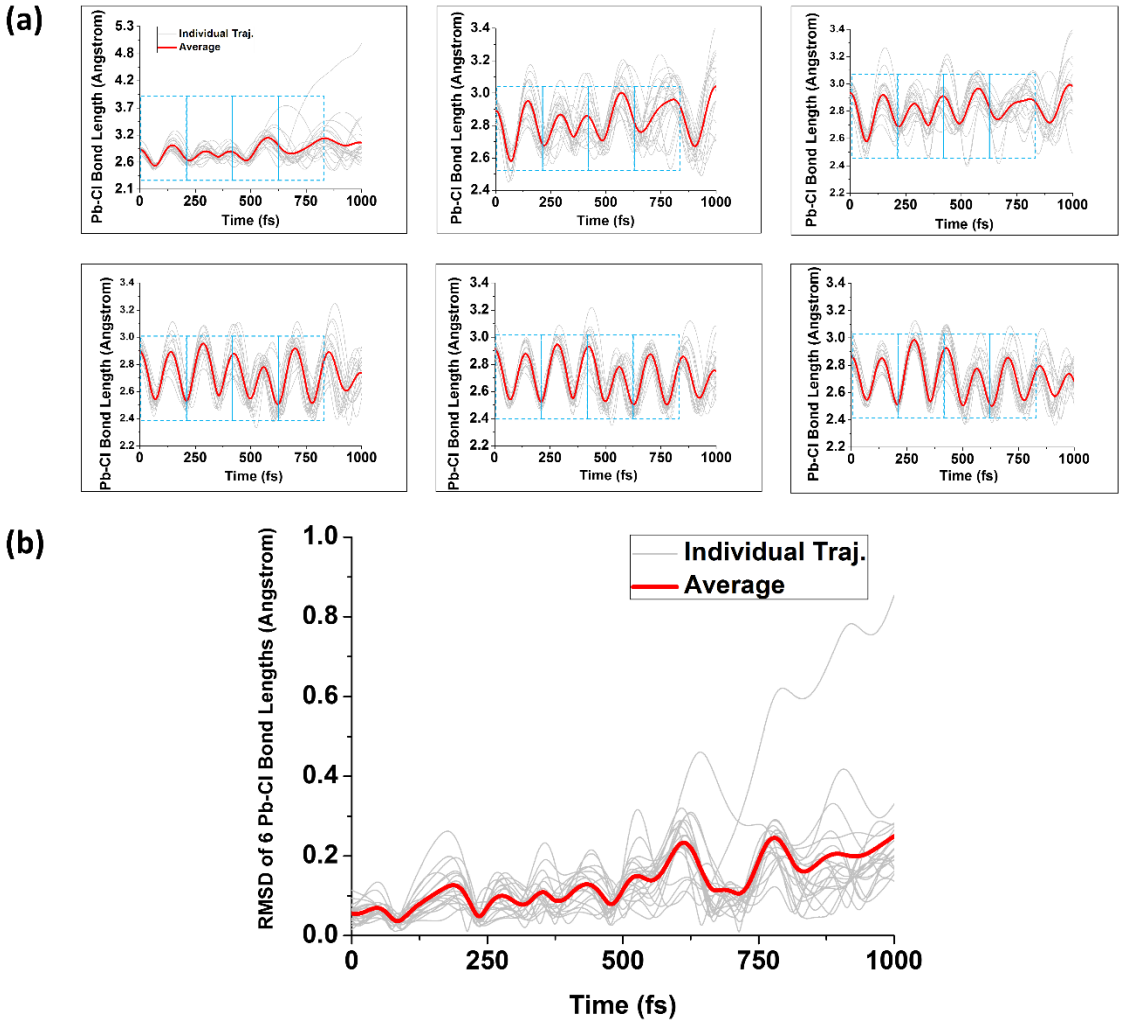


Figure 15-3. (a) Top six panels: six averaged Pb-Cl bond lengths as a function of time(bold red) with twenty individual initial nuclear basis functions shown in gray color. The small blue box in each panel spans the same period of time. (b) Bottom panel: the averaged RMSD of the six Pb-Cl bond lengths as a function of time (bold red) with twenty individual initial nuclear basis functions shown in gray color.

Figure 15-3 shows the Pb-Cl bonds change as a function of time. Top six panels show the six individual Pb-Cl bond length as a function of time. The six panels are ordered according to the averaged bond length during the 1 ps simulation for each initial nuclear basis functions. Each small blue box spans the same period for comparison. All six Pb-Cl bonds behave almost the synchronous Pb-Cl vibrational motion corresponds to 204.4 cm^{-1} frequency calculated at DFT/B3LYP/SBKJC_VDZ theoretical level. Averaged single Pb-Cl bond length oscillates between 2.50 to 3.12 Å (the ground state minimum geometry has Pb-Cl bond lengths close to 2.90 Å). This Pb-Cl bond oscillation is damped at the longer time of simulation. By comparing Figure 15-2 and Figure 15-3(a) top six panels, we found that the coherence oscillation of S_1-S_0 energy gap and Pb-Cl bond length change spans the same period of time. Figure 15-3 bottom panel shows the RMSD of six Pb-Cl bond lengths as a function of time. The increasing trend of RMSD indicates the $PbCl_6$ nuclear geometry distorts from perfect octahedral coordination at longer simulation time.

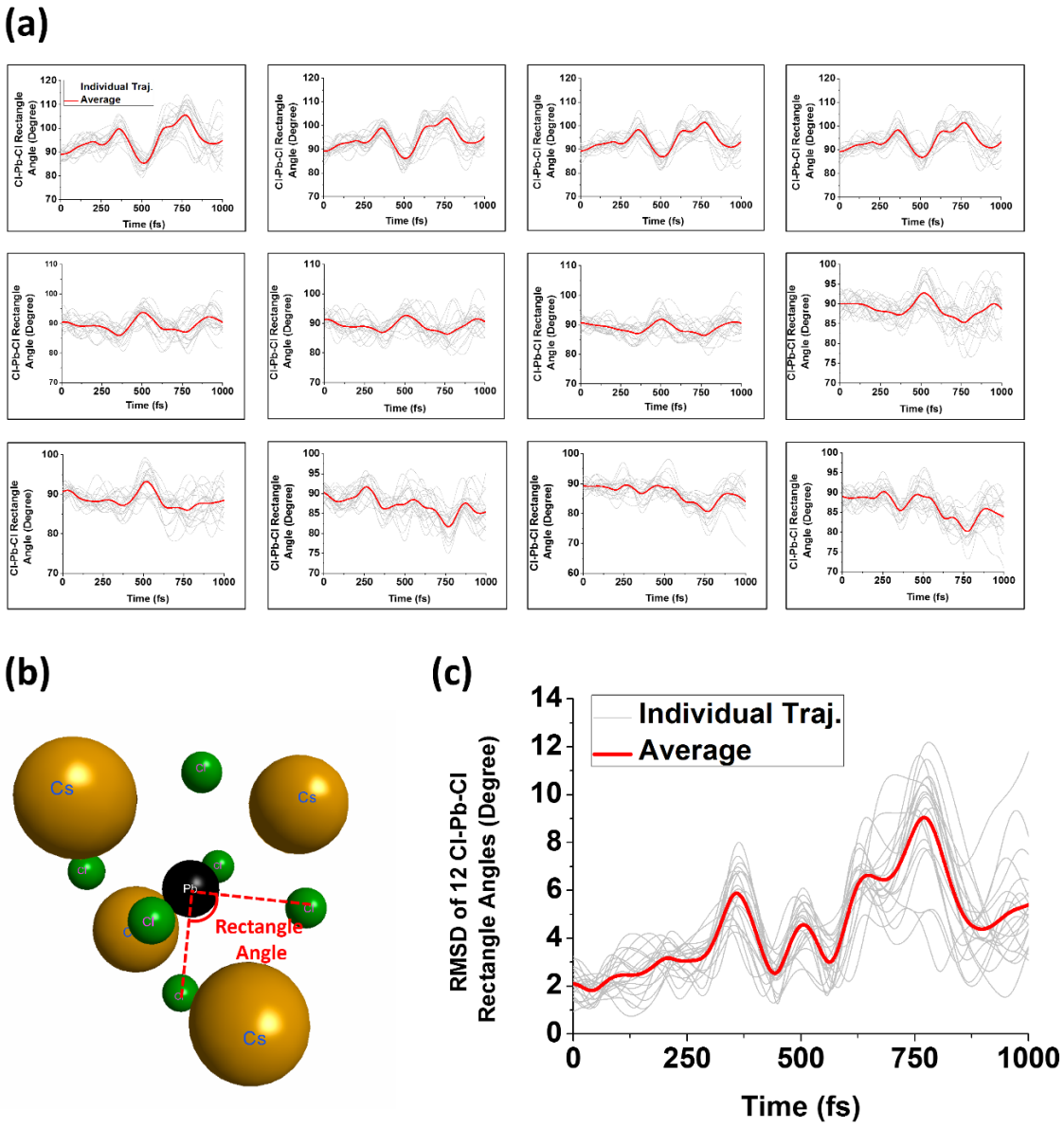


Figure 15-4. (a) Top twelve panels: twelve averaged Cl-Pb-Cl rectangle angles as a function of time(bold red) with twenty individual initial nuclear basis functions shown in gray color. (b) Bottom left panel indicates the representative atoms associated with a rectangle angle. (c) Bottom right panel: the averaged RMSD of the twelve Cl-Pb-Cl rectangle angles as a function of time (bold red) with twenty individual initial nuclear basis functions shown in gray color.

The geometry deformation of the cluster is further investigated by analyzing the Cl-Pb-Cl angle. There exist two types of Cl-Pb-Cl angles at perfect octahedral geometry, one is the rectangle angle (90°), the other corresponds to a straight angle (180°). Figure 15-4 shows the rectangle angle as a function of time. The top twelve panels show the individual Cl-Pb-Cl rectangle angle with individual trajectories vary from 69° to 113°, and the RMSD of the twelve rectangle angle changes from 0.7° and 12.2° as shown in Figure 15-4(c). The averaged Cl-Pb-Cl rectangle angle RMSD also suggests the non-symmetric distortion from octahedral geometry increases at the longer time. Figure 15-5(a) shows the three straight angles as a function of time with averaged value changes from 178° to 162° in 1 ps. Some of the individual trajectories show that the straight angle changes to 152° in 1 ps, which indicates a strong geometrical distortion from the perfect octahedral structure. The corresponding RMSD as a function of time is shown in Figure 15-5(c). The averaged RMSD changes from 0.7° to 3.1°. Note that the non-symmetric distortion mentioned above does not mean that the cluster is not distorted, for example, if all three straight Cl-Pb-Cl changes to 160°, the RMSD will be 0°, however, the cluster is still highly distorted from the perfect octahedral structure.

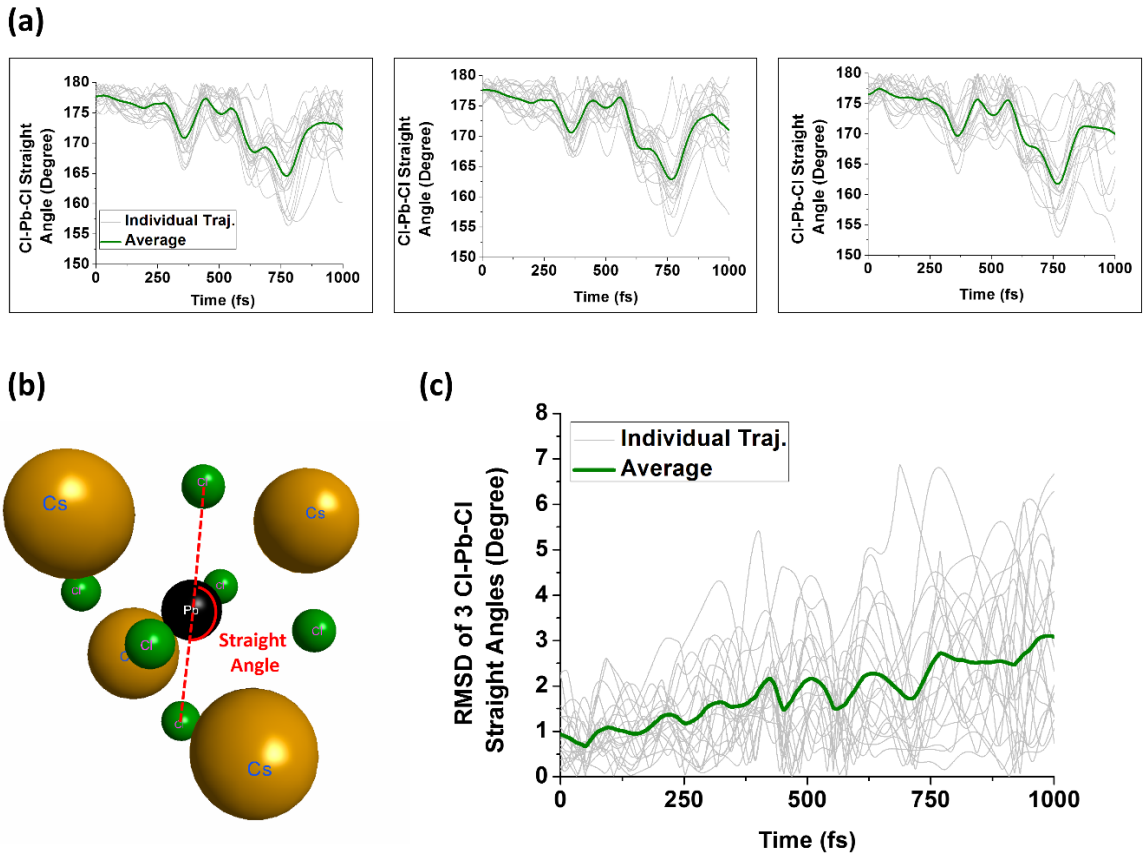


Figure 15-5. (a) Top three panels: three averaged Cl-Pb-Cl straight angles as a function of time(bold green) with twenty individual initial nuclear basis functions shown in gray color. (b) Bottom left panel indicates the representative atoms associated with a straight angle. (c) Bottom right panel: the averaged RMSD of the three Cl-Pb-Cl straight angles as a function of time (bold red) with twenty individual initial nuclear basis functions shown in gray color.

15.2.3 Potential energy surface analysis of CP cluster

It is surprised that the S_1-S_0 energy difference reaches below 0.5 eV without hitting the CI. At CI geometry, zero S_1-S_0 gap facilitates the population transfer. Also, the dynamical behavior involves Pb-Cl vibrational motion without any behavior related to Cs cation, despite the fact that the molecules are put on S_1 state at the beginning of the simulation whose character involves Cs s and p orbitals. To answer these two questions, we first analyze the S_1 state excitation character. As shown in Figure 15-1(b), the HOMO has an anti-bonding interaction between Pb s orbital and Cl p orbital. General chemistry tells us by removing an electron from an anti-bonding orbital stabilizes the molecule. This is exactly what happens in the excited state dynamics simulation. The shorter the Pb-Cl bond length, the stronger the anti-bonding interaction between the Pb s orbital and Cl p orbital, and hence, the more stable of the excited cluster. This explains the observation of Pb-Cl bond oscillation during our simulation. This also rationalizes why Cs orbitals are involved, but no dynamical behavior is observed. However, we would expect that at longer simulation time, Cs cation will possibly be involved into dynamics. By scanning the PES along symmetrical Pb-Cl bond lengths changing, we obtain a PES as shown in Figure 15-6. It is clear from the PES that the S_1-S_0 energy gap reduces as Pb-Cl bond length shortened. However, as reaching the repulsive region of the PES, this gap is never reduces to zero. However, we believe that by varying the geometry, the CI should be reached in that repulsive region of the PES. This high energy level of the CI also explains the fact that no population decay was observed during our AIMS simulation.

The blue dotted line in Figure 15-6 indicates the two ends of the PES well corresponding to 3.10 and 2.55 Å of Pb-Cl bond lengths. We have mentioned above that the averaged Pb-Cl bond lengths vary from 2.50 to 3.12 Å which is very close to what we have shown in Figure 15-6. Hence, the AIMS simulation can be understood as a coherence Pb-Cl bond vibration motion.

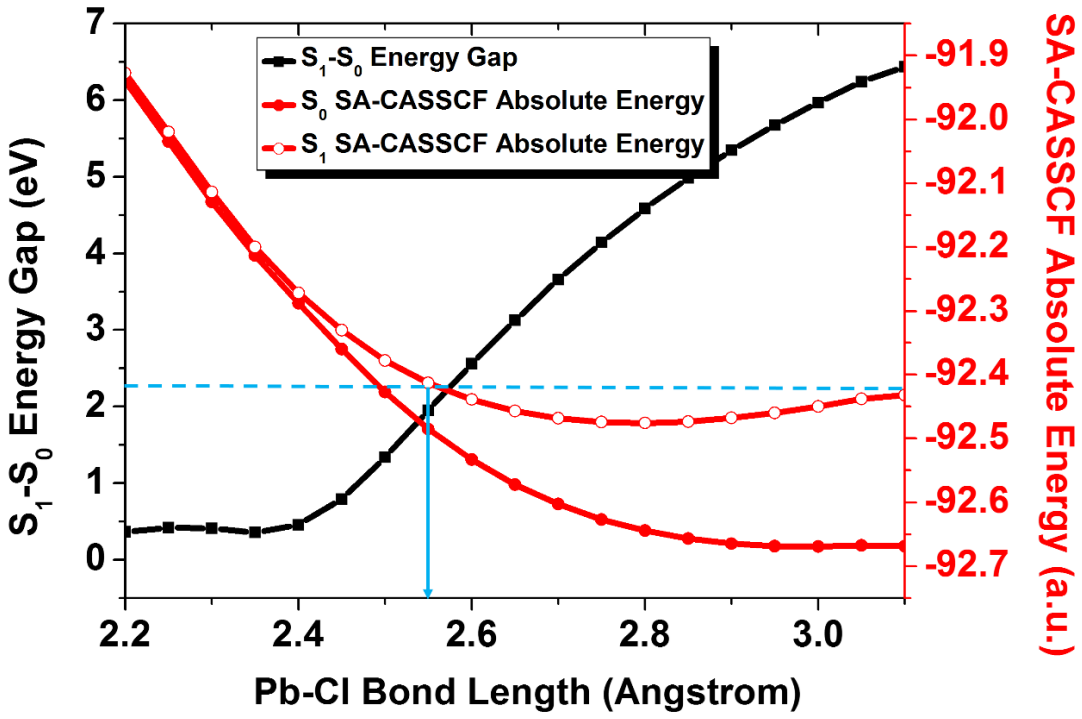


Figure 15-6. Scanned PES along a symmetric stretch of Pb-Cl bond lengths. The S_1 and S_0 state absolute SA-CASSCF energy PESs are shown as red curves; the S_1-S_0 energy gap is shown as the black curve. The blue dotted line indicates the two ends of the PES well where the right-hand side end corresponds to the FC geometry; the left-hand side end corresponds to the other geometry that has been reached during dynamical simulations.

As has been explained above, a stable excited CP cluster with shorter Pb-Cl bond length compared with FC geometry should be expected. We explore this possibility by optimizing the stable minimum on the first excited state PES (S_1 min) geometry at both SA-CASSCF and MS-CASPT2 theoretical levels. The results are summarized in Table 15-2 and Figure 15-7. As the geometry is very close to the FC geometry, we will denote this S_1 min geometry as FC-S_1 min in the following text. SA-CASSCF and MS-CASPT2 single point calculations performed on SA-CASSCF optimized FC-S_1 min geometry predict the 3.72 and 4.22 eV S_1 state energy above the ground state minimum. The small 0.5 eV discrepancy between the SA-CASSCF and MS-CASPT2 indicates the geometry optimized at SA-CASSCF theoretical level should be qualitatively correct. This result validates our excited state dynamics simulation. The accuracy of the SA-CASSCF PES is further assessed by re-optimizing the FC-S_1 min geometry at MS-CASPT2 theoretical level. The MS-CASPT2 optimized FC-S_1 min geometry has a shortest Pb-Cl bond length as 2.54 Å compared to 2.72 Å as that of SA-CASSCF optimized FC-S_1 min geometry. Both results show the shorter Pb-Cl bond compared to the FC geometry (2.90 Å). SA-CASSCF and MS-CASPT2 predict 4.30 and 3.05 eV S_1 state above the ground state minimum at MS-CASPT2 optimized FC-S_1 min geometry. This suggests a 1.25 eV relative large discrepancy between SA-CASSCF and MS-CASPT2 results. However, by comparing the SA-CASSCF S_1 state energy at SA-CASSCF optimized FC-S_1 min geometry, and MS-CASPT2 S_1 state energy at MS-CASPT2 optimized FC-S_1 min geometry, we found a 0.67 eV discrepancy, which suggests the SA-CASSCF PES employed in AIMS simulation should be qualitatively good.

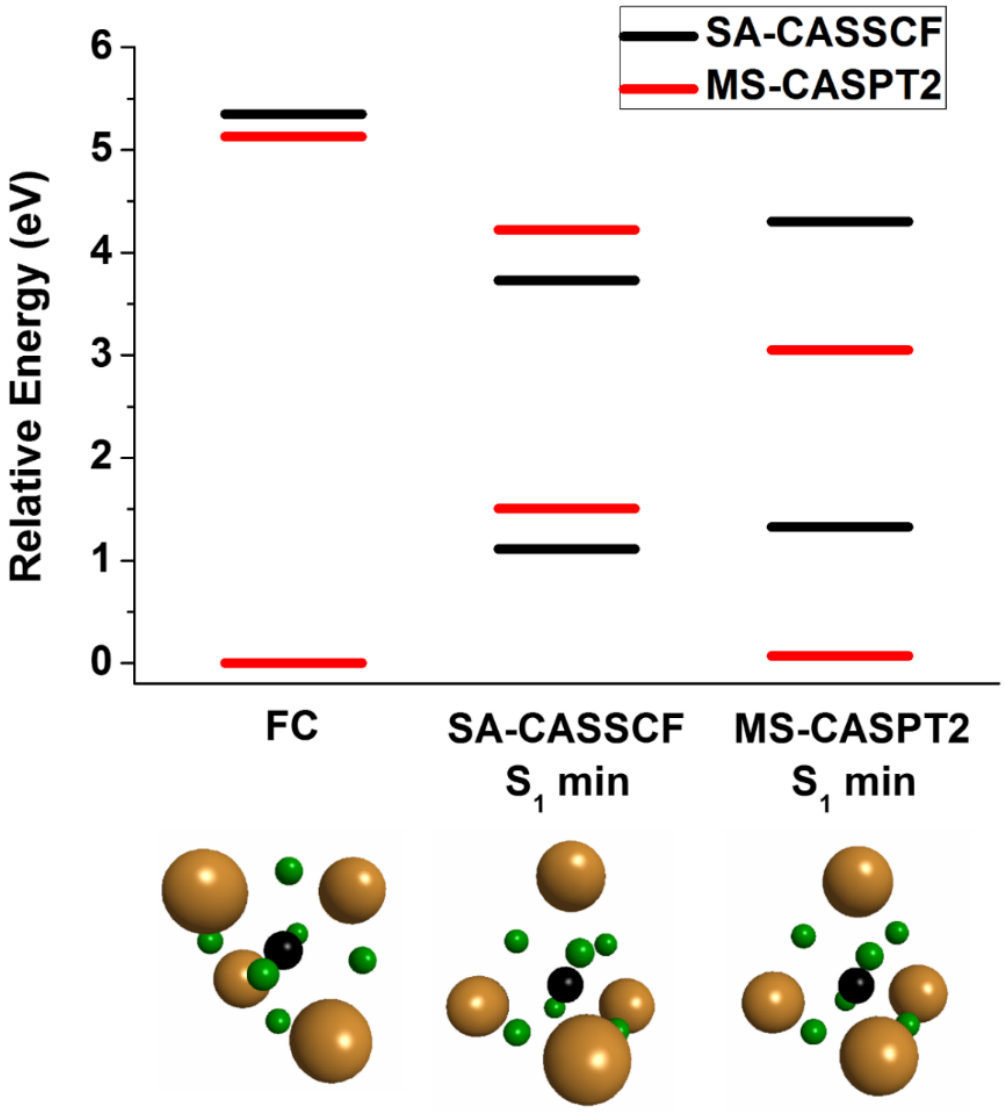


Figure 15-7. The S_0 and S_1 energies of FC geometry and FC-S_1 min geometries optimized at both SA-CASSCF and MS-CASPT2 theoretical levels. Black and red lines correspond to the energies of the single point calculation computed at the SA-CASSCF and MS-CASPT2 theoretical levels with SBKJC_VDZ and SBKJC_VDZ(p,2d) basis sets with ECPs. Energies are shifted so that the S_0 energy of the FC point is zero for all levels of theory. Optimized geometries are shown as insets.

It is still a question that whether a conical intersection exists at low energy level. We further explore the SA-CASSCF PES by optimizing minimal energy CI (MECI), and stable minimum on the first excited state PES (S1 min) geometries near the CI region of the PES if exists. The SA-CASSCF optimized MECI geometry has relative short Pb-Cl bonds (shortest Pb-Cl bond is 2.63 Å), and one of Cs atom is dissociated from the rest of the cluster as shown in the insets of Figure 15-8. SA-CASSCF and MS-CASPT2 predict the S_1 state of SA-CASSCF optimized MECI locates at 4.18 and 4.73 eV above the ground state minimum. MS-CASPT2 also predicts small S_1-S_0 energy gap (0.23 eV) at SA-CASSCF optimized MECI geometry. It is shown from the calculations that the MECI is at a very high energy which is far beyond the visible range band gap of perovskite material. Hence, it is concluded here that this type of MECI introduced by dissociates the Cs cation is not important in visible range photochemistry. A stable excited state minimum close to the MECI geometry was found at SA-CASSCF theoretical level. As this S_1 minimum geometry is close to the MECI geometry, we will denote this S_1 minimum as MECI-S_1 min in the following text. SA-CASSCF and MS-CASPT2 predict 4.15 and 4.70 eV S_1 state energy above the ground state minimum at SA-CASSCF optimized MECI-S_1 min. This energy is very close to that of MECI geometry. Moreover, the S_1-S_0 energy gap at SA-CASSCF optimized MECI-S_1 min is small (0.16 and 0.27 eV gap at SA-CASSCF and MS-CASPT2 theoretical levels respectively). Hence this MECI is possibly a peaked MECI.

MECI is re-optimized at MS-CASPT2 theoretical level. S_1 state energy was found at 5.11 and 3.57 eV at SA-CASSCF and MS-CASPT2 theoretical levels respectively. However,

we were not able to obtain a stable MECI-S_1 min geometry with lower energy compared to MECI at MS-CASPT2 theoretical level. This indicates that the MECI may be a peaked MECI, which is partially supported by the fact that SA-CASSCF optimized MECI-S_1 min is very close to MECI in terms of energy and geometry. These results are summarized in Table 15-2 and Figure 15-9.

Table 15-2. The S_0 and S_1 energies (in eV) at three important points on the PES of the CP cluster calculated at the CASSCF and MS-CASPT2 levels of theory. S_1-S_0 MECI and S_1min geometries are optimized at both the CASSCF and MS-CASPT2 levels of theory. Note that all SA-CASSCF calculations are performed with SBKJC_VDZ basis set while all MS-CASPT2 calculations are performed with SBKJC_VDZ(p,2d) basis set.

Geometry	Optimization Method	CASSCF S_0 Energy /SBKJC_VDZ	CASSCF S_1 Energy /SBKJC_VDZ	MS-CASPT2 S_0 Energy /SBKJC_VDZ (p,2d)	MS-CASPT2 S_1 Energy /SBKJC_VDZ (p,2d)
FC geometry	DFT/B3LYP	0	5.35	0	5.13
S_0-S_1 MECI	SA-CASSCF	4.15	4.18	4.50	4.73
	MS-CASPT2	5.04	5.11	3.55	3.57
MECI-S_1 min	SA-CASSCF	3.99	4.15	4.33	4.70
	MS-CASPT2	5.04	5.11	3.55	3.57
FC-S_1 min	SA-CASSCF	1.12	3.72	1.51	4.22
	MS-CASPT2	1.33	4.30	0.07	3.05
S_0-S_1 Pb-Cl MECI	SA-CASSCF	3.28	3.29	3.00	3.84
	MS-CASPT2	3.86	4.84	2.82	2.84

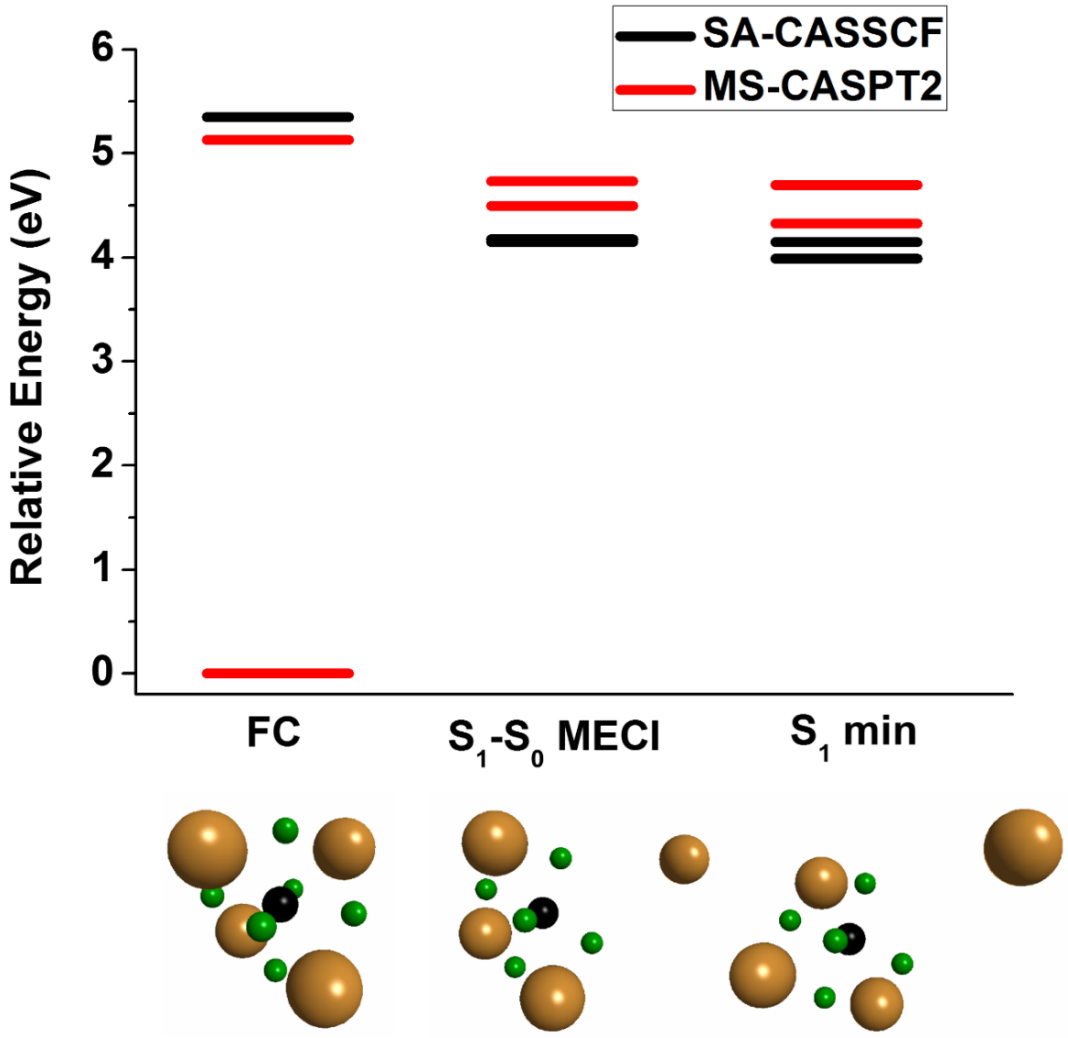


Figure 15-8. The S_0 and S_1 energies of three important points on the PES of the CP model cluster: the Franck-Condon (FC) point, S_0-S_1 MECI, S_1 min optimized on SA-CASSCF theoretical level. Black and red lines correspond to the energies of the single point calculation computed at the SA-CASSCF and MS-CASPT2 theoretical levels. Energies are shifted so that the S_0 energy of the FC point is zero for all levels of theory. Optimized geometries are shown as insets.

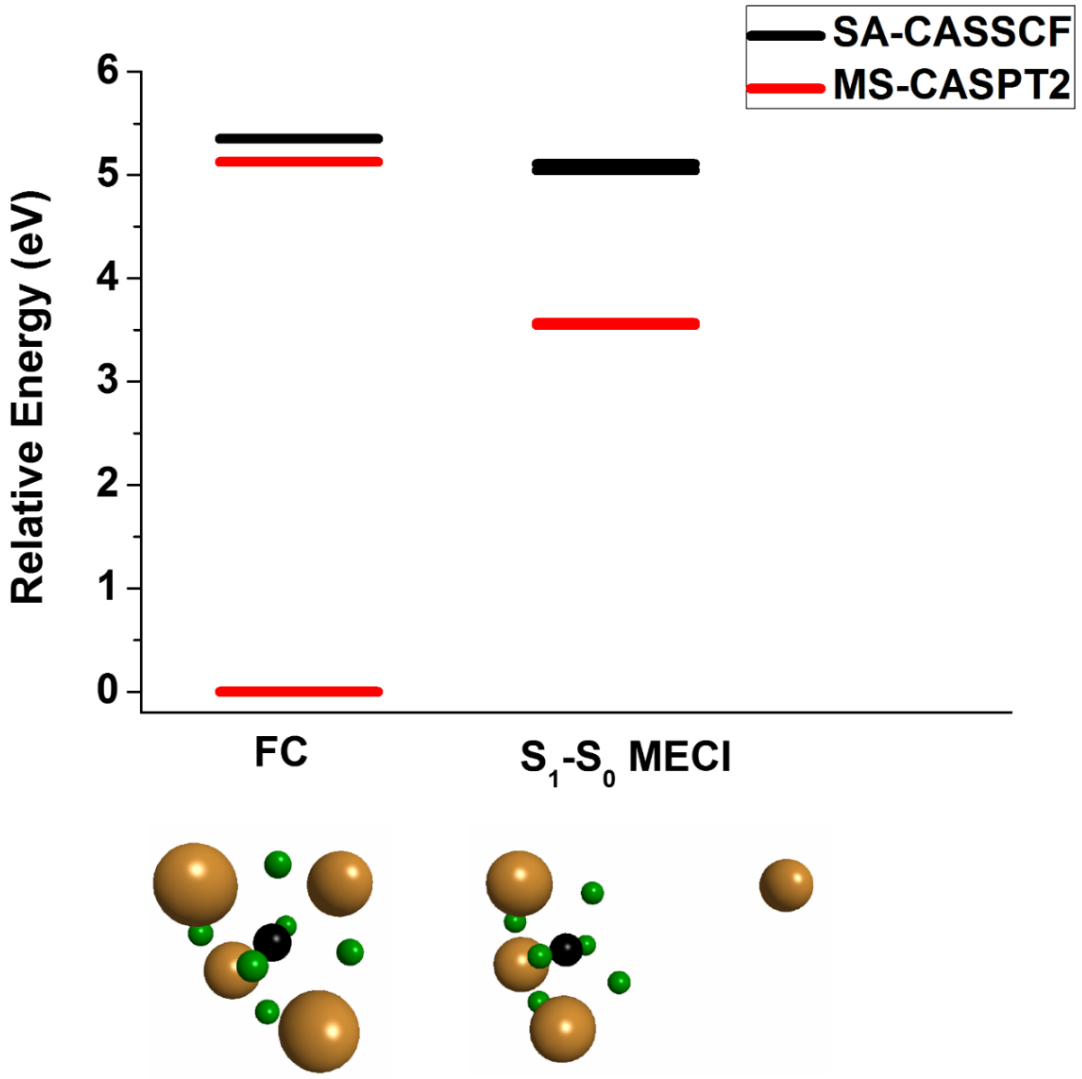


Figure 15-9. The S_0 and S_1 energies of three important points on the PES of the CP model cluster: the Franck-Condon (FC) point, S_0-S_1 MECI, S_1 min optimized on SA-CASSCF theoretical level. Black and red lines correspond to the energies of the single point calculation computed at the SA-CASSCF and MS-CASPT2 theoretical levels. Energies are shifted so that the S_0 energy of the FC point is zero for all levels of theory. Optimized geometries are shown as insets.

It is worth notating that the VBM and CBM of perovskite material are described by a mixed Pb s and Cl p orbital and Pb p orbital respectively. With this in mind, we explored another region of the SA-CASSCF PES of the CP cluster to investigate the possibility of existing of MECI with similar geometry that has dissociated Pb-Cl bond. We found SA-CASSCF MECI with dissociated Pb-Cl bond as shown in Figure 15-10 (we will denote this type of MECI as PbCl MECI in the following text). SA-CASSCF predicts 3.29 eV S_1 state energy for PbCl MECI as compared to 3.84 eV computed at MS-CASPT2 theoretical level. However, we do found that MS-CASPT2 predicts 0.84 eV relatively large S_1-S_0 energy gap at SA-CASSCF optimized PbCl MECI geometry. We further explore the possibility of existing PbCl MECI by re-optimizing the geometry at MS-CASPT2 theoretical level. MS-CASPT2 predicts 2.84 eV S_1 state energy compared with ground state minimum. This result indicates a 0.45 eV discrepancy between the SA-CASSCF and MS-CASPT2 theoretical levels at SA-CASSCF and MS-CASPT2 optimized Pb-Cl MECI geometry. The results are summarized in Table 15-2 and Figure 15-10.

Partially occupied natural orbitals (PONOs) are computed at SA-CASSCF/SBKJC_VDZ ECP and basis set theoretical level as shown in Figure 15-11. The PONOs at FC geometry are computed by diagonalization of S_1 density matrix while PONOs at both MECI geometries are calculated by diagonalizing averaged S_1-S_0 density matrix. At FC geometry, the nearly degenerate occupation number of two associated orbitals indicates the exacting mixing being meaningless, instead, the excitation character is described by Pb s and Cl p orbital to Cs s and p orbitals. At MECI geometry, the character of PONO

suggests that population decay could happen by transferring an electron from Cs s orbital to Pb s and Cl p mixed orbitals. This strong charge transfer character shows the ionic character of Cs atom. At PbCl MECI geometry, the PONOs show a strong charge transfer character between Pb p orbitals and Cl p orbitals.

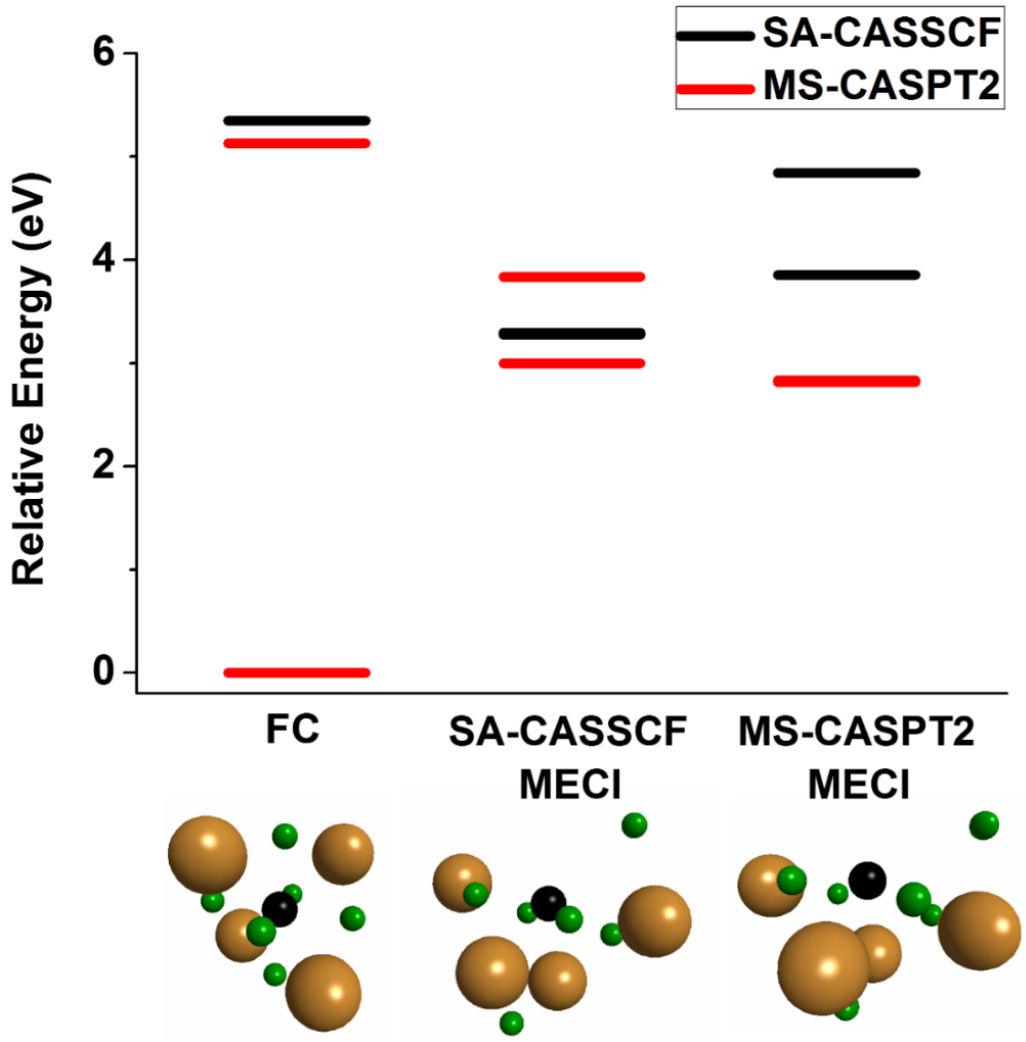


Figure 15-10. The S_0 and S_1 energies of FC geometry and Pb-Cl MECI geometries optimized at both SA-CASSCF and MS-CASPT2 theoretical levels. Black and red lines correspond to the energies of the single point calculation computed at the SA-CASSCF and

Figure 15-10 (cont'd). MS-CASPT2 theoretical levels with SBKJC_VDZ and SBKJC_VDZ(p,2d) basis sets with ECPs. Energies are shifted so that the S_0 energy of the FC point is zero for all levels of theory. Optimized geometries are shown as insets.

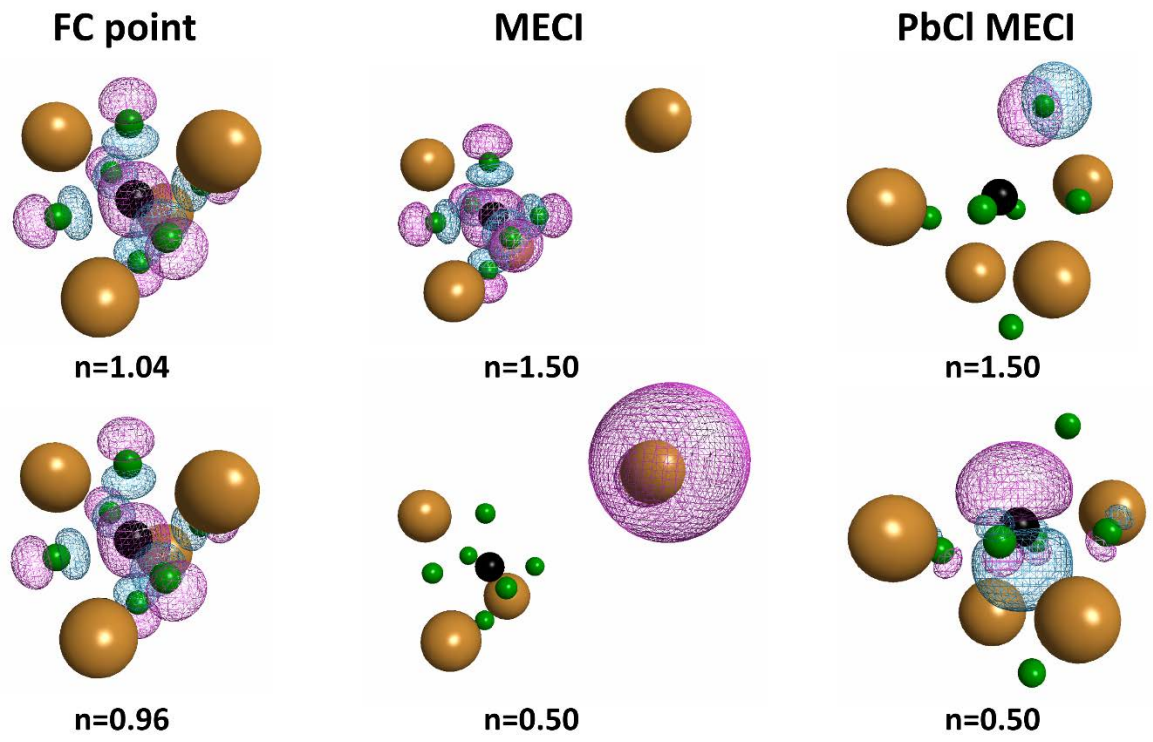


Figure 15-11. Partially occupied SA-CASSCF natural orbitals of silicon hydroxide cluster at FC (left), MECI (middle) and PbCl MECI(right) geometries are presented with occupation numbers. The FC natural orbitals are computed by diagonalization of the S_1 density matrix, while S_1–S_0 MECI natural orbitals are computed from the state-averaged S_1–S_0 density matrix.

15.3 Conclusions

The pervious calculations have shown the similar electronic structures of VBM and CBM for lead perovskite material $APbX_3$ at ground state minimum despite the various choices of A and X positions. As has been discussed in the introduction section, the VBM usually consists of mixed Pb s orbital and X p orbital with an anti-bonding interaction while the CBM is Pb p orbital. Our calculations show that CP model cluster has different excitation characters for low-lying excited state, this is attributed to the size effect.

In this work, we have shown the model cluster have lowest energy levels of MECI geometry at around 2.8 eV above the ground state minimum; this value is relatively insensitive to the size of the cluster as has been proven by our previous work.[70] The high energy level of MECI indicates that it is not significant in visible range band gap perovskite materials, and hence the efficient non-radiative decay pathway through CI is blocked. This result explains the experimental observation that single crystal perovskite has near unit quantum efficiency. Further investigation regarding the changing of the dielectric environment when the size of the cluster changes and how it relates to the energy level change of MECI is required. As has been confirmed by both experiment and theoretical calculations that the halide anions have a big effect in the material band gap, the effect of halide anion on the energy levels of MECI geometry will be summarized in the next chapter.

APPENDIX

APPENDIX

Table 15-S1. The absolute energies of the various optimized geometries of the CP cluster model at the SA-CASSCF level of theory (with active space and basis sets as described in the text).

SA-CASSCF	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point	-92.66448351	-92.46793386
SA-CASSCF optimized S_0/S_1 MECI	-92.51211935	-92.51100214
MS-CASPT2 optimized S_0/S_1 MECI	-92.47913719	-92.47682053
SA-CASSCF optimized MECI-S_1 min	-92.51787991	-92.51202877
MS-CASPT2 optimized MECI-S_1 min	N/A	N/A
SA-CASSCF optimized FC-S_1 min	-92.62346527	-92.52759184
MS-CASPT2 optimized FC-S_1 min	-92.61572189	-92.50649636
SA-CASSCF Optimized Pb-Cl S_0/S_1 MECI	-92.54412716	-92.54344926
MS-CASPT2 Optimized Pb-Cl S_0/S_1 MECI	-92.52267784	-92.48644599

Table 15-S2. The absolute energies of the various optimized geometries of the CP cluster model at the MS-CASPT2 level of theory (with active space and basis sets as described in the text).

MS-CASPT2	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point	-92.8303903	-92.6419483
SA-CASSCF optimized S_0/S_1 MECI	-92.6651312	-92.6564342
MS-CASPT2 optimized S_0/S_1 MECI	-92.6999625	-92.6992394
SA-CASSCF optimized MECI-S_1 min	-92.6714341	-92.6578257
MS-CASPT2 optimized MECI-S_1 min	N/A	N/A
SA-CASSCF optimized FC-S_1 min	-92.7750308	-92.6753326
MS-CASPT2 optimized FC-S_1 min	-92.8276924	-92.7182376
SA-CASSCF Optimized Pb-Cl S_0/S_1 MECI	-92.7201777	-92.689441
MS-CASPT2 Optimized Pb-Cl S_0/S_1 MECI	-92.7268742	-92.726182

Table 15-S3. The absolute energies of the FC geometry of the CP cluster model at the EOM-CCSD level of theory.

EOM-CCSD	S_0 Energy / a.u.	S_1 Energy / a.u.
Franck-Condon point	-93.30891191	-93.13320109

Optimized Geometries (In Angstrom)

FC cluster geometry

Cs	0.498073759	0.585691428	0.495232353
Cs	4.861922882	4.921801480	0.498018182
Cs	0.553721599	4.877690235	4.920997419
Cs	4.904574355	0.499847103	4.788441915
Cl	2.698034464	-0.182297802	2.636303747
Cl	2.663665804	2.769349818	-0.219230085
Cl	-0.189624424	2.728045589	2.717029380
Cl	5.601941865	2.708465633	2.638574579
Cl	2.719276120	5.608717819	2.726445829
Cl	2.740329707	2.678203231	5.570947303
Pb	2.705840826	2.718724975	2.683550168

SA-CASSCF Optimized S_0/S_1 MECI geometry

Cs	-2.271452652	-2.266124898	-2.258072482
Cs	5.377635292	5.378278544	0.638134724
Cs	0.659061739	5.370298250	5.371367971
Cs	5.371291137	0.658547178	5.373033579
Cl	3.034645881	0.197594674	3.038977575
Cl	3.046384029	3.045397100	0.195195327
Cl	0.198424647	3.032474678	3.037756147
Cl	5.614940851	2.934292020	2.934237799
Cl	2.933477494	5.614993243	2.933131460
Cl	2.926679438	2.926253250	5.627713429
Pb	2.808908261	2.807999298	2.808503528

MS-CASPT2 Optimized S_0/S_1 MECI geometry

Cs	-3.695725978	-3.691169534	-3.684222882
Cs	5.351721661	5.349659747	1.061666740
Cs	1.054012292	5.354354754	5.346241172
Cs	5.355594558	1.054886870	5.345456159
Cl	3.129425235	0.318069291	3.129806686
Cl	3.126916233	3.126272401	0.319011649
Cl	0.317975752	3.129107541	3.130176848

Cl	5.736760855	3.124308652	3.120522753
Cl	3.124634304	5.736258550	3.120311555
Cl	3.125473296	3.124952454	5.734564885
Pb	3.073198446	3.073294857	3.076448205

SA-CASSCF Optimized MECI-S_1 min geometry

Cs	-2.264356824	-2.259037064	-2.250977042
Cs	5.366096509	5.366655218	0.656721488
Cs	0.671242446	5.360451314	5.362490402
Cs	5.361416914	0.671033972	5.364001049
Cl	3.025056852	0.192780823	3.027008849
Cl	3.032830558	3.031710341	0.193002222
Cl	0.193508314	3.022972744	3.026019051
Cl	5.642788776	2.937919033	2.935367657
Cl	2.937174782	5.642552312	2.934313615
Cl	2.930137060	2.929778000	5.649522641
Pb	2.804100746	2.803186649	2.802509154

MS-CASPT2 Optimized FC-S_1 min geometry

Cs	-2.782542082	2.750466980	-0.735581208
Cs	3.599658650	1.586107643	-0.761669976
Cs	-0.572957374	-3.082244393	-2.394014071
Cs	-0.196089235	-1.140629333	3.757976420
Cl	0.441180191	2.454107662	-0.840750235
Cl	2.002445349	0.197555772	1.912923571
Cl	-0.479016827	-2.603035609	0.916303228
Cl	-1.801825286	0.905701307	1.926206327
Cl	1.689293854	-0.898643396	-1.813803409
Cl	-1.917114796	-0.228816393	-1.803520079
Pb	0.014706553	0.046446743	-0.145121235

SA-CASSCF Optimized Pb-Cl S_0/S_1 MECI geometry

Cs	0.724250271	0.597910508	0.978126826
Cs	4.561063125	5.366502436	0.168180780
Cs	0.654961480	4.212045520	4.784670063
Cs	3.188885760	-1.119025851	6.291652454
Cl	3.054642197	-0.74498570	3.067963918
Cl	3.524111704	2.315174721	0.153619218
Cl	-1.302677006	2.436266906	2.812937943
Cl	7.755030370	3.833766686	1.679855542
Cl	3.488484897	5.207674024	3.216856512
Cl	2.976579641	2.101497556	6.109762145
Pb	3.750014175	2.165525767	3.202151466

MS-CASPT2 Optimized Pb-Cl S_0/S_1 MECI geometry

Cs	1.089111573	0.633745412	1.047203691
Cs	3.780522496	5.033643146	0.437089473
Cs	1.079150818	4.147469620	4.773126554
Cs	2.306317154	-0.901396303	6.055355163
Cl	2.980825069	-0.752579767	3.186701810
Cl	3.868410098	2.021109389	0.189294464
Cl	-0.670150960	2.677819954	2.622255912
Cl	7.254488812	4.291949244	1.277951010
Cl	3.707649750	5.370838812	3.447002865
Cl	2.998654243	1.948371961	6.053097535
Pb	3.980365813	1.901381450	3.376699792

REFERENCES

REFERENCES

(1) Green, M. A.; Ho-Baillie, A.; Snaith, H. J. *Nat Photonics* **2014**, *8*, 506.

(2) Salim, T.; Sun, S. Y.; Abe, Y.; Krishna, A.; Grimsdale, A. C.; Lam, Y. M. *J Mater Chem A* **2015**, *3*, 8943.

(3) Zhao, Y. X.; Zhu, K. *J Phys Chem Lett* **2014**, *5*, 4175.

(4) Brittman, S.; Adhyaksa, G. W. P.; Garnett, E. C. *Mrs Commun* **2015**, *5*, 7.

(5) Poglitsch, A.; Weber, D. *J Chem Phys* **1987**, *87*, 6373.

(6) Mitzi, D. B.; Feild, C. A.; Harrison, W. T. A.; Guloy, A. M. *Nature* **1994**, *369*, 467.

(7) Mitzi, D. B.; Wang, S.; Feild, C. A.; Chess, C. A.; Guloy, A. M. *Science* **1995**, *267*, 1473.

(8) Kagan, C. R.; Mitzi, D. B.; Dimitrakopoulos, C. D. *Science* **1999**, *286*, 945.

(9) Kojima, A.; Teshima, K.; Shirai, Y.; Miyasaka, T. *J Am Chem Soc* **2009**, *131*, 6050.

(10) Kim, H. S.; Lee, C. R.; Im, J. H.; Lee, K. B.; Moehl, T.; Marchioro, A.; Moon, S. J.; Humphry-Baker, R.; Yum, J. H.; Moser, J. E.; Gratzel, M.; Park, N. G. Sci Rep-Uk 2012, 2.

(11) Lee, M. M.; Teuscher, J.; Miyasaka, T.; Murakami, T. N.; Snaith, H. J. Science 2012, 338, 643.

(12) Zhou, H. P.; Chen, Q.; Li, G.; Luo, S.; Song, T. B.; Duan, H. S.; Hong, Z. R.; You, J. B.; Liu, Y. S.; Yang, Y. Science 2014, 345, 542.

(13) Jeon, N. J.; Noh, J. H.; Yang, W. S.; Kim, Y. C.; Ryu, S.; Seo, J.; Seok, S. I. Nature 2015, 517, 476.

(14) Yang, W. S.; Noh, J. H.; Jeon, N. J.; Kim, Y. C.; Ryu, S.; Seo, J.; Seok, S. I. Science 2015, 348, 1234.

(15) http://www.nrel.gov/ncpv/images/efficiency_chart.jpg. Research Cell Efficiency Records.

(16) Tidhar, Y.; Edri, E.; Weissman, H.; Zohar, D.; Hodes, G.; Cahen, D.; Rybtchinski, B.;

Kirmayer, S. J Am Chem Soc 2014, 136, 13249.

(17) Azpiroz, J. M.; Mosconi, E.; Bisquert, J.; De Angelis, F. *Energ Environ Sci* **2015**, *8*, 2118.

(18) Nazeeruddin, M. K.; Snaith, H. *Mrs Bull* **2015**, *40*, 641.

(19) Eperon, G. E.; Stranks, S. D.; Menelaou, C.; Johnston, M. B.; Herz, L. M.; Snaith, H. J. Energ Environ Sci 2014, 7, 982.

(20) Koh, T. M.; Fu, K. W.; Fang, Y. N.; Chen, S.; Sum, T. C.; Mathews, N.; Mhaisalkar, S. G.; Boix, P. P.; Baikie, T. J Phys Chem C 2014, 118, 16458.

(21) Eperon, G. E.; Bryant, D.; Troughton, J.; Stranks, S. D.; Johnston, M. B.; Watson, T.; Worsley, D. A.; Snaith, H. J. J Phys Chem Lett 2015, 6, 129.

(22) Stoumpos, C. C.; Kanatzidis, M. G. *Accounts Chem Res* **2015**, *48*, 2791.

(23) Etgar, L.; Gao, P.; Xue, Z. S.; Peng, Q.; Chandiran, A. K.; Liu, B.; Nazeeruddin, M. K.; Gratzel, M. J Am Chem Soc 2012, 134, 17396.

(24) Jeng, J. Y.; Chiang, Y. F.; Lee, M. H.; Peng, S. R.; Guo, T. F.; Chen, P.; Wen, T. C. Adv Mater 2013, 25, 3727.

(25) Kim, H. S.; Lee, J. W.; Yantara, N.; Boix, P. P.; Kulkarni, S. A.; Mhaisalkar, S.; Gratzel, M.; Park, N. G. Nano Lett 2013, 13, 2412.

(26) Kim, H. S.; Park, N. G. *J Phys Chem Lett* **2014**, *5*, 2927.

(27) Umari, P.; Mosconi, E.; De Angelis, F. *Sci Rep-Uk* **2014**, *4*.

(28) Zhao, Y.; Zhu, K. *The Journal of Physical Chemistry C* **2014**, *118*, 9412.

(29) Saliba, M.; Matsui, T.; Seo, J.-Y.; Domanski, K.; Correa-Baena, J.-P.; Nazeeruddin, M. K.; Zakeeruddin, S. M.; Tress, W.; Abate, A.; Hagfeldt, A. Energ Environ Sci 2016.

(30) Lin, Q. Q.; Armin, A.; Nagiri, R. C. R.; Burn, P. L.; Meredith, P. *Nat Photonics* **2015**, *9*, 106.

(31) Niu, G. D.; Guo, X. D.; Wang, L. D. *J Mater Chem A* **2015**, *3*, 8970.

(32) Noel, N. K.; Stranks, S. D.; Abate, A.; Wehrenfennig, C.; Guarnera, S.; Haghighirad, A.

A.; Sadhanala, A.; Eperon, G. E.; Pathak, S. K.; Johnston, M. B.; Petrozza, A.; Herz, L. M.; Snaith, H. J. Energ Environ Sci 2014, 7, 3061.

(33) Hao, F.; Stoumpos, C. C.; Cao, D. H.; Chang, R. P.; Kanatzidis, M. G. Nat Photonics 2014, 8, 489.

(34) Kumar, M. H.; Dharani, S.; Leong, W. L.; Boix, P. P.; Prabhakar, R. R.; Baikie, T.; Shi, C.; Ding, H.; Ramesh, R.; Asta, M.; Graetzel, M.; Mhaisalkar, S. G.; Mathews, N. Adv Mater 2014, 26, 7122.

(35) Borriello, I.; Cantele, G.; Ninno, D. *Phys Rev B* **2008**, *77*.

(36) Tang, L. C.; Chang, Y. C.; Huang, J. Y.; Lee, M. H.; Chang, C. S. *Jpn J Appl Phys* **2009**, *48*.

(37) Huang, L. Y.; Lambrecht, W. R. L. *Phys Rev B* **2013**, *88*.

(38) Stoumpos, C. C.; Frazer, L.; Clark, D. J.; Kim, Y. S.; Rhim, S. H.; Freeman, A. J.; Ketterson, J. B.; Jang, J. I.; Kanatzidis, M. G. J Am Chem Soc 2015, 137, 6804.

(39) Brivio, F.; Butler, K. T.; Walsh, A.; van Schilfgaarde, M. *Phys Rev B* **2014**, *89*.

(40) Yin, W. J.; Shi, T. T.; Yan, Y. F. *Appl Phys Lett* **2014**, *104*.

(41) Filippetti, A.; Mattoni, A. *Phys Rev B* **2014**, *89*.

(42) Stoumpos, C. C.; Malliakas, C. D.; Kanatzidis, M. G. *Inorg Chem* **2013**, *52*, 9019.

(43) Noh, J. H.; Im, S. H.; Heo, J. H.; Mandal, T. N.; Seok, S. I. *Nano Lett* **2013**, *13*, 1764.

(44) Suarez, B.; Gonzalez-Pedro, V.; Ripolles, T. S.; Sanchez, R. S.; Otero, L.; Mora-Sero, I. J Phys Chem Lett 2014, 5, 1628.

(45) Kulkarni, S. A.; Baikie, T.; Boix, P. P.; Yantara, N.; Mathews, N.; Mhaisalkar, S. J Mater Chem A 2014, 2, 9221.

(46) Edri, E.; Kirmayer, S.; Cahen, D.; Hodes, G. *J Phys Chem Lett* **2013**, *4*, 897.

(47) Brenner, T. M.; Egger, D. A.; Kronik, L.; Hodes, G.; Cahen, D. Nature Reviews Materials 2016, 1, 15007.

(48) Even, J.; Pedesseau, L.; Dupertuis, M.-A.; Jancu, J.-M.; Katan, C. Phys Rev B 2012, 86,

205301.

(49) Yin, W. J.; Yang, J. H.; Kang, J.; Yan, Y. F.; Wei, S. H. *J Mater Chem A* **2015**, *3*, 8926.

(50) Colella, S.; Mosconi, E.; Fedeli, P.; Listorti, A.; Gazza, F.; Orlandi, F.; Ferro, P.; Besagni, T.; Rizzo, A.; Calestani, G. *Chem Mater* **2013**, *25*, 4613.

(51) Yin, W. J.; Shi, T. T.; Yan, Y. F. *Adv Mater* **2014**, *26*, 4653.

(52) Amat, A.; Mosconi, E.; Ronca, E.; Quarti, C.; Umari, P.; Nazeeruddin, M. K.; Gratzel, M.; De Angelis, F. Nano Lett 2014, 14, 3608.

(53) Pellet, N.; Gao, P.; Gregori, G.; Yang, T. Y.; Nazeeruddin, M. K.; Maier, J.; Gratzel, M. Angew Chem Int Edit 2014, 53, 3151.

(54) Yamada, K.; Nakada, K.; Takeuchi, Y.; Nawa, K.; Yamane, Y. B Chem Soc Jpn 2011, 84, 926.

(55) Comin, R.; Walters, G.; Thibau, E. S.; Voznyy, O.; Lu, Z. H.; Sargent, E. H. J Mater Chem C 2015, 3, 9480.

(56) Saba, M.; Quochi, F.; Mura, A.; Bongiovanni, G. *Accounts Chem Res* **2016**, *49*, 166.

(57) Jung, H. S.; Park, N. G. *Small* **2015**, *11*, 10.

(58) Butler, K. T.; Frost, J. M.; Walsh, A. *Mater Horiz* **2015**, *2*, 228.

(59) Ball, J. M.; Lee, M. M.; Hey, A.; Snaith, H. J. *Energ Environ Sci* **2013**, *6*, 1739.

(60) Dong, Q. F.; Fang, Y. J.; Shao, Y. C.; Mulligan, P.; Qiu, J.; Cao, L.; Huang, J. S. Science 2015, 347, 967.

(61) Yang, B.; Dyck, O.; Poplawsky, J.; Keum, J.; Puretzky, A.; Das, S.; Ivanov, I.; Rouleau, C.; Duscher, G.; Geohegan, D.; Xiao, K. J Am Chem Soc 2015, 137, 9210.

(62) Burschka, J.; Pellet, N.; Moon, S. J.; Humphry-Baker, R.; Gao, P.; Nazeeruddin, M. K.; Gratzel, M. Nature 2013, 499, 316.

(63) Xiao, Z. G.; Dong, Q. F.; Bi, C.; Shao, Y. C.; Yuan, Y. B.; Huang, J. S. Adv Mater 2014, 26, 6503.

(64) Ponseca, C. S.; Savenije, T. J.; Abdellah, M.; Zheng, K. B.; Yartsev, A.; Pascher, T.; Harlang, T.; Chabera, P.; Pullerits, T.; Stepanov, A.; Wolf, J. P.; Sundstrom, V. J Am Chem Soc 2014, 136, 5189.

(65) Stranks, S. D.; Burlakov, V. M.; Leijtens, T.; Ball, J. M.; Goriely, A.; Snaith, H. J. Phys Rev Appl 2014, 2.

(66) Shi, D.; Adinolfi, V.; Comin, R.; Yuan, M. J.; Alarousu, E.; Buin, A.; Chen, Y.; Hoogland, S.; Rothenberger, A.; Katsiev, K.; Losovyj, Y.; Zhang, X.; Dowben, P. A.; Mohammed, O. F.; Sargent, E. H.; Bakr, O. M. Science 2015, 347, 519.

(67) Shu, Y. A.; Levine, B. G. *J Chem Phys* **2013**, *139*.

(68) Shu, Y. A.; Levine, B. G. *J Phys Chem C* **2014**, *118*, 7669.

(69) Shu, Y. N.; Levine, B. G. *J Phys Chem C* **2015**, *119*, 1737.

(70) Shu, Y. N.; Fales, B. S.; Levine, B. G. *Nano Lett* **2015**, *15*, 6247.

(71) Yarkony, D. R. *Accounts Chem Res* **1998**, *31*, 511.

(72) Levine, B. G.; Martinez, T. J. *Annu Rev Phys Chem* **2007**, *58*, 613.

(73) Marchioro, A.; Teuscher, J.; Friedrich, D.; Kunst, M.; van de Krol, R.; Moehl, T.; Gratzel, M.; Moser, J. E. Nat Photonics 2014, 8, 250.

(74) Yamada, Y.; Nakamura, T.; Endo, M.; Wakamiya, A.; Kanemitsu, Y. J Am Chem Soc 2014, 136, 11610.

(75) Ben-Nun, M.; Quenneville, J.; Martinez, T. J. *J Phys Chem A* **2000**, *104*, 5161.

(76) Levine, B. G.; Coe, J. D.; Virshup, A. M.; Martinez, T. J. *Chemical Physics* **2008**, *347*, 3.

(77) Heller, E. J. *J Chem Phys* **1981**, *75*, 2923.

(78) Roos, B. O.; Taylor, P. R. *Chemical Physics* **1980**, *48*, 157.

(79) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J Am Chem Soc* **1980**, *102*, 939.

(80) Stevens, W. J.; Krauss, M.; Basch, H.; Jasien, P. G. *Can J Chem* **1992**, *70*, 612.

(81) Stevens, W. J.; Basch, H.; Krauss, M. *J Chem Phys* **1984**, *81*, 6026.

(82) Stanton, J. F.; Bartlett, R. J. *The Journal of chemical physics* **1993**, *98*, 7029.

(83) Labello, N. P.; Ferreira, A. M.; Kurtz, H. A. *J Comput Chem* **2005**, *26*, 1464.

(84) Labello, N. P.; Ferreira, A. M.; Kurtz, H. A. *Int J Quantum Chem* **2006**, *106*, 3140.

(85) Becke, A. D. *Journal of Chemical Physics* **1993**, *98*, 1372.

(86) Lee, C. T.; Yang, W. T.; Parr, R. G. *Physical Review B* **1988**, *37*, 785.

(87) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Journal of Physical Chemistry 1994, 98, 11623.

(88) Meek, G. A.; Levine, B. G. *J Phys Chem Lett* **2014**, *5*, 2351.

(89) Thompson, A. L.; Punwong, C.; Martinez, T. J. *Chem Phys* **2010**, *370*, 70.

(90) Finley, J.; Malmqvist, P. A.; Roos, B. O.; Serrano-Andres, L. Chem Phys Lett 1998, 288, 299.

(91) Levine, B. G.; Coe, J. D.; Martinez, T. J. *Journal of Physical Chemistry B* **2008**, *112*, 405.

(92) Werner, H. J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Wiley Interdisciplinary Reviews: Computational Molecular Science 2012, 2, 242.

(93) Eckert, F.; Pulay, P.; Werner, H. J. *J Comput Chem* **1997**, *18*, 1473.

(94)Werner, H.; Knowles, P.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G. There is no corresponding record for this reference 2010.

(95) Knowles, P. J.; Werner, H. J. *Theor Chim Acta* **1992**, *84*, 95.

(96) Shiozaki, T.; Gyorffy, W.; Celani, P.; Werner, H. J. *J Chem Phys* **2011**, *135*.

(97) Hampel, C.; Peterson, K. A.; Werner, H. J. *Chem Phys Lett* **1992**, *190*, 1.

(98) Werner, H. J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schutz, M. Wires Comput Mol Sci 2012, 2, 242.

(99) Celani, P.; Werner, H. J. *J Chem Phys* **2000**, *112*, 5546.

(100) Korona, T.; Werner, H. J. *J Chem Phys* **2003**, *118*, 3006.

(101) Werner, H. J.; Knowles, P. J. *The Journal of chemical physics* **1985**, *82*, 5053.

(102) Woodward, P. M. *Acta Crystallogr B* **1997**, *53*, 32.

(103) Choi, H.; Jeong, J.; Kim, H. B.; Kim, S.; Walker, B.; Kim, G. H.; Kim, J. Y. Nano Energy 2014, 7, 80.

(104) Lee, J. W.; Kim, D. H.; Kim, H. S.; Seo, S. W.; Cho, S. M.; Park, N. G. Adv Energy Mater 2015, 5.

(105) Yi, C. Y.; Luo, J. S.; Meloni, S.; Boziki, A.; Ashari-Astani, N.; Gratzel, C.; Zakeeruddin, S. M.; Rothlisberger, U.; Gratzel, M. Energ Environ Sci 2016, 9, 656.

(106) McMeekin, D. P.; Sadoughi, G.; Rehman, W.; Eperon, G. E.; Saliba, M.; Horantner, M. T.; Haghighirad, A.; Sakai, N.; Korte, L.; Rech, B.; Johnston, M. B.; Herz, L. M.; Snaith, H. J. Science 2016, 351, 151.

(107) Li, Z.; Yang, M. J.; Park, J. S.; Wei, S. H.; Berry, J. J.; Zhu, K. *Chem Mater* **2016**, *28*, 284.

(108) Lang, L.; Yang, J. H.; Liu, H. R.; Xiang, H. J.; Gong, X. G. *Phys Lett A* **2014**, *378*, 290.

(109) Murtaza, G.; Ahmad, I. *Physica B: Condensed Matter* **2011**, *406*, 3222.

(110) Hoffman, R. “Solids and surfaces: a chemist's view of bonding in extended structures,” DTIC Document, 1988.

Chapter 16

New Paradigm for Non-Radiative Recombination Mechanism in Optoelectronic Materials

"All science is either physics or stamp collecting"

-Ernest Rutherford (1871-1937)

Being one of the most important concerns of the optoelectronic materials, the efficiency of the material is decided by a series complex photophysical processes including absorption, charge separation, and radiative/non-radiative decay. It is hence significant to obtain a comprehensive understanding of the non-radiative recombination mechanism. Early stages of illustrating non-radiative recombination process employs Fermi's golden rule (FGR)[1-3] based on the static picture without considering the nuclear motion as shown in Figure16-1(a). Defects induced in the materials are treated as a trapped state inserted between valence band maximum (VBM) and conduction band minimum (CBM) which is based on the pioneering works by Shockley, Read and Hall (SRH).[4,5] However, the SRH did not describe any detailed photophysical mechanism of the process itself. Various attempts have been tried in understanding the underlying mechanisms.[6-8] The currently widely accepted one is the multiphonon emission (MPE) mechanism.[7,9] MPE mechanism is based on the lattice relaxation and involves the breakdown of the Born-

Oppenheimer approximation as shown in Figure 16-2. And hence nuclear motion is being considered in the MPE mechanism. The non-radiative recombination mechanism proposed by us has the similar spirit as MPE, as our model involves a successive processes (as shown in Figure 16-1(b)) of exciton generation at quantum confined delocalized states, exciton localization into defect state and non-radiative recombination if possible. The differences between our model and MPE lie at, 1). The importance of the energy level of conical intersection (CI) has not been illustrated in MPE mechanism, 2). The MPE mechanism involves multiphonon releasing, while in our model, the relaxation involves localized nuclear deformation, despite the existing of defect or not, 3). Thanks to the advances in electronic structure theories, the current model employs computations of the potential energy surfaces (PESs) directly through *ab initio* electronic structure theories as a function of nuclear coordinates. While in MPE, the potential well (defined in the original paper[8] as the sum of electronic and lattice vibration energies) is a function of lattice coordinate which is approximated as a quadratic form, 4). MPE utilizes the single electron energies while our model employs many electron energies which accounts for the strong electron correlation near the intersection point.

In Chapter 7-15, we have investigated excited state dynamics of various optoelectronic materials to understand the possible mechanisms for non-radiative recombination process. They are generally in three categories, silicon nanocrystal (SiNC) with oxygen containing defects, tantalum nitride (Ta_3N_5) and lead halide perovskites (LHP). These materials are some of the representatives in the optoelectronic material world as

silicon is the current main component of photovoltaic cells (PV cells), Ta_3N_5 is becoming a promising photocatalyst in water splitting and LHP has been the one with most rapid efficiency increasing solar cell photosensitizer in the recent years. In this chapter, I will summarize our reported results to see how our proposed general mechanism of non-radiative recombination for optoelectronic materials can be applied to explain experimental observations with some of them have been in controversial for long time.

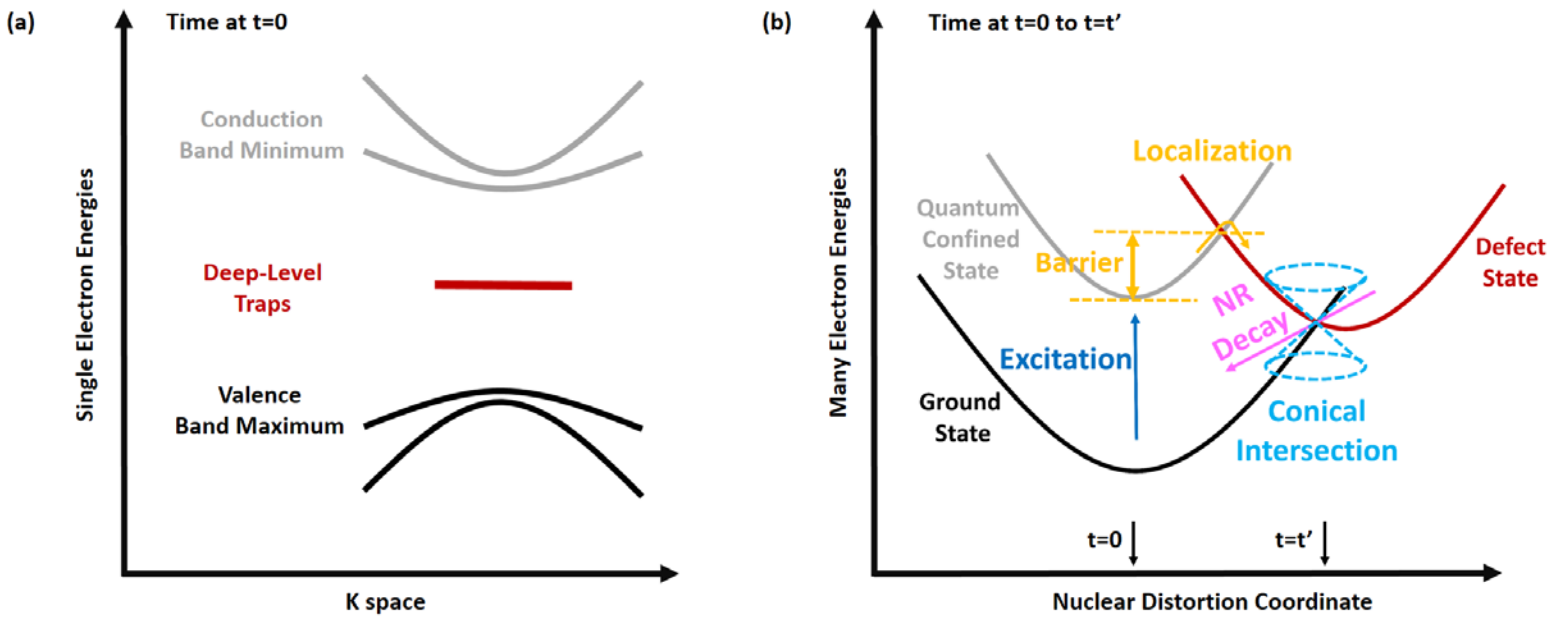


Figure 16-1. (a) Band structure at a fixed time t=0, with valence bands, conduction bands and the trapped state shown in black, grey and red respectively. (b) Potential energy surface as a function of nuclear distortion with ground state, quantum confined excited state and defect state shown in black, grey and red respectively. The nuclear distortion is also a function of time. From time t=0 to t=t', it corresponds to a series processes of generated exciton localize to a defect state and go through non-radiative decay to ground state through the conical intersection between the ground state and defect state.

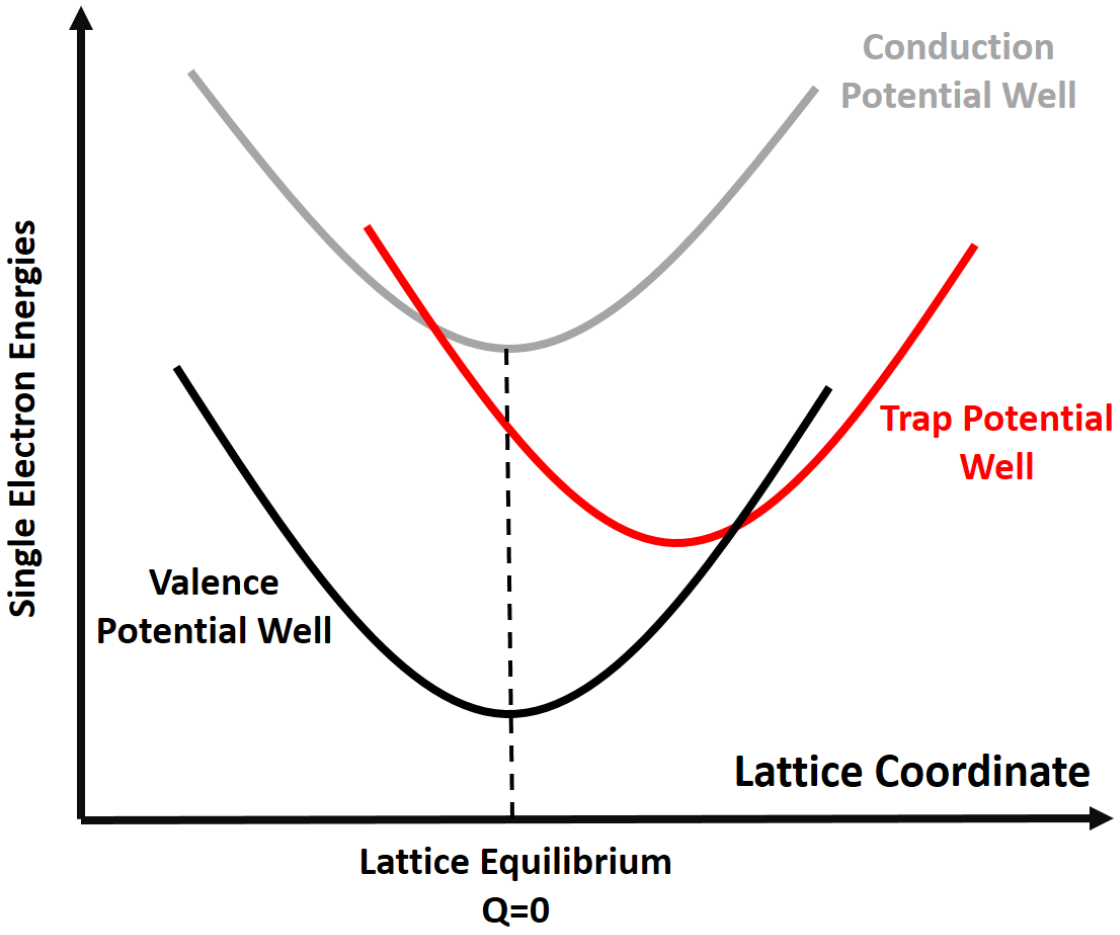


Figure 16-2. An energy level diagram shows the MPE mechanism. The valence potential well, conduction potential well and the trap potential well are shown in black, grey and red respectively. The potential well is defined as a sum of electronic and lattice vibration energies. The potential wells are plotted as a function of lattice coordinate.

16.1 Conical Intersection Facilitates the Non-Radiative Recombination in Optoelectronic Materials

Conical intersections, points of degeneracy between the potential energy surfaces (PESs) of different electronic states, facilitate the non-radiative decay process in molecular systems are well known.[10-12] Defects induced in nanocrystals (NC) has the similar characteristics as molecules, as they involve localized electronic structures with similar amounts of energy to be excited. Figure 16-3(a) shows the excitation energies of both quantum confined delocalized first excited state and defect localized first excited state as a function of cluster size. The two types of states are calculated separately at the optimized cluster geometry with or without defect. Geometries are optimized at density functional theory (DFT) with CAM-B3LYP[13] functional and STO-3G basis set. The excited states are calculated with the time-dependent DFT (TDDFT)[14,15] and 6-31G* basis set. The associated excited state character is described in Figure 16-3(b). We see that as the cluster size increases from 10 to 269 silicon atoms, the quantum confined state is dramatically affected by the size of the cluster, changing from 6.55 to 3.55 eV. However, the defect localized state has localized electronic structures and is insensitive to the size effect. The excitation energy of the defect localized state changes from 3.51 to 3.16 eV. This little difference may be attributed to the change of the dielectric environment when the cluster size changes. Hence, due to this localized electronic structure, it is very reasonable to believe the possibility that defect will undergo localized nuclear deformation to reach CI geometry once the excited electron localized to the defect state.

Based on this assumption, we have investigated various materials to prove this possibility as will be discussed in the next section. The follow-up researches in our group found that this localized nuclear deformation driven by the excited state relaxation does not only happen when a defect is induced, it also happens for pristine pure cluster like silicon nanocrystal (SiNC) capped with hydrogen atoms. In addition, a real physical nanocrystal (NC) sample synthesized from the experiment is an ensemble with finite temperature. Hence the thermodynamic fluctuation cannot be prevented. And therefore, NC with high symmetry does not really exist in experiment. This suggests the existence of excited state with localized electronic structure due to the distortion of the nuclear geometry from high symmetry and can be treated as a type of defect as well. Combined with both facts, it is believed by us that localized nuclear deformation is a general process once the NC is excited, and if the conical intersection is accessible, efficient non-radiative recombination is expected.

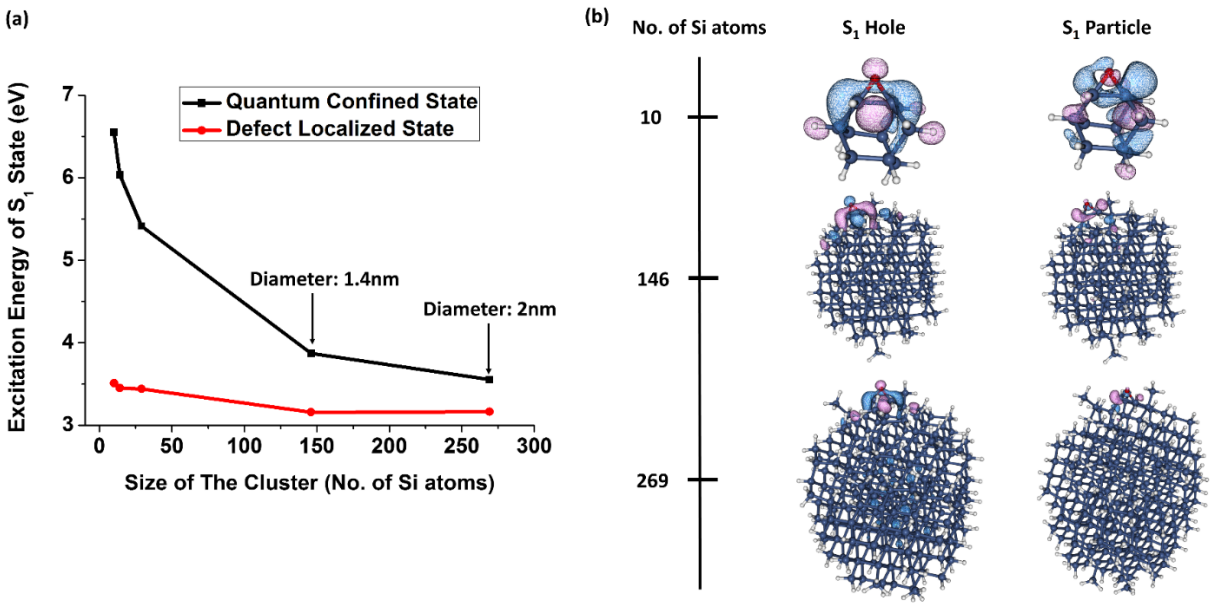


Figure 16-3. (a) First excited state of SiNC and SiNC with epoxide defect as a function of different cluster sizes represented by black and red respectively. These two states are representative state for quantum confined delocalized state and localized defect state. (b) The corresponding excitation characters of SiNC with epoxide defect with 10, 146 and 269 silicon atoms cluster size. It is clear that the excitation characters are localized at the defect. However, as can been seen from the hole of 269 silicon atom cluster, the electronic structure of the hole has a mixing character involving epoxide defect and quantum confined state, this is due to the fact that as the cluster size increases, the excitation energy of quantum confined state reduces, and they are close enough to mix at the 269 silicon atom cluster. This is also the reason why the energy of defect localized state of 269 silicon atom cluster risen a little bit compared to 146 silicon atom cluster.

So, what are the requirements for the CI being accessible? We proposed at least the following three requirements should be fulfilled, (a) a CI is present between the ground and localized excited state, (b) the energy of that CI is below the vertical excitation energy of the lowest electronic excited state, and (c) there is not a large energetic barrier on the excited state PES preventing the access from the Franck-Condon point.

The requirement (b) is a key factor in the photophysical process and has important effect on experimental observation. Let us understand the importance of CI energy with an example. Assume a rapid relaxation of the exciton to the lowest quantum confined state according to Kasha's rule.[16] For example if the MECI energy level is high (e.g., 4 eV, corresponds to 310 nm), then visible range photophysical process may not be affected by such a MECI, due to the excitons locate on the first quantum confined state do not have enough energy to reach the MECI, and hence the high internal quantum efficiency is expected. If the MECI energy level is in the visible range (e.g., 2 eV, 600 nm), it can be deduced that those excitons with higher energy than 2 eV, will likely be quenched through MECI and those excitons with energy far lower than 2 eV will be emissive.

The barrier in requirement (c) corresponds to the difference between the quantum confined state minimum and the crossing of quantum confined and defect localized states as shown in Figure 16-1(b). There exist two ways that we can approximate it. The first one is to utilize the Marcus theory[17,18] and approximate the quadratic form of the quantum confined state PES near equilibrium geometry, this is how we worked out the activation barrier as a function of band gap previously.[19] Alternatively, it is possible to

adopt a lower level of theory to approximate it from *ab* initio calculations.

16.2 Some Successful Applications of Localized Nuclear Deformation Induced MECI Prompted Non-Radiative Recombination Mechanism

16.2.1 SiNC with oxygen containing defects

The research of SiNC is dated back to 1990s when Canham et al. [20] discovered its visible light range photoluminescence (PL). Since then, the applications of SiNC ranges from light emitting devices to photovoltaics. [21-28] Due to the quantum confinement effect, the PL peak can be tuned from bulk silicon band gap of 1.1 eV to approximately 5 eV by changing the NC size.[29,30] The pristine SiNC has near unity PL quantum yields, but the yields is significantly lowered upon oxidation.[31] Figure 16-4(b) shows the PL peaks of pristine and oxidized SiNC as a function of particle size (the figure is reproduced from the experimental data of paper: MV Wolkin, J Jorne, PM Fauchet, G Allan, C Delerue, Phys. Rev. Lett. 82, 197). The PL energies of oxidized SiNCs with gaps larger than 2.1 eV shift to the red and appear to become insensitive to the average particle size as shown by the red shade in Figure 16-4(b).[29,32] The resulting orange PL is known as the slow (or S-) band PL of nanostructured silicon due to its microsecond lifetime.[33] It is known in the community that defect induced trap state provides pathways for non-radiative recombination.[5,34] However, it has been in long controversial in the community that which type of oxygen containing defect is responsible for the PL change.[29,35-47]

We have studied various oxygen containing defects from epoxides to silicon oxygen

double bond with the model SiNC cluster as has been discussed in detail in Chapter 7-13. We also refer the reader Chapter 1 for a detailed introduction. Combined with our model in understanding the non-radiative recombination for optoelectronic materials and the various energy levels of MECIs introduced by oxygen containing defects in SiNC as shown in Figure 16-4(a), we have concluded that the size insensitive PL peaks change is caused by the oxygen containing defect induced MECIs. There exists various oxygen containing defect MECIs as shown in 16-4(a) with their energy levels right around 2.1eV (considering with some errors). As has been discussed in section 16.1, those excitons in the ensemble have higher energy than MECIs will possibly be quenched leaving those lower energy excitons to emit.

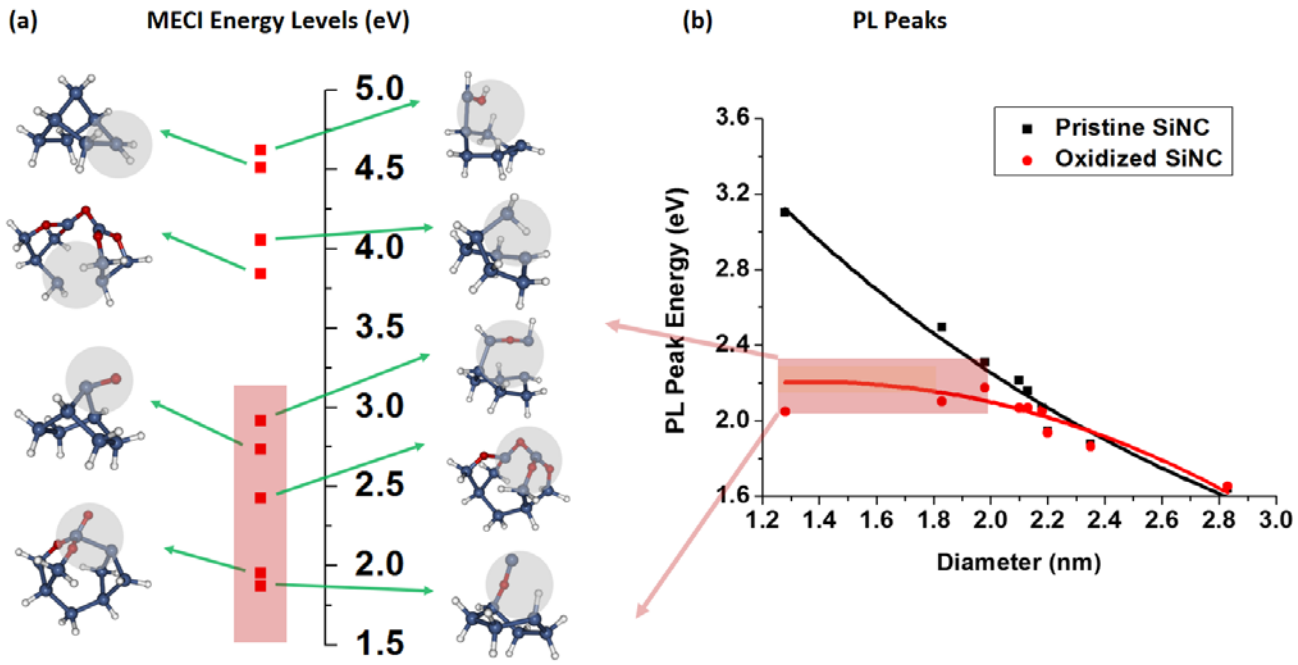


Figure 16-4. (a) Various MECIs energy levels of model silicon cluster with oxygen defects. With the green arrow points to the specific MECI structure and the grey shadow indicates the localized part where strong nuclear deformation is observed. The red rectangle shade under the energy level diagram corresponds the MECIs that are responsible for experimental observed PL peaks change after the SiNC is oxidized. (b) The PL peaks of pristine SiNC and oxidized SiNC as a function of SiNC size. The red shade close to the energy level of MECIs locate in (a). This explains why the excitons have energy level that is higher that

Figure 16-4(cont'd). ~2eV (as the red shade indicates) are quenched. (b) is reproduced from the experimental data of paper: Electronic states and luminescence in porous silicon quantum dots: the role of oxygen, MV Wolkin, J Jorne, PM Fauchet, G Allan, C Delerue, Phys. Rev. Lett. 82, 197.

16.2.2 Tantalum nitride

Tantalum nitride (Ta_3N_5) is a promising photocatalyst for water splitting with the visible light range band gap (2.1eV). [48-66] The successful reduction of the Ta_3N_5 band gap from tantalum oxide is due to the raising of the potential of N p orbitals compared to O p orbitals.[66,67] However, sacrificial reagent is required to prevent the charge recombination.[48,68-70] And to the best of our knowledge, the mechanism for the possible non-radiative recombination process has not been investigated.

We studied the non-radiative recombination process of Ta_3N_5 with our model cluster. The energy level of MECI locates around 1.7eV as compared with ground state minimum, which is right below the band gap of the bulk material. And our simulation suggests this MECI is a peaked MECI, which means efficient non-radiative recombination is expected for bulk materials.

16.2.3 Lead Halide perovskites

The application of organic-inorganic hybrid halide perovskite (HHP) material in solar cell photosensitizer is becoming an active field of research.[71] The efficiency of the perovskite solar cell increases from around 10% in 2012[72,73] to 22.1% in 2016,[74] which represents one of the

categories of photosensitizers with the fastest efficiency increasing rate. The attractive properties of perovskite materials come from three parts, 1). HHP is believed to be direct band gap semiconductor,[75-82] with band gap can be made in the visible range[72,83-87] or close to the 1.1-1.5 eV optimal band gap value based on Shockley-Queisser model,[88] 2). the optical properties and device stabilities are tunable by adopting different cations and anions into the fixed perovskite AMX_3 structure, with A is organic or inorganic cation, M is the metal cation, and X is the halide anion,[83,84,86,89-102] 3). Single crystal HHP has low recombination rate with nearly unity internal quantum efficiency[103-107] and high charge carrier mobility.[108-110]

We have studied the excited state dynamics of lead halide perovskite (LHP) material with our model cluster A_4PbX_6. Specifically, we optimized the MECI geometries and calculated their energy levels with respect to the ground state minimum. The MECI geometries have similar structures with breaking Pb-X bonds. Moreover, we found the MECI energy level largely depends on the halide, while organic or inorganic position A cation has little effect on the MECI energy level. The MECI energy levels of different LHP models we studied are summarized in Figure 16-5(a). The MECI geometries of A_4PbCl_6, A_4PbBr_6 and A_4PbI_6 have energy levels around 3.5, 3.0 and 2.0 eV respectively. Compared with the blue, green and red colors band gaps of $APbCl_3$, $APbBr_3$ and $APbI_3$ material respectively, [99-102] the MECI energy levels are right above the band gaps. This explains the near unity internal quantum efficiency of the LHP.

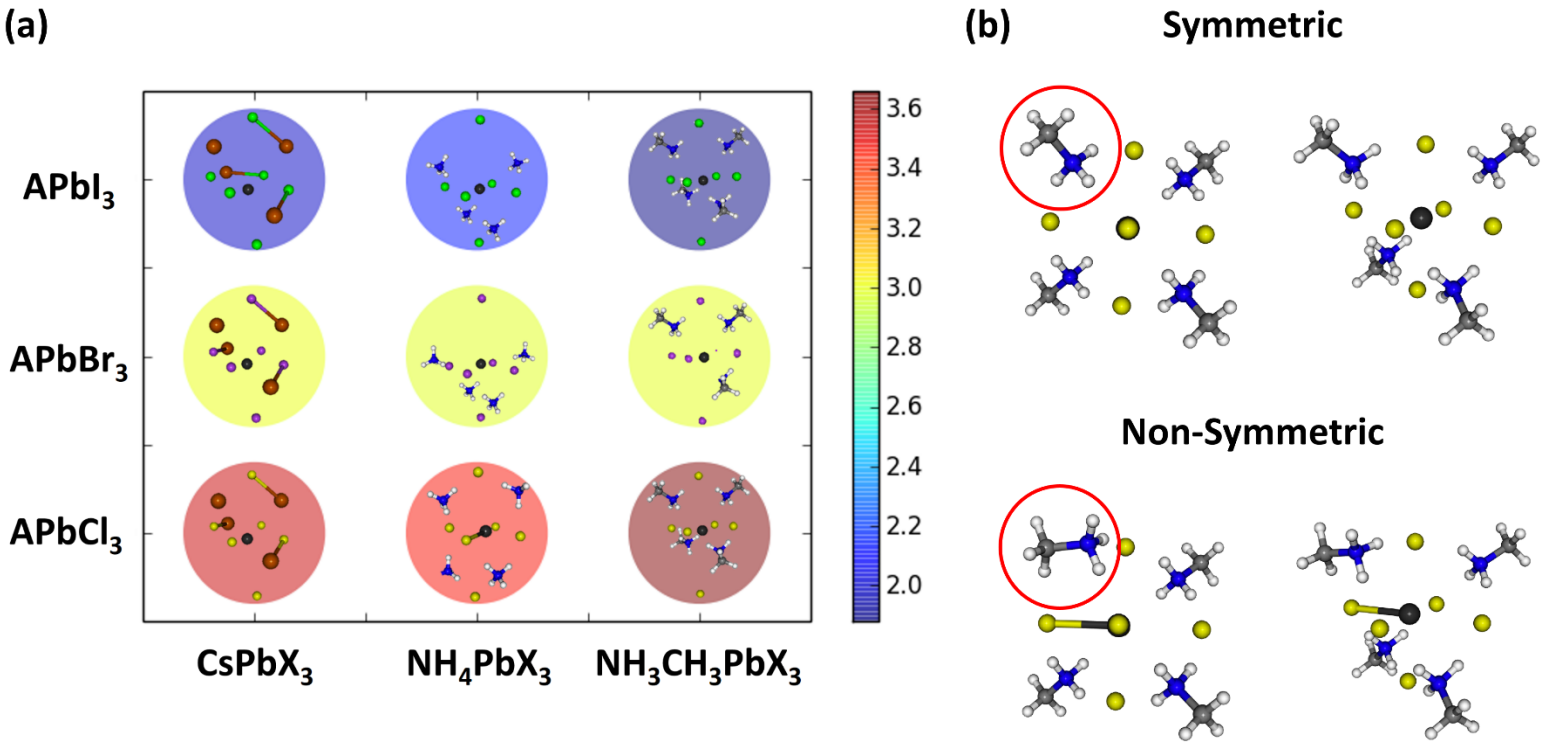


Figure 16-5. (a). A matrix showing MECI energy levels of $APbX_3$ perovskite material. With each row represents the same type of halide anion but different types of A position cation, each column represents the same type of A position cation but different types of halide anions. The color of the dot represents the energy levels of MECI geometries, and the insets show the MECI geometries. The right hand side color coordinate indicates the energy level in the unit of eV. (b). The ground state minimum

Figure 16-5 (cont'd). geometries of symmetric and non-symmetric conformations of $(NH_3CH_3)_4PbCl_6$ model cluster with two perspectives. The circled $NH_3CH_3^+$ cation indicates the one that has been rotated from symmetric conformation.

It is also noteworthy from our calculations that the conformational change of the ground state LHP with large A position cations, for example, $NH_3CH_3PbCl_3$, has little effect on the vertical excitation energy and MECI energy level. As we rotate one of the $NH_3CH_3^+$ cation in our $(NH_3CH_3)_4PbCl_6$ model cluster from a symmetric structure, the vertical excitation energy changes from 6.476 to 6.248 eV and MECI energy levels changes from 3.659 to 3.507 eV. The ground state minimum geometries of symmetric and non-symmetric $(NH_3CH_3)_4PbCl_6$ model clusters are shown in Figure 16-6(b).

We refer the reader Chapter 15 for a detailed description of the excited state dynamics of the two model clusters.

16.3 Conclusions and Outlook

The non-radiative recombination mechanism for optoelectronic materials is of great importance for solar cells and organic light emitting diodes. Both devices are becoming more and more complex in recent years, for example, multi-junction solar cells require photosensitizers with different band gaps to overcome the Shockley-Queisser limit and the current record keeper

is the four-junction solar cell with 46% efficiency. And hence a comprehensive understanding of the efficient limiting process will help us optimize the material efficiency systematically.

In the past five years, we have been dedicated to understanding the non-radiative recombination process by modeling the excited state dynamics of model systems. Based on the fact that MECI is usually achieved by local nuclear geometrical deformation of NC, the energy level of MECI should be size insensitive. Thus the energy level of MECI obtained from our model cluster should be comparable with that of real size NC. With the investigation of a couple of systems, we proposed a general non-radiative recombination mechanism for optoelectronic materials involving CIs. The MECI will dramatically affect the PL spectrum if it is energetically accessible by the excitons. And hence development of efficient solar cell photosensitizers should take the energy level of MECI into account as well.

The proposed mechanism has been employed in understanding the non-radiative decay processes for various materials including oxidized silicon, tantalum nitride, and perovskites. The current computational framework and the proposed mechanism of non-radiative recombination process in materials provides, 1) a routine in investigating the excited state dynamics of materials, 2) a general way to predict whether an efficient non-radiative recombination is achievable for materials upon excitation with the knowledge of band gap, 3) a new strategy in efficient optoelectronic materials design. Previous chapters including Chapter 7 to 15 have proved the usefulness of this computational framework.

The current studies are all based on modeling of the whole cluster up to nano-sized cluster with the benefits of new developments in multi-reference electronic structure theories,[111] efficient and accurate non-adiabatic dynamics,[112-114] and TeraChem graphics processing unit accelerated quantum chemistry software package.[115-119] From the electronic structure theory perspective, it will be important to investigate the proposed mechanism in periodic systems in the future, which requires incorporating the many-body wavefunction ideas into quasi-particle pictures (quasi-particles are counter parts of the orbitals in many-body wavefunction methods, while in many-body wavefunction ansatz, we have both orbitals constituted determinant and configuration interactions among multiple determinants, we refer the readers who are interested in a detailed discussion to Chapter 5 and 6). From the material design perspective, an automated optimization scheme for MECI should be developed to incorporate with the computational screening strategies such that material candidates can be screened through computational methods and provide promising candidates for experimental investigation (we suggest readers who are interested in the computational screening topic to Chapter 17 and 18).

REFERENCES

REFERENCES

(1) Bransden, B. H.; Joachain, C. J. *Quantum mechanics*; Pearson Education, 2000.

(2) Dirac, P. A. M. *Proceedings of the Royal Society of London Series a-Containing Papers of a Mathematical and Physical Character* **1927**, *114*, 243.

(3) Bixon, M.; Jortner, J. *Journal of Chemical Physics* **1968**, *48*, 715.

(4) Shockley, W.; Read, W. T. *Physical Review* **1952**, *87*, 835.

(5) Hall, R. N. *Physical Review* **1952**, *87*, 387.

(6) Lax, M. *Physical Review* **1960**, *119*, 1502.

(7) Lang, D. V.; Henry, C. H. *Bulletin of the American Physical Society* **1975**, *20*, 390.

(8) Henry, C. H.; Lang, D. V. *Physical Review B* **1977**, *15*, 989.

(9) Luque, A.; Marti, A.; Antolin, E.; Tablero, C. *Physica B-Condensed Matter* **2006**, *382*, 320.

(10) Bernardi, F.; Olivucci, M.; Robb, M. A. *Chemical Society Reviews* **1996**, *25*, 321.

(11) Yarkony, D. R. *Reviews of Modern Physics* **1996**, *68*, 985.

(12) Levine, B. G.; Martinez, T. J. *Annual Review of Physical Chemistry* **2007**, *58*, 613.

(13) Yanai, T.; Tew, D. P.; Handy, N. C. *Chemical Physics Letters* **2004**, *393*, 51.

(14) Petersilka, M.; Gossmann, U. J.; Gross, E. K. U. *Physical Review Letters* **1996**, *76*, 1212.

(15) Casida, M. E. *Recent Advances in Density Functional Methods:(Part I)* **1995**, *1*, 155.

(16) Kasha, M. *Discussions of the Faraday society* **1950**, *9*, 14.

(17) Marcus, R. A. *Journal of Chemical Physics* **1956**, *24*, 966.

(18) Barbara, P. F.; Meyer, T. J.; Ratner, M. A. *Journal of Physical Chemistry* **1996**, *100*, 13148.

(19) Shu, Y. N.; Fales, B. S.; Levine, B. G. *Nano Letters* **2015**, *15*, 6247.

(20) Canham, L. T. *Applied Physics Letters* **1990**, *57*, 1046.

(21) Erdem, T.; Demir, H. V. *Nature Photonics* **2011**, *5*, 126.

(22) Canham, L. *Nature* **2000**, *408*, 411.

(23) Walters, R. J.; Bourianoff, G. I.; Atwater, H. A. *Nature Materials* **2005**, *4*, 143.

(24) Cui, Y.; Wei, Q. Q.; Park, H. K.; Lieber, C. M. *Science* **2001**, *293*, 1289.

(25) He, Y.; Zhong, Y. L.; Peng, F.; Wei, X. P.; Su, Y. Y.; Lu, Y. M.; Su, S.; Gu, W.; Liao, L. S.; Lee, S. T. Journal of the American Chemical Society 2011, 133, 14192.

(26) Erogbogbo, F.; Yong, K. T.; Roy, I.; Hu, R.; Law, W. C.; Zhao, W. W.; Ding, H.; Wu, F.; Kumar, R.; Swihart, M. T.; Prasad, P. N. Acs Nano 2011, 5, 413.

(27) Pikulev, V.; Kuznetsov, S.; Saren, A.; Gardin, Y. E.; Gurtov, V. physica status solidi (a) 2007, 204, 1266.

(28) Cuony, P.; Alexander, D. T.; Perez-Wurfl, I.; Despeisse, M.; Bugnon, G.; Boccard, M.; Söderström, T.; Hessler-Wyser, A.; Hébert, C.; Ballif, C. Advanced Materials 2012, 24, 1182.

(29) Wolkin, M. V.; Jorne, J.; Fauchet, P. M.; Allan, G.; Delerue, C. Physical Review Letters 1999, 82, 197.

(30) Reboredo, F. A.; Franceschetti, A.; Zunger, A. *Physical Review B* **2000**, *61*, 13073.

(31) Jurbergs, D.; Rogojina, E.; Mangolini, L.; Kortshagen, U. *Appl Phys Lett* **2006**, *88*.

(32) von Behren, J.; van Buuren, T.; Zacharias, M.; Chimowitz, E. H.; Fauchet, P. M. *Solid State Communications* **1998**, *105*, 317.

(33) Garcia, C.; Garrido, B.; Pellegrino, P.; Ferre, R.; Moreno, J. A.; Morante, J. R.; Pavesi, L.; Cazzanelli, M. Applied Physics Letters 2003, 82, 1595.

(34) Shockley, W.; Read, W. T. *Physical Review* **1952**, *87*, 835.

(35) Puzder, A.; Williamson, A. J.; Grossman, J. C.; Galli, G. Journal of Chemical Physics 2002, 117, 6721.

(36) Vasiliev, I.; Chelikowsky, J. R.; Martin, R. M. *Physical Review B* **2002**, *65*.

(37) Puzder, A.; Williamson, A. J.; Grossman, J. C.; Galli, G. Journal of the American Chemical Society 2003, 125, 2786.

(38) Luppi, M.; Ossicini, S. *Physical Review B* **2005**, *71*.

(39) Chen, X. B.; Pi, X. D.; Yang, D. R. *Journal of Physical Chemistry C* **2010**, *114*, 8774.

(40) Pennycook, T. J.; Hadjisavvas, G.; Idrobo, J. C.; Kelires, P. C.; Pantelides, S. T. Physical Review B 2010, 82.

(41) Daldosso, N.; Luppi, M.; Ossicini, S.; Degoli, E.; Magri, R.; Dalba, G.; Fornasini, P.; Grisenti, R.; Rocca, F.; Pavesi, L.; Boninelli, S.; Priolo, F.; Spinella, C.; Iacona, F. Physical Review B 2003, 68.

(42) Ramos, L. E.; Furthmuller, J.; Bechstedt, F. *Physical Review B* **2004**, *70*.

(43) Ramos, L. E.; Furthmuller, J.; Bechstedt, F. *Physical Review B* **2005**, *71*.

(44) Guerra, R.; Degoli, E.; Ossicini, S. *Physical Review B* **2009**, *80*.

(45) Konig, D.; Rudd, J.; Green, M. A.; Conibeer, G. Solar Energy Materials and Solar Cells 2009, 93, 753.

(46) Seino, K.; Bechstedt, F.; Kroll, P. *Nanotechnology* **2009**, *20*.

(47) Li, T. S.; Gygi, F.; Galli, G. *Physical Review Letters* **2011**, *107*.

(48) Hitoki, G.; Ishikawa, A.; Takata, T.; Kondo, J. N.; Hara, M.; Domen, K. Chemistry Letters 2002, 736.

(49) Zhang, Q. H.; Gao, L. *Langmuir* **2004**, *20*, 9821.

(50) Li, Y. B.; Takata, T.; Cha, D.; Takanabe, K.; Minegishi, T.; Kubota, J.; Domen, K. Advanced Materials 2013, 25, 125.

(51) Ma, S. S. K.; Hisatomi, T.; Maeda, K.; Moriya, Y.; Domen, K. Journal of the American Chemical Society 2012, 134, 19993.

(52) Cong, Y. Q.; Park, H. S.; Wang, S. J.; Dang, H. X.; Fan, F. R. F.; Mullins, C. B.; Bard, A. J. Journal of Physical Chemistry C 2012, 116, 14541.

(53) Dabirian, A.; van de Krol, R. *Applied Physics Letters* **2013**, *102*.

(54) Ma, S. S. K.; Maeda, K.; Hisatomi, T.; Tabata, M.; Kudo, A.; Domen, K. Chemistry-a European Journal 2013, 19, 7480.

(55) Dang, H. X.; Hahn, N. T.; Park, H. S.; Bard, A. J.; Mullins, C. B. Journal of Physical Chemistry C 2012, 116, 19225.

(56) Nurlaela, E.; Ould-Chikh, S.; Harb, M.; del Gobbo, S.; Aouine, M.; Puzenat, E.; Sautet, P.; Domen, K.; Basset, J. M.; Takanabe, K. Chemistry of Materials 2014, 26, 4812.

(57) Hara, M.; Hitoki, G.; Takata, T.; Kondo, J. N.; Kobayashi, H.; Domen, K. Catalysis Today 2003, 78, 555.

(58) Liu, G.; Shi, J.; Zhang, F.; Chen, Z.; Han, J.; Ding, C.; Chen, S.; Wang, Z.; Han, H.; Li, C. Angewandte Chemie International Edition 2014, 53, 7295.

(59) Yuliati, L.; Yang, J.-H.; Wang, X.; Maeda, K.; Takata, T.; Antonietti, M.; Domen, K. Journal of Materials Chemistry 2010, 20, 4295.

(60) Cong, Y.; Park, H. S.; Dang, H. X.; Fan, F.-R. F.; Bard, A. J.; Mullins, C. B. Chemistry of Materials 2012, 24, 579.

(61) Ho, C. T.; Low, K. B.; Klie, R. F.; Maeda, K.; Domen, K.; Meyer, R. J.; Snee, P. T. Journal of Physical Chemistry C 2011, 115, 647.

(62)Ishikawa, A.; Takata, T.; Kondo, J. N.; Hara, M.; Domen, K. Journal of Physical Chemistry B 2004, 108, 11049.

(63) Tabata, M.; Maeda, K.; Higashi, M.; Lu, D.; Takata, T.; Abe, R.; Domen, K. Langmuir 2010, 26,

9161.

(64) Li, Y.; Zhang, L.; Torres-Pardo, A.; González-Calbet, J. M.; Ma, Y.; Oleynikov, P.; Terasaki, O.; Asahina, S.; Shima, M.; Cha, D. Nature Communications 2013, 4.

(65) Feng, X.; LaTempa, T. J.; Basham, J. I.; Mor, G. K.; Varghese, O. K.; Grimes, C. A. Nano Letters 2010, 10, 948.

(66) Chun, W. J.; Ishikawa, A.; Fujisawa, H.; Takata, T.; Kondo, J. N.; Hara, M.; Kawai, M.; Matsumoto, Y.; Domen, K. Journal of Physical Chemistry B 2003, 107, 1798.

(67) Maeda, K.; Domen, K. *Journal of Physical Chemistry C* **2007**, *111*, 7851.

(68) Li, Y.; Takata, T.; Cha, D.; Takanabe, K.; Minegishi, T.; Kubota, J.; Domen, K. Advanced Materials 2013, 25, 125.

(69) Yokoyama, D.; Hashiguchi, H.; Maeda, K.; Minegishi, T.; Takata, T.; Abe, R.; Kubota, J.; Domen, K. Thin Solid Films 2011, 519, 2087.

(70) Higashi, M.; Domen, K.; Abe, R. *Energy & Environmental Science* **2011**, *4*, 4138.

(71) Green, M. A.; Ho-Baillie, A.; Snaith, H. J. *Nature Photonics* **2014**, *8*, 506.

(72) Kim, H. S.; Lee, C. R.; Im, J. H.; Lee, K. B.; Moehl, T.; Marchioro, A.; Moon, S. J.; Humphry-Baker, R.; Yum, J. H.; Moser, J. E.; Gratzel, M.; Park, N. G. Scientific Reports 2012, 2.

(73) Lee, M. M.; Teuscher, J.; Miyasaka, T.; Murakami, T. N.; Snaith, H. J. *Science* **2012**, *338*, 643.

(74) http://www.nrel.gov/ncpv/images/efficiency_chart.jpg. Research Cell Efficiency Records.

(75) Borriello, I.; Cantele, G.; Ninno, D. *Physical Review B* **2008**, *77*.

(76) Tang, L. C.; Chang, Y. C.; Huang, J. Y.; Lee, M. H.; Chang, C. S. Japanese Journal of Applied Physics 2009, 48.

(77) Umari, P.; Mosconi, E.; De Angelis, F. *Scientific Reports* **2014**, *4*.

(78) Huang, L. Y.; Lambrecht, W. R. L. *Physical Review B* **2013**, *88*.

(79) Stoumpos, C. C.; Frazer, L.; Clark, D. J.; Kim, Y. S.; Rhim, S. H.; Freeman, A. J.; Ketterson, J. B.; Jang, J. I.; Kanatzidis, M. G. Journal of the American Chemical Society 2015, 137, 6804.

(80) Brivio, F.; Butler, K. T.; Walsh, A.; van Schilfgaarde, M. *Physical Review B* **2014**, *89*.

(81) Yin, W. J.; Shi, T. T.; Yan, Y. F. *Applied Physics Letters* **2014**, *104*.

(82) Filippetti, A.; Mattoni, A. *Physical Review B* **2014**, *89*.

(83) Stoumpos, C. C.; Malliakas, C. D.; Kanatzidis, M. G. *Inorganic Chemistry* **2013**, *52*, 9019.

(84) Noh, J. H.; Im, S. H.; Heo, J. H.; Mandal, T. N.; Seok, S. I. *Nano Letters* **2013**, *13*, 1764.

(85) Suarez, B.; Gonzalez-Pedro, V.; Ripolles, T. S.; Sanchez, R. S.; Otero, L.; Mora-Sero, I. Journal of Physical Chemistry Letters 2014, 5, 1628.

(86) Kulkarni, S. A.; Baikie, T.; Boix, P. P.; Yantara, N.; Mathews, N.; Mhaisalkar, S. J Mater Chem A 2014, 2, 9221.

(87) Edri, E.; Kirmayer, S.; Cahen, D.; Hodes, G. *Journal of Physical Chemistry Letters* **2013**, *4*, 897.

(88) Brenner, T. M.; Egger, D. A.; Kronik, L.; Hodes, G.; Cahen, D. *Nature Reviews Materials* **2016**, *1*, 15007.

(89) Eperon, G. E.; Stranks, S. D.; Menelaou, C.; Johnston, M. B.; Herz, L. M.; Snaith, H. J. Energy & Environmental Science 2014, 7, 982.

(90) Yin, W. J.; Yang, J. H.; Kang, J.; Yan, Y. F.; Wei, S. H. Journal of Materials Chemistry A 2015, 3, 8926.

(91) Colella, S.; Mosconi, E.; Fedeli, P.; Listorti, A.; Gazza, F.; Orlandi, F.; Ferro, P.; Besagni, T.; Rizzo, A.; Calestani, G. Chemistry of Materials 2013, 25, 4613.

(92) Yin, W. J.; Shi, T. T.; Yan, Y. F. *Advanced Materials* **2014**, *26*, 4653.

(93) Stoumpos, C. C.; Kanatzidis, M. G. *Accounts of Chemical Research* **2015**, *48*, 2791.

(94) Amat, A.; Mosconi, E.; Ronca, E.; Quarti, C.; Umari, P.; Nazeeruddin, M. K.; Gratzel, M.; De Angelis, F. Nano Letters 2014, 14, 3608.

(95) Pellet, N.; Gao, P.; Gregori, G.; Yang, T. Y.; Nazeeruddin, M. K.; Maier, J.; Gratzel, M. Angewandte Chemie-International Edition 2014, 53, 3151.

(96) Mitzi, D. B.; Feild, C. A.; Harrison, W. T. A.; Guloy, A. M. *Nature* **1994**, *369*, 467.

(97) Yamada, K.; Nakada, K.; Takeuchi, Y.; Nawa, K.; Yamane, Y. Bulletin of the Chemical Society of Japan 2011, 84, 926.

(98) Noel, N. K.; Stranks, S. D.; Abate, A.; Wehrenfennig, C.; Guarnera, S.; Haghighirad, A. A.; Sadhanala, A.; Eperon, G. E.; Pathak, S. K.; Johnston, M. B.; Petrozza, A.; Herz, L. M.; Snaith, H. J. Energy & Environmental Science 2014, 7, 3061.

(99) Comin, R.; Walters, G.; Thibau, E. S.; Voznyy, O.; Lu, Z. H.; Sargent, E. H. Journal of Materials Chemistry C 2015, 3, 9480.

(100) Saba, M.; Quochi, F.; Mura, A.; Bongiovanni, G. Accounts of Chemical Research 2016, 49, 166.

(101) Jung, H. S.; Park, N. G. *Small* **2015**, *11*, 10.

(102) Butler, K. T.; Frost, J. M.; Walsh, A. *Materials Horizons* **2015**, *2*, 228.

(103) Ball, J. M.; Lee, M. M.; Hey, A.; Snaith, H. J. *Energy & Environmental Science* **2013**, *6*, 1739.

(104) Dong, Q. F.; Fang, Y. J.; Shao, Y. C.; Mulligan, P.; Qiu, J.; Cao, L.; Huang, J. S. Science 2015, 347, 967.

(105) Yang, B.; Dyck, O.; Poplawsky, J.; Keum, J.; Puretzky, A.; Das, S.; Ivanov, I.; Rouleau, C.; Duscher, G.; Geohegan, D.; Xiao, K. Journal of the American Chemical Society 2015, 137, 9210.

(106) Burschka, J.; Pellet, N.; Moon, S. J.; Humphry-Baker, R.; Gao, P.; Nazeeruddin, M. K.; Gratzel, M. Nature 2013, 499, 316.

(107) Xiao, Z. G.; Dong, Q. F.; Bi, C.; Shao, Y. C.; Yuan, Y. B.; Huang, J. S. Advanced Materials 2014, 26, 6503.

(108) Ponseca, C. S.; Savenije, T. J.; Abdellah, M.; Zheng, K. B.; Yartsev, A.; Pascher, T.; Harlang, T.; Chabera, P.; Pullerits, T.; Stepanov, A.; Wolf, J. P.; Sundstrom, V. Journal of the American Chemical Society 2014, 136, 5189.

(109) Stranks, S. D.; Burlakov, V. M.; Leijtens, T.; Ball, J. M.; Goriely, A.; Snaith, H. J. Physical Review Applied 2014, 2.

(110) Shi, D.; Adinolfi, V.; Comin, R.; Yuan, M. J.; Alarousu, E.; Buin, A.; Chen, Y.; Hoogland, S.; Rothenberger, A.; Katsiev, K.; Losovyj, Y.; Zhang, X.; Dowben, P. A.; Mohammed, O. F.; Sargent, E. H.; Bakr, O. M. Science 2015, 347, 519.

(111) Shu, Y. A.; Hohenstein, E. G.; Levine, B. G. *Journal of Chemical Physics* **2015**, *142*.

(112) Levine, B. G.; Coe, J. D.; Virshup, A. M.; Martinez, T. J. *Chem Phys* **2008**, *347*, 3.

(113) Ben-Nun, M.; Quenneville, J.; Martinez, T. J. *Journal of Physical Chemistry A* **2000**, *104*, 5161.

(114) Meek, G. A.; Levine, B. G. *Journal of Physical Chemistry Letters* **2014**, *5*, 2351.

(115) Ufimtsev, I. S.; Martinez, T. J. *Journal of Chemical Theory and Computation* **2008**, *4*, 222.

(116) Ufimtsev, I. S.; Martinez, T. J. *Journal of Chemical Theory and Computation* **2009**, *5*, 3138.

(117) Ufimtsev, I. S.; Martinez, T. J. *Journal of Chemical Theory and Computation* **2009**, *5*, 2619.

(118) Isborn, C. M.; Luehr, N.; Ufimtsev, I. S.; Martinez, T. J. *Journal of Chemical Theory and Computation* **2011**, *7*, 1814.

(119) Fales, B. S.; Levine, B. G. *Journal of Chemical Theory and Computation* **2015**, *11*, 4708.

Chapter 17

Simulated Evolution of Fluorophores for Light Emitting Diodes

"A very great deal more truth can become known than can be proven."

-Richard P. Feynman (1918-1988)

This chapter is reproduced from paper: Simulated evolution of fluorophores for light emitting diodes, Y. Shu and B. G. Levine, J. Chem. Phys., 142, 104104 (2015).

Organic light-emitting diodes (OLEDs) are a sustainable, low cost light source used in displays and other applications.[1] Unfortunately, simple fluorescence-based first generation OLEDs are limited to 25% internal quantum efficiency (IQE) because random pairing of electrons and holes results in three times more triplet excitations than singlet ones.[2] However, one can circumvent this limitation by developing schemes in which the formation of triplet states results in efficient emission as well. For example, by introducing heavy atoms with strong spin-orbit coupling, phosphorescent second generation OLEDs achieve nearly 100% IQE and have external quantum efficiencies (EQEs) as high as 20-30%.[3-9] Copper (I) organometallic compounds have been demonstrated to achieve 16% EQE via a process known as thermally activated (or E-type) delayed

fluorescence (TADF),[10-14] and similarly impressive results have recently been achieved in metal-free systems by the Adachi group and others, with IQEs up to 100% and EQEs up to 25% reported.[15-22] Thus, the development of fluorophores capable of TADF is an active and exciting area of research.

TADF proceeds as follows: As noted above, roughly 75% of excitons formed in an OLED are spin triplets. Upon relaxation, these states exist in a relatively dark first triplet excited state, T_1. If the lowest singlet excited state, S_1, is only slightly above T_1 in energy (or even slightly below[23]) then S_1 can be thermally populated from T_1, and fluorescence is therefore possible even from excitons which were initially formed in the triplet spin state. Thus TADF requires the fluorophore to have two properties: a small S_1-T_1 gap and a large S_0-S_1 transition dipole moment. A small S_1-T_1 gap is most easily achieved if S_1 and T_1 are intramolecular charge transfer states, in which the exchange interaction between the electron and hole is weak because they are spatially separated. On the other hand, a large S_0-S_1 transition dipole moment, and thus efficient fluorescence from S_1, is possible only when the HOMO and LUMO densities overlap in space. These conflicting requirements make TADF a difficult behavior to achieve in practice, though theoretical results suggest that small S_1-T_1 gaps can be accessed more easily than previously thought.[23] Other more complicated requirements for TADF have been proposed, including that higher energy triplet states, such as heteroatom lone pair (n) to π^* excitations, serve as intermediates for TADF.[22] It has also been suggested that kinetics

may favor the formation of singlet excitons over triplet ones, even when the singlet energy is considerably above that of the triplet.[24]

Having recognized small S_1-T_1 gap and large S_0-S_1 transition dipole moment to be desirable features for fluorophores for OLEDs, it is possible to develop a computational screening or inverse design strategy to identify promising compounds. The pharmaceutical industry has long used high-throughput computational approaches to identify compounds for further experimental study.[25] More recently, such strategies have gained popularity for the design of molecules and materials for optoelectronic applications. In groundbreaking studies, such approaches have identified organic[26-33] and inorganic[34-36] materials well suited for photovoltaic applications, as well as photocatalysts.[37-39] Other targets have included porous materials,[40-48] alloys,[49-54] materials for electrochemical applications,[55-59] and topological insulators.[60] Similar approaches have been used to develop expansive libraries of synthetically accessible molecules.[61] Screening strategies often employ global optimization, data mining, or hierarchical screening procedures[52] and may also replace realistic chemical space, which is discrete due to the integral charges of nuclei, with a description of the potential that can change continuously from that of one molecule to that of another.[30,62-67] In materials science, standard tools for high-throughput studies are now becoming available,[68,69] while in chemistry automated approaches for the discovery of networks of chemical reaction pathways (rather than chemicals themselves) have recently been developed.[70-74]

Herein we present a flexible approach to automated material design and apply it to search a large chemical space for molecules that are likely to exhibit TADF, with the goal of identifying fluorophores for OLEDs. In the algorithm presented here, we represent molecules using an intuitive tree data structure, where edges represent chemical bonds between fragments. The fragments are drawn from a user-chosen set and bound to each other according to user-chosen rules. Together, these define a limited chemical space in which to search, which may be as narrow or as general as is appropriate for the problem at hand and the available computational resources. In the next section, we present how genetic algorithms (GA), global optimizers originally conceived as models of natural selection,[75] can be used in conjunction with this data structure to identify molecules with desirable properties. In Section 3, we analyze the performance of our algorithm and describe the candidate molecules that it has identified. In the final section we draw conclusions and discuss future prospects.

17.1 Method

17.1.1 Tree-based genetic algorithm

In order to search a chemical space for molecules with desirable properties, we must first choose a means to represent the molecules to be optimized and define the chemical space to be searched. We represent molecules as tree data structures, and we will refer to the tree structure defining a specific molecule as a *gene*. A tree structure is composed of nodes which are connected hierarchically by edges. Each node in our tree

represents a single molecular fragment, which can be drawn from a user-defined set of fragments. Each edge represents a chemical bond between fragments, but unlike chemical bonds, the edges of a tree structure are directed; all edges are associated with a *parent* node and a *child* node. One node of the tree is the root node, which has no parent, while all other nodes have exactly one parent. In addition to defining a set of fragments, the user defines a set of rules governing the allowed connectivity between these fragments. Each fragment is defined to have a fixed number of sites capable of binding child fragments. For each site, a list of fragments which may be bound to that site is defined. Together, the list of available fragments and the list of rules governing their bonding define the chemical space to be searched. This space may be either finite or infinite. (Note that the simulated evolution of tree structures is an essential element of genetic programming, the goal of which is to optimize computer programs, rather than molecules, to perform specific tasks,[76] but owing to the dissimilarity between computer programs and molecules, our algorithm is quite different that those typically used in that field.)

The goal of the GA is to find the gene that maximizes the *fitness*, F, which can be any function of molecular properties computable from electronic structure or other calculations. The only requirement on the definition of the fitness function is that $0 \leq F \leq 1$ must hold for all genes. In order to begin an optimization, a *pool* of N_{pool} unique genes is randomly drawn from our molecular space. This is the first *generation*. Each successive generation is determined by a stochastic process composed of three steps

illustrated in Figure 17-1: *selection*, *crossover*, and *mutation*. In the selection step, a set of N_{pool} molecules, termed the *selected set*, is created by drawing genes from the pool as follows. A random integer between one and N_{pool} is chosen, and the molecule with this index is tentatively selected. The fitness of this molecule is compared to a random number between zero and one, r. If $F > r$, this molecule is added to the selected set; if not the tentative selection is discarded and another random molecule is chosen. This procedure is repeated until the selected set contains N_{pool} genes, at which point we move to the next step. Note that this process may result in several identical copies of the same gene being included in the selected set.

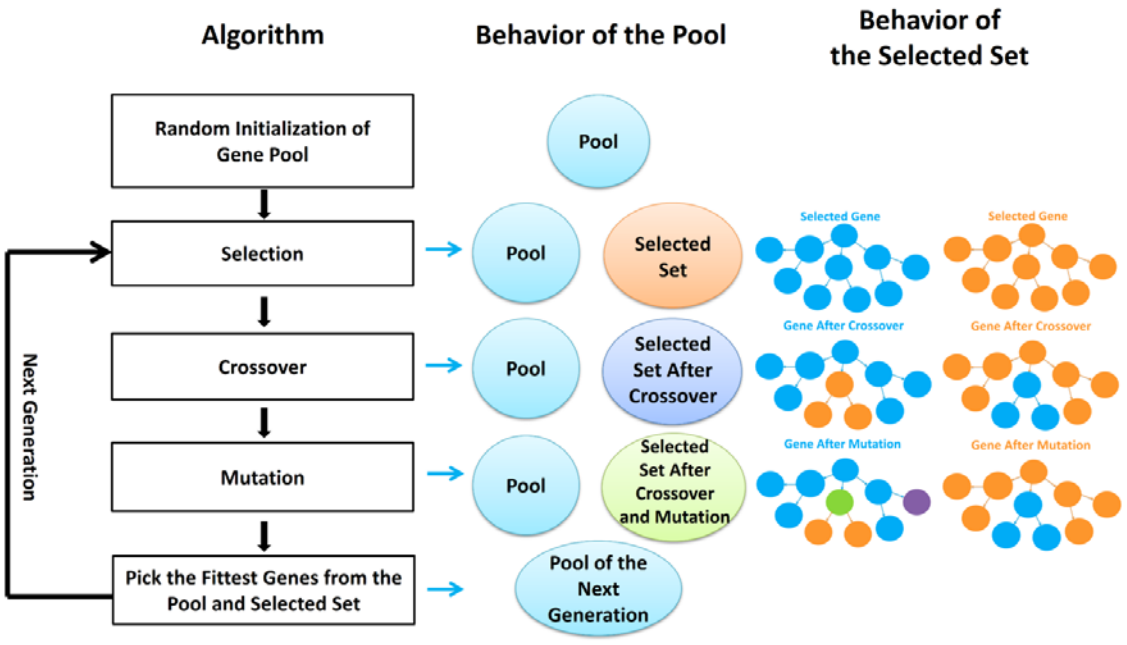


Figure 17-1. Illustration of the tree-based genetic algorithm employed in this work.

Next is the crossover step. In this step, in analogy to the mating of organisms, pairs of molecules in the selected set randomly exchange molecular fragments in hopes of increasing their fitness. A fixed, user-defined number of exchanges, N_{cross}, is undertaken for each generation. For each exchange, a pair of genes, A and B, in the selected set is randomly chosen. In each gene, an edge is then chosen at random. Genes A and B are then replaced with two new genes created by binding the parent of the chosen edge in A to the child of the chosen edge in B, and vice versa. Aside from the two chosen edges, all other edges remain intact during the crossover; that is to say that A and B are divided into subtrees which are reconnected to form new genes, as illustrated in Figure 17-1. In the case that these new genes are disallowed by the user defined rules governing binding between fragments the exchange is aborted and the two chosen genes remain unchanged.

After N_{cross} exchanges are attempted, we proceed to the mutation step. For each fragment in each of the genes in the selected set a random number between zero and one is chosen. If this number is less than a user defined mutation probability, P_{mut}, then a mutation is performed. A fragment is chosen at random to replace the mutating fragment, choosing only from those fragments which are allowed to be bonded to the parent at the present site. If the mutating fragment has one or more children, these children are retained to the extent allowed by the bonding rules. If the mutated fragment has more sites than the fragment it is replacing, then random fragments are chosen to bond to the newly created sites. Similarly, if after mutation one of the children

of the mutated fragment is not allowed by the bonding rules, then a random fragment is chosen to bond to this site, thus ensuring a legal molecule.

Once selection, crossover, and mutation are complete, the fitness values of the modified genes in the selected set are computed. The fitness values of all $2N_{pool}$ genes in the set comprising the pool and the selected set are then compared, and the N_{pool} unique genes with the largest fitness values become the pool of the next generation. As such, each generation is guaranteed to be, on average, at least as fit as the previous generation. By iteratively creating generation after generation, highly fit genes are identified.

As described below, the calculation of the fitness in this study requires time consuming electronic structure calculations. The value of the fitness for a particular gene may be required several times in a single optimization, as the same molecule may evolve multiple times. As such, in our implementation we save all computed fitness values to a database, and each time the fitness of a gene is requested this database is searched. If the fitness of an identical gene has already been computed, it is read from the database rather than recomputed. Thus, our algorithm calculates the fitness of each unique gene only once during an optimization. In addition, the database created by an optimization is persistent and can be reused to accelerate future optimizations or for subsequent data analysis. We also note that the computation of the fitness of the selected set is trivially parallelizable. As such, we have parallelized our algorithm using a master/slave model; a single master process performs the GA optimization, but

distributes the necessary fitness evaluations to a set of waiting slave processes which perform them in parallel.

17.1.2 Definition of molecular space

In order to search for candidate fluorophores which may exhibit TADF, we define a molecular space likely to contain molecules with low lying intramolecular charge transfer excitations similar to those studied by Adachi and coworkers.[15] Each molecule in our space has a central electron acceptor group surrounded by a combination of electron donors and electron withdrawing groups. Figure 17-2 shows the fragments defining our molecular space divided into three categories: electron acceptors, electron donors, and electron withdrawing groups. The rules defining bonding are as follows: Each gene will have an electron acceptor as its root. These electron acceptors each have between four and six sites, marked by Xs in Figure 17-2, which may be bound to either electron donors or electron withdrawing groups. The binding sites on these child fragments are marked by Ys in Figure 17-2. Thus, every bond between fragments connects one X site to one Y site. The donors and electron withdrawing groups have no sites to allow additional bonding in our present work, though in general our algorithm would allow these more complex structures. In total, there are $1.26*10^6$ possible molecules in our chemical space.

For each fragment in Figure 17-2, a two letter abbreviation is assigned which allows us to introduce the following shorthand notation for whole genes. The structure of the tree is indicated by square brackets, which enclose a fragment and the entire sub-

tree rooted at that fragment. For example, suppose we define a molecular space based on two fragments: hydrogen atoms and oxygen atoms. A water molecule could be abbreviated [O[H][H]], indicating that the oxygen atom is the root, and the two hydrogen atoms are children of (bonded to) the oxygen atom. Recognizing that chemical bonding is not directional, a second representation of the water molecule would be [H[O[H]]], where now a hydrogen atom is the root. For fragments with multiple bonding sites, the adjacent children are bonded to physically adjacent sites. For example, considering a molecular space defined by a benzene ring (Bz) bonded to either hydrogen or fluorine atoms, one representation of o-difluorobenzene would be [Bz[F][F][H][H][H][H]], while p-difluorobenzene could be represented [Bz[F][H][H][F][H][H]]. For the acceptor groups with heteroatoms, the first site is adjacent to the heteroatom. Thus [Fr[F][H][H][H]] would represent 2-fluorofuran, and [Fr[H][F][H][H]] would represent 3-fluorofuran.

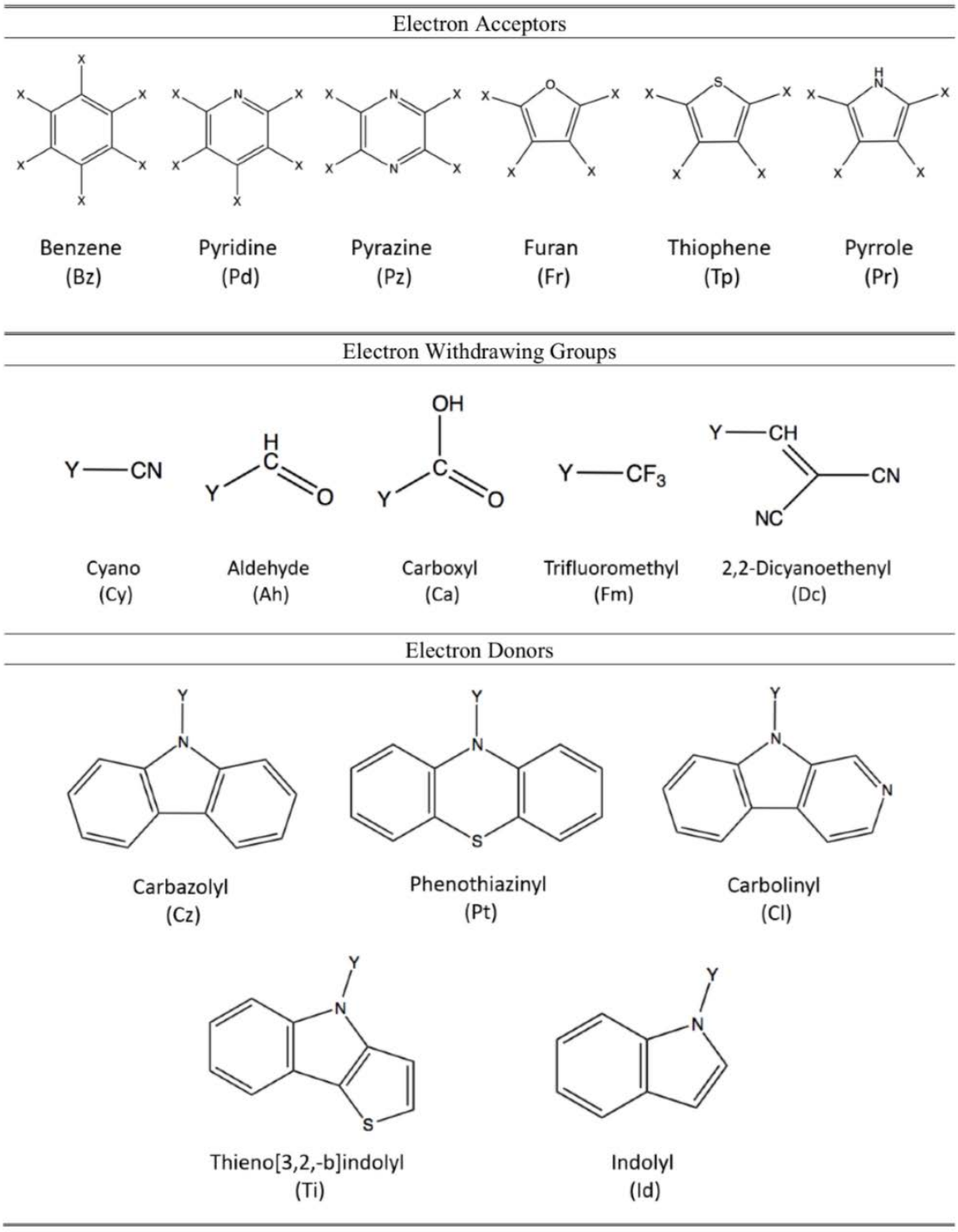


Figure 17-2. The pictured fragments define the molecular space explored in this work. The X sites on the acceptor groups are allowed to form bonds with the Y sites on the withdrawing and donor groups.

17.2 Fitness Evaluation

The fitness of a given gene is computed in a series of steps. First, a crude approximation to the molecular geometry of the compound is constructed as follows. Prior to genetic optimization, reasonable structures of each individual fragment are predefined in Z-matrix internal coordinates by the user. For each fragment, the user also defines bond distances, angles, and dihedral angels which can be used to connect that fragment to a site. These parameters are fragment specific, but are general to all sites. In practice, we choose angles and dihedrals which do not lead to perfect symmetry (e.g. we avoid dihedral angles of 0º and 90º) to reduce the probability that subsequent geometry optimizations become trapped at unstable points on the PES. Using these parameters, a Z-matrix for the entire molecule can be automatically constructed for any gene. The initial structure is then optimized at the B3LYP/STO-3G level of theory.[77] The minimal basis set is chosen for efficiency. At the optimized structure, the S_0 state is computed at the CAM-B3LYP/6-31G** level of theory, the T_1 state is computed using restricted open-shell (RO-) CAM-B3LYP/6-31G**, and the S_1 state is computed using Tamm-Dancoff time-dependent (TD) density functional theory at the CAM-B3LYP/6-31G** level of theory.[78] From these calculations, the S_1-T_1 gap,

$$\Delta E_{S_1-T_1} = E_{S_1} - E_{T_1} \tag{18-1}$$

and the norm of the S_0-S_1 transition dipole moment, $\left\|\vec{\mu}_{S_1}\right\|$, are both computed. In order to optimize for small $\Delta E_{S_1-T_1}$ and large $\left\|\vec{\mu}_{S_1}\right\|$ we choose the fitness to be defined

$$F = e^{-500\left(\Delta E_{S_1-T_1}\right)^2 - 0.059\left\|\vec{\mu}_{S_1}\right\|^{-1}} \quad (18\text{-}2)$$

This function increases when increasing $\left\|\vec{\mu}_{S_1}\right\|$ or decreasing $\left|\Delta E_{S_1-T_1}\right|$. It decreases rapidly to zero as $\left\|\vec{\mu}_{S_1}\right\|$ approaches zero. The constants (500 and 0.059) are in atomic units and were chosen with the goal of providing a balanced treatment of the two molecular properties we wish to optimize. (Note that 0.059 a.u. of dipole = 0.15 Debye.) With these parameters, genes with $\left|\Delta E_{S_1-T_1}\right|$ greater than 0.5 eV or $\left\|\vec{\mu}_{S_1}\right\|$ less 1.0 Debye, which will likely have a low propensity for TADF, will also have low fitness values. The exponential form of the fitness together with the positivity of $\left\|\vec{\mu}_{S_1}\right\|$ and $\left(\Delta E_{S_1-T_1}\right)^2$ guarantees that $0 \le F \le 1$. Future work will utilize a less empirical definition of the fitness based on an expression for the rate of TADF.

It should be pointed out that the use of the S_1 and T_1 energies computed at the S_0 minimum is an approximation. Nuclear relaxation on the excited state is fast compared to intersystem crossing, therefore it would be ideal to compute the S_1 and T_1 energies at the minimum energy geometries of these states rather than the Franck-Condon point. In the case that the S_1 and T_1 PESs are parallel, however, the gap is geometry independent. It is not unreasonable to suppose that this is roughly the case, as S_1 and T_1 generally correspond to charge transfer states of similar character, differing only in the spin coupling of the electron and hole. Assuming a weak geometry dependence of the exchange interaction (which itself is quite weak in ideal TADF fluorophores), it is thus not unreasonable to approximate the more relevant relaxed S_1-T_1 gap by the more conveniently computed S_1-T_1 gap at the S_0 minimum.

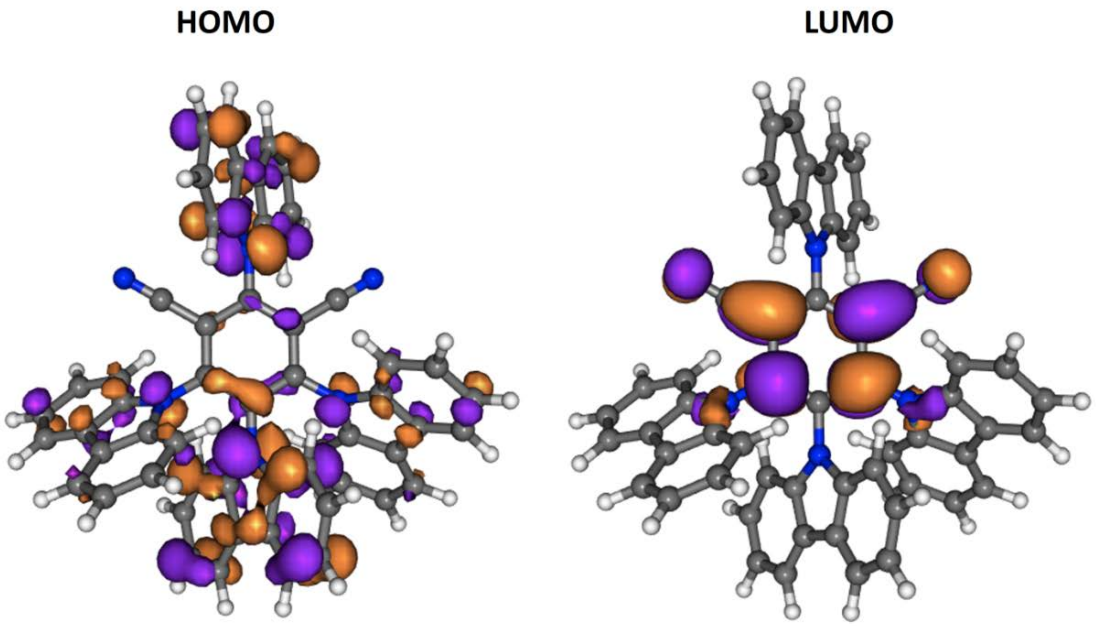


Figure 17-3. HOMO and LUMO orbitals of the 4CzIPN molecule computed at the CAM-B3LYP/6-31G** level of theory.

To calibrate this fitness, we consider one of the carbazolyl dicyanobenzene dyes shown by Adachi and coworkers to undergo TADF, abbreviated 4CzIPN (and represented as [Bz[Cy][Cz][Cy][Cz][Cz][Cz]] in our notation).[15] The 4CzIPN molecule is shown in Figure 17-3 with its highest occupied (HO-) and lowest unoccupied molecular orbitals (LUMO), which correspond to the dominant TDDFT amplitude for excitation to S_1. This state has charge transfer character, as expected. The computed values of $\Delta E_{S_1-T_1}$ and $\left\|\vec{\mu}_{S_1}\right\|$ for 4CzIPN are 0.41 eV and 1.95 Debye, respectively. Computed according to the above procedure, the fitness of 4CzIPN is 0.828. Therefore in this work we will consider molecules with fitness values of 0.8 or higher as reasonable candidate fluorophores.

Our GA optimizations require significant computer time, and as such we take advantage of graphics processing unit (GPU) acceleration to compute the fitness; all of the calculations in this work were performed using the implementations of DFT and TDDFT in a development version of the TeraChem GPU-accelerated quantum chemistry software package.[79-82]

17.3 Results and Discussion

17.3.1 Overview of GA results

Applying our GA methodology to the chemical space described above, we have run eight optimizations, utilizing different values for N_{pool} and P_{mut}. In all cases, N_{cross} is set to 3. All optimizations were stopped when they reached 200 generations. In total, the fitness values of 7518 molecules were computed, 3792 of which had $F > 0.8$. Thus 50.4% of the molecules computed are reasonable candidates for further testing. To evaluate the efficiency of the sampling we have achieved, we compare this percentage to a randomly sampled set of molecules from our space: the 110 randomly chosen molecules used to initialize our optimizations. Of these only 13, or 12%, have $F > 0.8$. Thus, the GA algorithm employed here provides a considerably more efficient sampling of the available molecular space than naïve random sampling of molecules from the same space. We summarize the set of molecules whose properties were computed in our study in Figure 17-4, which plots all computed genes according to $\left\| \vec{\mu}_{S_1} \right\|$ and $\Delta E_{S_1-T_1}$. The total fitness is indicated by the color of each point, and the Pareto optimal frontier (the set of

genes for which no gene has been observed to have both greater $\left\|\vec{\mu}_{S_1}\right\|$ and smaller $\Delta E_{S_1-T_1}$) is indicated by a dashed red line. This plot confirms our assertion that high $\left\|\vec{\mu}_{S_1}\right\|$ and low $\Delta E_{S_1-T_1}$ are difficult to achieve simultaneously; the molecules with the largest $\left\|\vec{\mu}_{S_1}\right\|$ tend also to have larger $\Delta E_{S_1-T_1}$. Table S1 in Supplementary Materials lists all molecules in the Pareto optimal frontier.[83]

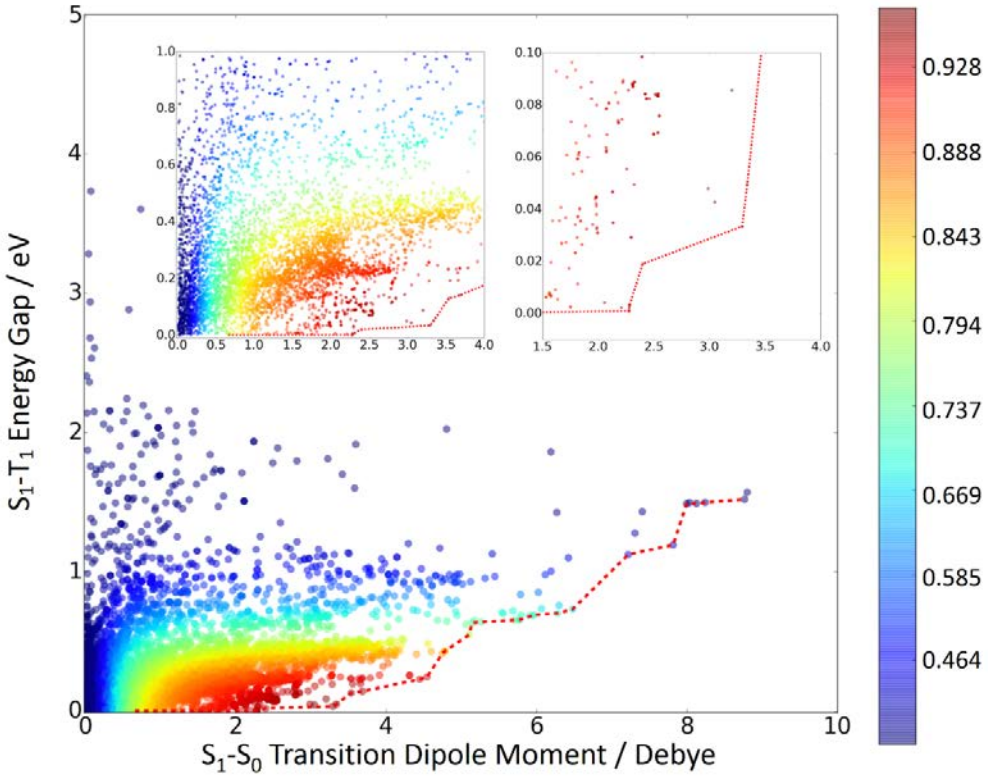


Figure 17-4. A scatter plot of the S_1-T_1 energy gap and S_1-S_0 transition dipole moment of the 7518 molecules computed in this study. The color of the points indicates the fitness values of the genes according to the scale on the right. The red dashed line is the Pareto optimal frontier of the data set. The two inset graphs show the same data on different scales for clarity.

17.3.2 GA performance

Before considering the chemical nature of the viable candidates we have identified, we analyze the performance of our algorithm as a function of N_{pool} and P_{mut}. Figure 17-5(a) presents the average fitness as a function of generation for a series of optimizations with P_{mut} ranging from 0.01 to 0.15 and $N_{pool} = 10$. Figure 17-5(b) presents the same data for a series of optimizations with N_{pool} ranging from 10 to 30 and $P_{mut} = 0.05$. In all cases, the average fitness increases rapidly in the initial generations, before leveling out. There is no obvious pattern relating either P_{mut} or N_{pool} to the initial rate of increase or the average fitness after 200 generations. Table 17-1 presents the three fittest genes from each simulation. Again, there is no obvious correlation between P_{mut} and the fittest genes identified. However, it is noteworthy that the optimization utilizing the largest pool ($N_{pool} = 30$) resulted in the identification of several genes with $F > 0.95$, the fittest of any identified in this study. This is not reflected in the average fitness values in Figure 17-5 because the pool in the $N_{pool} = 30$ optimization contains several lower-fitness genes which result in a lower average than some of the optimizations using smaller pools.

Table 17-1. Information pertaining to the eight individual optimizations performed in this work is presented, including the user defined parameters (N_{pool} and P_{mut}), the total number of fitness evaluations performed, and the structures, fitness values, singlet-triplet energy gaps, and S_1 transition dipole moments of the three fittest genes computed in each optimization.

N_{pool}	P_{mut}	# of Evals.	Fittest Genes	Fitness Value	$\Delta E_{S_1-T_1}$ / meV	μ_{S_1} / Debye
30	0.05	1656	[Pd[Pt][Id][Cl][Ti][Ti]]	0.955	33.3	3.29
			[Pd[Pt][Id][Cz][Ti][Ti]]	0.955	49.2	3.40
			[Pd[Pt][Id][Cz][Cz][Ti]]	0.951	42.8	3.05
20	0.05	1197	[Pd[Ti][Cy][Cz][Cy][Cl]]	0.940	69.3	2.54
			[Bz[Cz][Cy][Id][Cy][Cz][Cy]]	0.938	83.4	2.54
			[Pd[Cz][Cy][Cz][Cy][Cl]]	0.938	128.5	2.85
10	0.15	1287	[Bz[Cy][Cl][Cy][Cz][Cy][Cz]]	0.939	116.6	2.80
			[Bz[Cy][Cz][Cy][Cz][Cy][Id]]	0.938	84.2	2.54
			[Bz[Cy][Id][Cy][Cz][Cy][Cz]]	0.938	83.9	2.52
10	0.10	1086	[Pd[Ti][Cl][Cz][Cy][Cl]]	0.940	111.9	2.79
			[Pd[Ti][Cy][Cz][Cy][Cl]]	0.940	69.3	2.54
			[Pd[Ti][Cy][Cz][Cy][Cz]]	0.939	75.8	2.55
10	0.07	607	[Fr[Fm][Dc][Dc][Pt]]	0.912	101.9	1.77
			[Fr[Ti][Dc][Dc][Pt]]	0.890	195.8	1.66
			[Fr[Ca][Dc][Dc][Pt]]	0.884	153.6	1.40
10	0.05	704	[Bz[Cy][Cl][Cy][Cz][Cy][Cz]]	0.939	116.6	2.80
			[Bz[Cy][Cz][Cy][Cz][Cy][Id]]	0.938	84.2	2.54
			[Bz[Cy][Id][Cy][Cz][Cy][Cz]]	0.938	83.9	2.52
10	0.03	582	[Pd[Cz][Cy][Ti][Cy][Ti]]	0.939	68.9	2.50
			[Bz[Cz][Cy][Id][Cy][Cz][Cy]]	0.938	83.3	2.54
			[Bz[Cz][Cy][Cz][Cy][Id][Cy]]	0.938	83.3	2.54
10	0.01	399	[Pd[Ti][Cy][Cz][Cy][Cz]]	0.939	75.8	2.55
			[Pd[Cz][Cy][Ti][Cy][Ti]]	0.939	68.9	2.50
			[Bz[Cz][Cy][Cz][Cy][Id][Cy]]	0.938	83.3	2.54

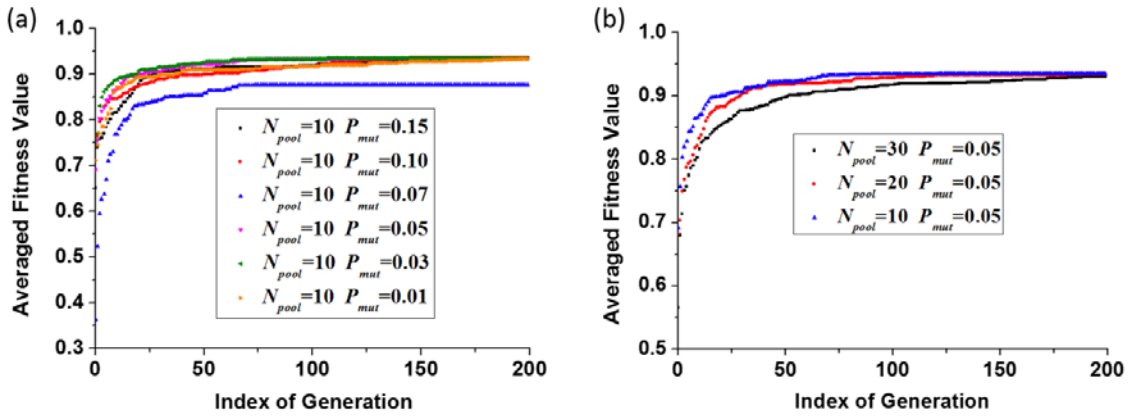


Figure 17-5. The average fitness of the pool as a function of generation for series of GA optimizations with (a) identical N_{pool}, but varying P_{mut} and (b) identical P_{mut}, but varying N_{pool}.

Interestingly, though many of the optimizations converge to similar solutions, the optimization with $P_{mut} = 0.7$ and $N_{pool} = 10$ has become trapped in a lower-fitness region of chemical space. Specifically, the fittest genes explored by this optimization are furan-based molecules with $F < 0.92$, while the other seven optimizations found higher-fitness genes ($F > 0.92$) based on benzene or pyridine acceptors rather than furan. This points out that even though GAs are global optimizers, optimizations with the parameters explored in this work can easily become stuck in local minima. Though this may seem a disadvantage, it is worth noting that in the present work it is not our goal to find the fittest gene so much as it is to find a diverse set of reasonably fit genes. As such, the identification of furan-based fluorophores with $F > 0.80$ is fortuitous. This suggests that when diversity is desired, running a set of uncoupled GA optimizations with smaller

pools may be preferable to running a single optimization with a larger pool. In a future study we will investigate strategies for ensuring the diversity of our identified candidate pool.

How the result of each optimization depends on the initial content of the pool has also been investigated. Figure 17-6 shows a scatter plot relating the initial content of the pool to the final content. For each optimization, a point is present for each of the three most common acceptor groups: benzene, pyridine, and furan. On the x-axis is the fraction of genes in the pool based on that acceptor at the beginning of the optimization, and the y-axis is the fraction of genes in the pool based on that acceptor at the end of the pool. If the results of the optimization depend strongly on the initial content of the pool, we would expect to see a positive correlation, but, no such correlation is observed.

Table 17-1 also reports the total number of unique genes explored by each simulation. There is a clear increase in this number with P_{mut}, from 399 for $P_{mut} = 0.01$ to 1287 for $P_{mut} = 0.15$. This is significant, since it is the number of unique fitness evaluations that determines the total computational cost of the optimization, not the number of generations. In this sense, a smaller value of P_{mut} can be seen as advantageous, as it allows for a similarly fast increase in the average fitness of the pool for a lower computational cost. On the other hand, at lower P_{mut}, a smaller swath of molecular space is explored, and as such a smaller and less diverse final set of fit genes are likely to be identified. So, the ideal value of P_{mut} depends, at least in part, on

whether the primary goal of the optimization is to efficiently identify the very fittest genes or to identify a large and diverse pool of reasonably fit genes.

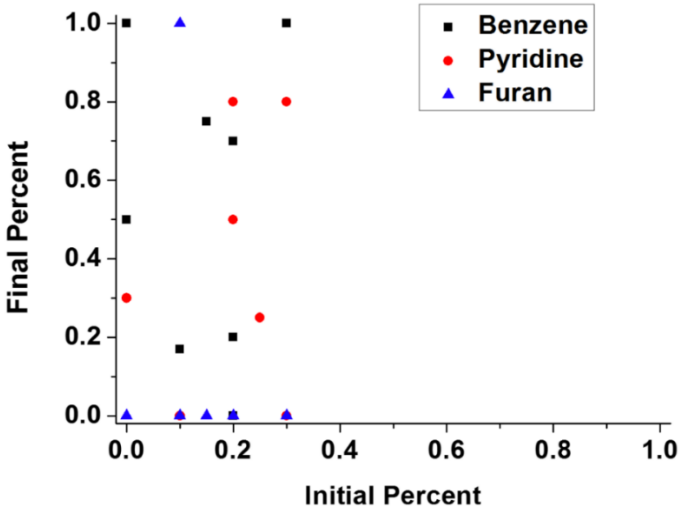


Figure 17-6. This scatter plot relates the initial content of the pool to the final content. Three points corresponding to the three most common acceptor groups (benzene, black; pyridine, red; furan, blue) are plotted for each optimization, with the fraction of genes based on that acceptor in the initial pool on the x-axis and the same fraction in the final pool on the y-axis.

17.3.3 Candidate fluorophores

Now we turn our attention to the chemical nature of the candidate fluorophores identified in this study. With the database created by our eight genetic optimizations, it is possible to investigate the structural similarities among promising TADF fluorophores, with the hope of guiding future experimental studies. Figure 17-7(a) and (b) show matrices containing the total number of fit ($F > 0.8$) genes identified that contain a

specific pair of fragments. For example, three fit molecules containing both a pyrrole acceptor and a cyano withdrawing group were identified, as indicated by the top left corner of Figure 17-7(a). Figure 17-7(a) has the central electron acceptors on the y-axis and the donors and withdrawing groups on the x-axis, while Figure 17-7(b) shows the donor and withdrawing groups on both the x- and y-axes. The diagonal elements of Figure 17-7(b) report the total number of fit genes containing two (or more) copies of the corresponding fragment. Additionally, Table 17-1 presents the fittest three molecules observed in each optimization. A more complete list of the fittest molecules observed in these optimizations is presented in Table S2.

Figure 17-7 shows that, among the various central electron acceptors, benzene and pyridine were included in the majority of the viable candidates that were identified. Nearly every benzene-centered fluorophore contains at least one cyano group, i.e. 2168 candidates contain at least one cyano group out of 2263 total benzene-based candidates. The other electron withdrawing groups are present in far fewer candidates. Among donor groups, the carbazolyl group is nearly as common as the cyano group, being included in 2068 benzene-based candidates. This is not surprising, given that Adachi and coworkers have already demonstrated that carbozolyl dicyanobenzenes undergo efficient TADF.[15] In fact, many of the fittest genes observed in our optimizations are similar to these fluorophores (e.g. [Bz[Cz][Cy][Id][Cy][Cz][Cy]], [Bz[Cy][Cl][Cy][Cz][Cy][Cz]], and [Bz[Cy][Cz][Cy][Cz][Cy][Id]]; all presented in Table 17-1), but unlike Adachi's dicyanobenzene fluorophores, the most fit molecules are often 1,3,5-tricyanobenzene-

based. Many viable candidates also include thienoindolyl, indolyl, and carbolinyl donor groups, and further investigation of such molecules would be a promising avenue for future research.

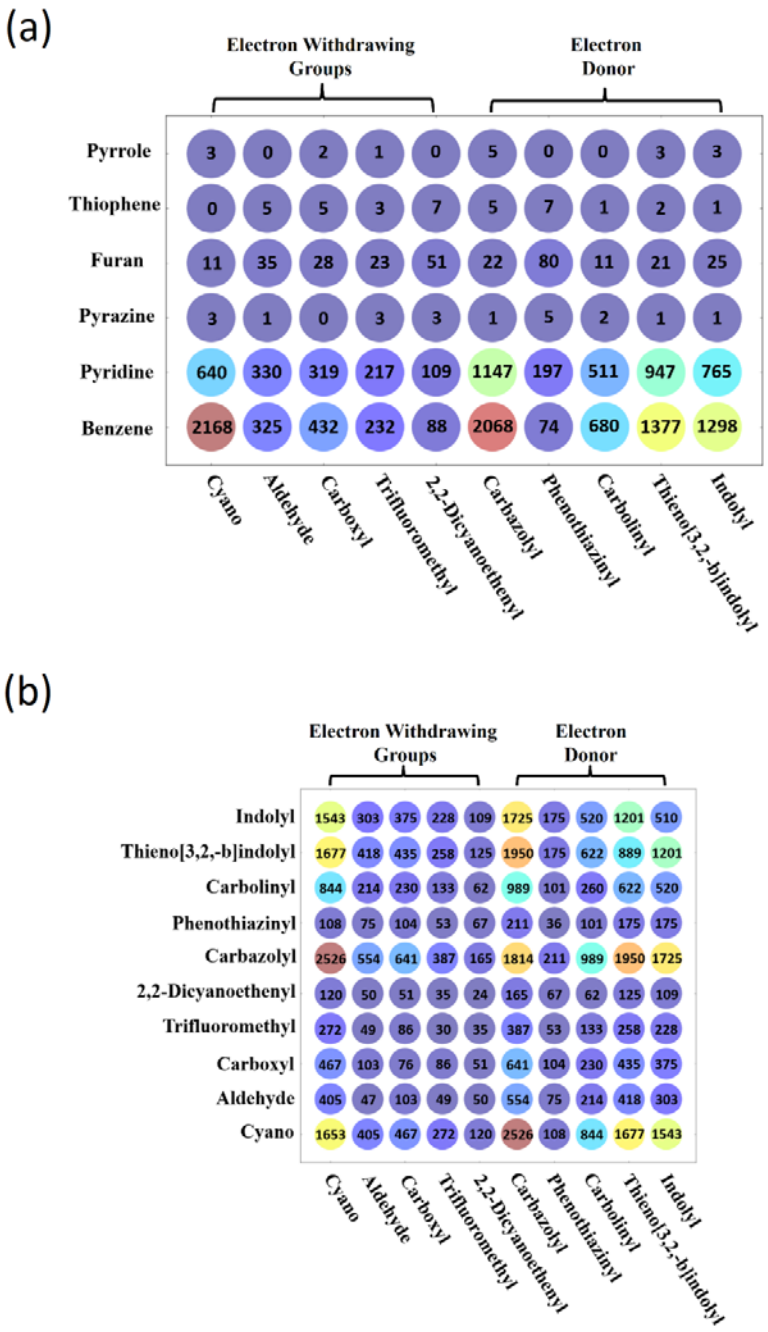


Figure 17-7. Matrices showing the numbers of fit molecules (those with $F > 0.8$)

Figure 17-7 (cont'd). containing specific pairs of fragments. (a) shows the number of fit molecules containing the acceptor group on the y-axis and the substituent (donor or withdrawing group) on the x-axis. (b) shows the number of fit molecules containing one copy of each of the substituents on the x- and y-axes. Colored dots guide the eye by reflecting the value of each element, with smaller elements indicated by colder colors and larger elements by warmer colors.

Though cyano groups are often present in pyridine-based fluorophores, they do not appear to be essential like they are in the benzene-based molecules; 640 out of 1418 pyridine-based candidates contain these groups. As with benzene-based dyes, carbazolyl, thienoindolyl, indolyl, and carbolinyl groups appear in a large number of pyridine-based candidates (1147, 947, 765, and 511, respectively). Several of the fittest molecules in Table 17-1 are 3,5-dicyanopyridine- (e.g. [Pd[Ti][Cy][Cz][Cy][Cl]]) or 3-cyanopyridine-based (e.g. [Pd[Ti][Cl][Cz][Cy][Cl]]), while others (e.g. [Pd[Pt][Id][Cl][Ti][Ti]], [Pd[Pt][Id][Cz][Ti][Ti]], and [Pd[Pt][Id][Cz][Cz][Ti]], the three fittest genes found in our study) contain no withdrawing groups of any kind. Based on this work, we suggest that additional investigation of pyridine-based fluorophores bound to various electron donors—with or without cyano withdrawing groups—would be promising. In fact, since we completed our study, carbazolyl pyridines have been reported as suitable host materials for phosphorescent blue OLEDs, due to their high triplet energy and blue-to-ultraviolet emission.[84]

As noted above, 88 candidates based on a furan acceptor were identified. Interestingly, only 11 of these contain cyano groups, and only 22 contain carbazolyl donors. However, 51 contain at least one 2,2-dicyanoethenyl withdrawing group and 80 contain a phenothiazinyl donor, making these the most common substituents in furan-based candidates. This is in stark contrast to the benzene- and pyridine-based candidates discussed above, for which these are the least frequently observed fragments in their respective categories, and suggests a new target for experimental investigation of TADF-capable fluorophores. The specificity of the interaction between donor, acceptor, and withdrawing group is surprising and underlines the value of the type of exploratory algorithm employed here. The fittest fluorophores of this type (e.g. [Fr[Fm][Dc][Dc][Pt]], [Fr[Ti][Dc][Dc][Pt]]), have 2,2-dicyanoethenyl groups at the 3 and 4 positions of the furan ring. Only a total of 22 genes with $F > 0.8$ based on the other acceptor groups—pyrazine, thiophene, or pyrrole—were observed.

Figure 17-8 presents a histogram of the S_1 vertical excitation energies of the molecules with $F > 0.8$. It is noteworthy that these cover the entire visible range, along with near-infrared (down to 1.6 eV) and ultraviolet (up to 4.6 eV) energies. Thus, the set of candidates identified here likely covers a wide range of desirable emission wavelengths. Note that the vertical excitation energy itself does not correspond to the emission energy, as it neglects the Stokes shift, and the application of TDDFT, even with a range-corrected functional such as CAM-B3LYP, likely results in substantial errors in the calculated vertical

excitation energies. Further study would be needed to predict the emission energy with the level of accuracy required to predict luminescence color.

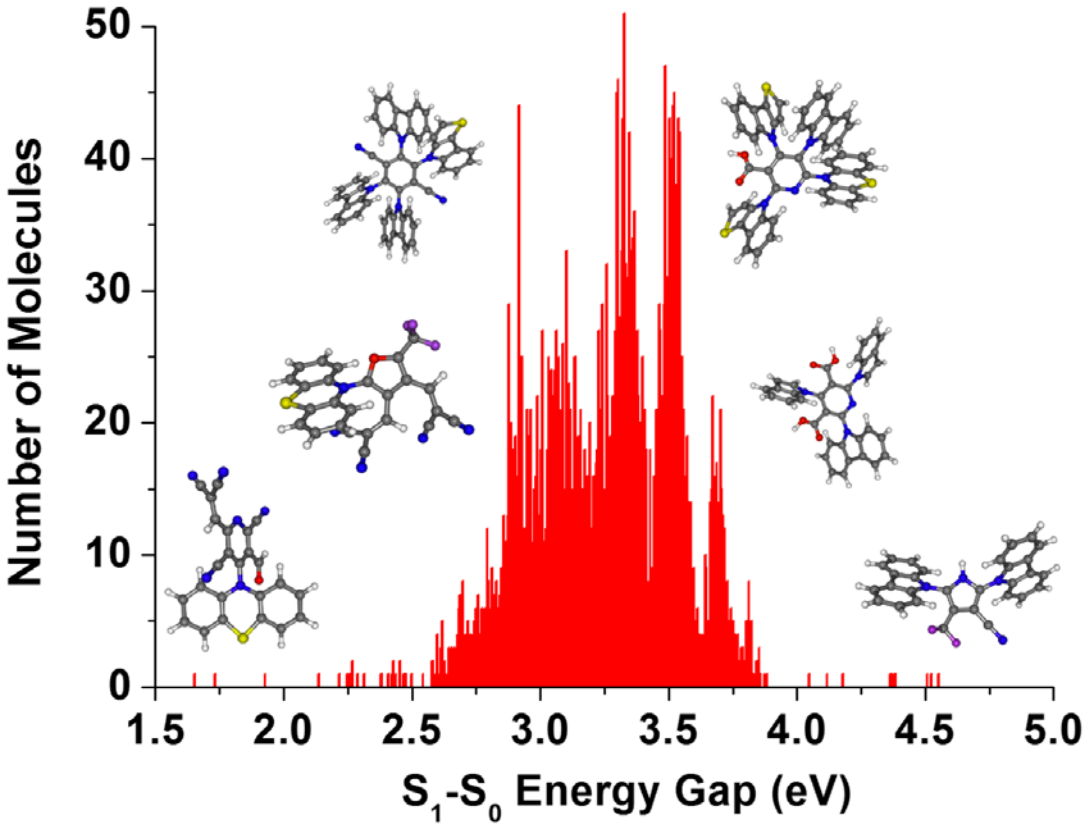


Figure 17-8. This histogram presents the 3792 molecules with $F > 0.8$ binned according to their respective S_1 vertical excitation energies. Insets show representative molecules from left to right having calculated excitations energies of 1.65 eV, 2.79eV, 3.10eV, 3.53eV, 3.77eV, and 4.55 eV.

The candidates identified by our optimizations vary considerably in the ease with which they can be synthesized. In particular, highly asymmetric molecules with many

different fragments may be more challenging to synthesize than those built from multiple identical fragments. As such, we have highlighted those molecules with $F > 0.8$ that contain at most one variety of withdrawing group and one variety of donor in Table S3. In total, there are 65 such candidates, 64 of which are benzene-, pyridine-, or furan-based. The sole exception is a carbazolyl dicyanopyrazine ([Pz[Cz][Cz][Cy][Cy]]). Also noteworthy are a small number of pyridine-based molecules with aldehyde or carboxylic acid withdrawing groups instead of cyano groups, (e.g. [Pd[Cz][Cz][Ah][Cz][Cz]] and [Pd[Cz][Ca][Ca][Ca][Cz]]). In the future we will incorporate a more sophisticated descriptor of synthetic accessibility, such as that developed by Virshup, et al.,[61] directly into the fitness function used in our algorithm.

17.4 Conclusions

We have applied a genetic algorithm to identify a set of fluorophores which show promise for application in OLEDs. Specifically, our goal has been to identify molecules likely to undergo thermally activated delayed fluorescence by optimizing for small S_1-T_1 gap and large S_1 transition dipole moment. In addition to carbazolyl cyanobenzenes, which have previously been reported to undergo TADF, our evolutionary optimizations have identified several classes of molecules which may undergo TADF: pyridines bound to various electron donors (without withdrawing group), 3-cyano- or 3,5-dicyanopyridines bound to various electron donors, carbazolyl pyridines with carboxyl or aldehyde withdrawing groups, carbazolyl cyanopyrazines, and phenothiazinyl 3,4-di(2,2-

dicyanoethenyl)-furans. We hope that the present work will stimulate experimental efforts to study these fluorophores.

More generally, our GA approach is found to provide efficient sampling of chemical space. By representing the molecule with a tree structure, with each node representing a molecular fragment drawn from a user-defined set, our approach is scalable and flexible, allowing the user to search either a very broad or very narrow chemical space. Our implementation takes advantage of advanced hybrid parallel computer architectures by farming individual fitness evaluations out to any number of graphics processing units. This approach can be tailored both to problems where the target of the study is to identify a small number of nearly optimal molecules or a large and diverse set of viable candidates for further study, and future work will consider how to optimize our procedure for either of these goals.

REFERENCES

REFERENCES

(1) Grimsdale, A. C.; Chan, K. L.; Martin, R. E.; Jokisz, P. G.; Holmes, A. B. *Chem Rev* **2009**, *109*, 897.

(2) Brown, A. R.; Bradley, D. D. C.; Burroughes, J. H.; Friend, R. H.; Greenham, N. C.; Burn, P. L.; Holmes, A. B.; Kraft, A. Appl Phys Lett 1992, 61, 2793.

(3) Baldo, M. A.; Lamansky, S.; Burrows, P. E.; Thompson, M. E.; Forrest, S. R. Appl Phys Lett 1999, 75, 4.

(4) Adachi, C.; Baldo, M. A.; Thompson, M. E.; Forrest, S. R. *J Appl Phys* **2001**, *90*, 5048.

(5) Tanaka, D.; Sasabe, H.; Li, Y. J.; Su, S. J.; Takeda, T.; Kido, J. Jpn J Appl Phys 2 2007, 46, L10.

(6) Pawlik, T. D.; Kondakova, M. E.; Giesen, D. J.; Deaton, J. C.; Kondakov, D. Y. Sid Int Symp Dig Tec 2008, 39, 613.

(7) Chopra, N.; Lee, J.; Zheng, Y.; Eom, S. H.; Xue, J. G.; So, F. *Appl Phys Lett* **2008**, *93*.

(8) Xiao, L. X.; Chen, Z. J.; Qu, B.; Luo, J. X.; Kong, S.; Gong, Q. H.; Kido, J. J. Adv Mater 2011, 23, 926.

(9) Udagawa, K.; Sasabe, H.; Cai, C.; Kido, J. *Adv Mater* **2014**, *26*, 5062.

(10) Deaton, J. C.; Switalski, S. C.; Kondakov, D. Y.; Young, R. H.; Pawlik, T. D.; Giesen, D. J.; Harkins, S. B.; Miller, A. J. M.; Mickenberg, S. F.; Peters, J. C. J Am Chem Soc 2010, 132, 9499.

(11) Tsuboyama, A.; Kuge, K.; Furugori, M.; Okada, S.; Hoshino, M.; Ueno, K. Inorg Chem 2007, 46, 1992.

(12) Hashimoto, M.; Igawa, S.; Yashima, M.; Kawata, I.; Hoshino, M.; Osawa, M. J Am Chem Soc 2011, 133, 10348.

(13) Czerwieniec, R.; Yu, J. B.; Yersin, H. *Inorg Chem* **2011**, *50*, 8293.

(14) Czerwieniec, R.; Kowalski, K.; Yersin, H. *Dalton T* **2013**, *42*, 9826.

(15) Uoyama, H.; Goushi, K.; Shizu, K.; Nomura, H.; Adachi, C. *Nature* **2012**, *492*, 234.

(16) Zhang, Q. S.; Li, J.; Shizu, K.; Huang, S. P.; Hirata, S.; Miyazaki, H.; Adachi, C. J Am Chem Soc 2012, 134, 14706.

(17) Mehes, G.; Nomura, H.; Zhang, Q. S.; Nakagawa, T.; Adachi, C. Angew Chem Int Edit 2012, 51, 11311.

(18) Li, J.; Nakagawa, T.; MacDonald, J.; Zhang, Q. S.; Nomura, H.; Miyazaki, H.; Adachi, C. Adv Mater 2013, 25, 3319.

(19) Zhang, Q. S.; Li, B.; Huang, S. P.; Nomura, H.; Tanaka, H.; Adachi, C. Nat Photonics 2014, 8, 326.

(20)Nakanotani, H.; Higuchi, T.; Furukawa, T.; Masui, K.; Morimoto, K.; Numata, M.; Tanaka, H.; Sagara, Y.; Yasuda, T.; Adachi, C. Nat Commun 2014, 5.

(21) Cho, Y. J.; Yook, K. S.; Lee, J. Y. *Adv Mater* **2014**, *26*, 6642.

(22) Dias, F. B.; Bourdakos, K. N.; Jankus, V.; Moss, K. C.; Kamtekar, K. T.; Bhalla, V.; Santos, J.; Bryce, M. R.; Monkman, A. P. Adv Mater 2013, 25, 3707.

(23) Difley, S.; Beljonne, D.; Van Voorhis, T. *J Am Chem Soc* **2008**, *130*, 3420.

(24) Shuai, Z.; Beljonne, D.; Silbey, R. J.; Bredas, J. L. *Phys Rev Lett* **2000**, *84*, 131.

(25) Fink, T.; Reymond, J. L. *J Chem Inf Model* **2007**, *47*, 342.

(26) Hachmann, J.; Olivares-Amaya, R.; Atahan-Evrenk, S.; Amador-Bedolla, C.; Sanchez-Carrera, R. S.; Gold-Parker, A.; Vogt, L.; Brockway, A. M.; Aspuru-Guzik, A. J Phys Chem Lett 2011, 2, 2241.

(27) Martsinovich, N.; Troisi, A. *J Phys Chem C* **2011**, *115*, 11781.

(28) O'Boyle, N. M.; Campbell, C. M.; Hutchison, G. R. *J Phys Chem C* **2011**, *115*, 16200.

(29) Sokolov, A. N.; Atahan-Evrenk, S.; Mondal, R.; Akkerman, H. B.; Sanchez-Carrera, R. S.; Granados-Focil, S.; Schrier, J.; Mannsfeld, S. C. B.; Zoombelt, A. P.; Bao, Z. N.; Aspuru-Guzik, A. Nat Commun 2011, 2.

(30) Xiao, D. Q.; Martini, L. A.; Snoeberger, R. C.; Crabtree, R. H.; Batista, V. S. J Am Chem

Soc 2011, 133, 9014.

(31) Kanal, I. Y.; Owens, S. G.; Bechtel, J. S.; Hutchison, G. R. *J Phys Chem Lett* **2013**, *4*, 1613.

(32) Hachmann, J.; Olivares-Amaya, R.; Jinich, A.; Appleton, A. L.; Blood-Forsythe, M. A.; Seress, L. R.; Roman-Salgado, C.; Trepte, K.; Atahan-Evrenk, S.; Er, S.; Shrestha, S.; Mondal, R.; Sokolov, A.; Bao, Z. A.; Aspuru-Guzik, A. Energ Environ Sci 2014, 7, 698.

(33) Ornso, K. B.; Pedersen, C. S.; Garcia-Lastra, J. M.; Thygesen, K. S. Phys Chem Chem Phys 2014, 16, 16246.

(34) Castelli, I. E.; Olsen, T.; Datta, S.; Landis, D. D.; Dahl, S.; Thygesen, K. S.; Jacobsen, K. W. Energ Environ Sci 2012, 5, 5814.

(35) Castelli, I. E.; Landis, D. D.; Thygesen, K. S.; Dahl, S.; Chorkendorff, I.; Jaramillo, T. F.; Jacobsen, K. W. Energ Environ Sci 2012, 5, 9034.

(36) Yu, L. P.; Zunger, A. *Phys Rev Lett* **2012**, *108*.

(37) Greeley, J.; Jaramillo, T. F.; Bonde, J.; Chorkendorff, I. B.; Norskov, J. K. Nat Mater 2006, 5, 909.

(38) Wu, Y. B.; Lazic, P.; Hautier, G.; Persson, K.; Ceder, G. *Energ Environ Sci* **2013**, *6*, 157.

(39) Wu, Y. B.; Ceder, G. *J Phys Chem C* **2013**, *117*, 24710.

(40) Kim, J.; Martin, R. L.; Rubel, O.; Haranczyk, M.; Smit, B. J Chem Theory Comput 2012, 8, 1684.

(41) Martin, R. L.; Willems, T. F.; Lin, L. C.; Kim, J.; Swisher, J. A.; Smit, B.; Haranczyk, M. Chemphyschem 2012, 13, 3595.

(42) Sikora, B. J.; Wilmer, C. E.; Greenfield, M. L.; Snurr, R. Q. *Chem Sci* **2012**, *3*, 2217.

(43) Wilmer, C. E.; Leaf, M.; Lee, C. Y.; Farha, O. K.; Hauser, B. G.; Hupp, J. T.; Snurr, R. Q. Nat Chem 2012, 4, 83.

(44) Martin, R. L.; Lin, L. C.; Jariwala, K.; Smit, B.; Haranczyk, M. J Phys Chem C 2013, 117, 12159.

(45) Colon, Y. J.; Snurr, R. Q. *Chem Soc Rev* **2014**, *43*, 5735.

(46) Martin, R. L.; Simon, C. M.; Smit, B.; Haranczyk, M. *J Am Chem Soc* **2014**, *136*, 5006.

(47) Krishna, R.; Long, J. R. *J Phys Chem C* **2011**, *115*, 12941.

(48) Krishna, R.; van Baten, J. M. *Phys Chem Chem Phys* **2011**, *13*, 10593.

(49) Franceschetti, A.; Zunger, A. *Nature* **1999**, *402*, 60.

(50) Drautz, R.; Diaz-Ortiz, A.; Fahnle, M.; Dosch, H. *Phys Rev Lett* **2004**, *93*.

(51) Abraham, N. L.; Probert, M. I. J. *Phys Rev B* **2006**, *73*.

(52) Bligaard, T.; Andersson, M. P.; Jacobsen, K. W.; Skriver, H. L.; Christensen, C. H.; Norskov, J. K. Mrs Bull 2006, 31, 986.

(53) d'Avezac, M.; Zunger, A. *Phys Rev B* **2008**, *78*.

(54) Yao, Y.; Tse, J. S.; Tanaka, K. *Phys Rev B* **2008**, *77*.

(55) Ceder, G.; Chiang, Y. M.; Sadoway, D. R.; Aydinol, M. K.; Jang, Y. I.; Huang, B. Nature 1998, 392, 694.

(56) Hautier, G.; Jain, A.; Chen, H. L.; Moore, C.; Ong, S. P.; Ceder, G. J Mater Chem 2011, 21, 17147.

(57) Mueller, T.; Hautier, G.; Jain, A.; Ceder, G. *Chem Mater* **2011**, *23*, 3854.

(58) Burkhardt, S. E.; Lowe, M. A.; Conte, S.; Zhou, W. D.; Qian, H. L.; Rodriguez-Calero, G. G.; Gao, J.; Hennig, R. G.; Abruna, H. D. Energ Environ Sci 2012, 5, 7176.

(59) Hautier, G.; Jain, A.; Ong, S. P.; Kang, B.; Moore, C.; Doe, R.; Ceder, G. Chem Mater 2011, 23, 3495.

(60) Yang, K. S.; Setyawan, W.; Wang, S. D.; Nardelli, M. B.; Curtarolo, S. Nat Mater 2012, 11, 614.

(61) Virshup, A. M.; Contreras-Garcia, J.; Wipf, P.; Yang, W. T.; Beratan, D. N. J Am Chem Soc 2013, 135, 7296.

(62) von Lilienfeld, O. A.; Lins, R. D.; Rothlisberger, U. *Phys Rev Lett* **2005**, *95*.

(63) von Lilienfeld, O. A.; Tuckerman, M. E. *J Chem Theory Comput* **2007**, *3*, 1083.

(64) Hu, X. Q.; Beratan, D. N.; Yang, W. T. *J Chem Phys* **2008**, *129*.

(65) Hu, X. Q.; Beratan, D. N.; Yang, W. T. *J Chem Phys* **2009**, *131*.

(66) von Lilienfeld, O. A. *International Journal of Quantum Chemistry* **2013**, *113*, 1676.

(67) Montavon, G.; Rupp, M.; Gobre, V.; Vazquez-Mayagoitia, A.; Hansen, K.; Tkatchenko, A.; Müller, K.-R.; von Lilienfeld, O. A. New Journal of Physics 2013, 15, 095003.

(68) Setyawan, W.; Curtarolo, S. *Comp Mater Sci* **2010**, *49*, 299.

(69) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Comp Mater Sci 2013, 68, 314.

(70) Wang, L.-P.; Titov, A.; McGibbon, R.; Liu, F.; Pande, V. S.; Martínez, T. J. Nat Chem 2014, 6, 1044.

(71) Sun, Z.; Winschel, G. A.; Zimmerman, P. M.; Nagorny, P. Angewandte Chemie International Edition 2014, 53, 11194.

(72) Rappoport, D.; Galvin, C. J.; Zubarev, D. Y.; Aspuru-Guzik, A. J Chem Theory Comput 2014, 10, 897.

(73) Zimmerman, P. M. *Journal of computational chemistry* **2013**, *34*, 1385.

(74) Maeda, S.; Morokuma, K. *J Chem Theory Comput* **2012**, *8*, 380.

(75) Holland, J. H. Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence; U Michigan Press, 1975.

(76) Cramer, N. L. “A representation for the adaptive generation of simple sequential programs”; Proceedings of the First International Conference on Genetic Algorithms, 1985.

(77) Becke, A. D. *J Chem Phys* **1993**, *98*, 1372.

(78) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem Phys Lett* **2004**, *393*, 51.

(79) Ufimtsev, I. S.; Martinez, T. J. *J Chem Theory Comput* **2008**, *4*, 222.

(80) Ufimtsev, I. S.; Martinez, T. J. *J Chem Theory Comput* **2009**, *5*, 1004.

(81) Ufimtsev, I. S.; Martinez, T. J. *J Chem Theory Comput* **2009**, *5*, 2619.

(82) Isborn, C. M.; Luehr, N.; Ufimtsev, I. S.; Martínez, T. J. J Chem Theory Comput 2011, 7, 1814.

(83) See Supplementary Material Document for the genes that fall on the Pareto optimal frontier, a complete list of the fittest genes from the GA optimizations, and a table of genes thought to be more synthetically accessible. For information on Supplementary Material, see; http://www.aip.org/pubservs/epaps.html.

(84) Tang, C.; Bi, R.; Tao, Y.; Wang, F.; Cao, X.; Wang, S.; Jiang, T.; Zhong, C.; Zhang, H.; Huang, W. Chem Commun 2015, 51, 1650.

Chapter 18

Efficient Prediction of Physical Properties Based On Machine Learning Techniques

"The single flower contains more brightness than a hundred flowers"

-Yasunari Kawabata (1899-1972)

In Chapter 17, we have applied genetic algorithm to efficiently search a very big molecular configuration space to find the optimal candidates for organic light emitting diodes (OLEDs) emitters based on thermally activated delayed fluorescence (TADF) mechanism. The most time consuming step in such material optimization process is to evaluate the chemical and physical properties. The property evaluation step requires *ab initio* quantum chemistry calculations. Hence, an efficient property evaluation step can be very useful. In this chapter, I will introduce some work we have done about how we applied machine learning techniques to build a model and provide us a fast property evaluation step.

Machine learning is becoming an important filed of research with extensive industrial applications in the recent decades applied in data analysis[1-6] and artificial intelligence.[7-11] Especially the recent news of AlphaGo beating Lee Se-dol a famous 9-

dan Go game player, the learning algorithm attracted world attention.[12] Machine learning techniques solve several categories of problems.[13] The regression problem is a basic type of supervised machine learning problem; the clustering problem is a basic type of unsupervised machine learning problem; and visualization of the data from very high dimensions. That is to say, the goal of supervised learning algorithm is to obtain the map from input to output, while the goal of unsupervised learning algorithm is to find out the structure of the input data. For example, given $y(x) = 2x + \varepsilon$, with ε represents small random variation. Every time we calculate a *y* value based on *x* and a random small ε value, we will have one point on the *x*-*y* scatter plot. Such that we can fit these limited given data with the linear relation assumption, found the relation between *x* and *y* as $y(x) = 2x$. With that in hand, we will be able to predict other *y* values given an *x* value. Thus, the heart of the machine learning problem can be treated as a prediction of the future with the past. We are not going to introduce the whole field here as it is impossible to make it into one chapter, we recommend the readers who are interested refer to the textbooks.[13-16]

The goals of the current chapter are, 1). Replace the quantum chemistry calculation with a simple supervised learning algorithm, 2). Explore the structure similarities in promising candidates for a further understanding of the structure of the data with visualization from high dimensions.

18.1 Geometrical Many-Body Expansion Model And Molecular Representation

18.1.1 Geometrical many-body expansion model

It is noteworthy at the beginning that the current model is built to treat the same problem we tried to solve in Chapter 17. And hence every molecule is represented as a tree structure with one branching layer. The extension of the current model to multi-layer tree structure will be trivial. As has been introduced in Chapter 17, the molecules in our calculated database are composed of a root fragment (will be simply denoted as root in the following text) and some fragments linked to the root. The choices of root and fragments are limited as shown in Figure 17-2. Figure 18-1 shows the graphical representation of a molecule which is composed of a central root, and six fragments (A, B, A, C, E, B).

Because the fragments are all limited to certain choices, the molecular configuration space is limited as well. If we treat each fragment or root as a "*body*", we can build a geometrical many-body expansion to specifically represent a certain molecule. The chemical and physical properties of the molecules can be written as a function of such many-body expansion, and hence the fitness value. The many-body expansion $F(\sigma)$ can be written as,

$$F(\sigma) = \sum_i C_i^1 \sigma_i^1 + \sum_{i,j} C_{ij}^2 \sigma_{ij}^2 + \sum_{i,j,k} C_{ijk}^3 \sigma_{ijk}^3 + \ldots \quad (18\text{-}1)$$

where C_i^1 , C_{ij}^2 and C_{ijk}^3 are one-body, two-body and three-body interaction coefficients respectively. We can have four-body, five-body, etc. interactions as well.

The maximum number of N for the N-body interactions will be the available positions of a certain root can have, with those positions are where the fragments linked to. For example, benzene root shown in Figure 17-2 have six positions, and hence a complete representation of a molecule that employs benzene as root can expand one-body, two-body until six-body interactions. The many-body coefficients represent how important of a certain kind of interaction is to the function $F(\sigma)$. The counter σ_i^1, σ_{ij}^2, σ_{ijk}^3 are used to indicate whether and how many times of certain type of many-body interactions exists in the molecule. Here we also use σ to represent all the counters. For example, molecule [Bz[Cy][Cl][Cy][Cz][Cy][Cz]] has one-body interaction Cy, Cl, Cy, Cz, Cy, Cz, hence $\sigma_{Cy}^1 = 3$, $\sigma_{Cz}^1 = 2$ and $\sigma_{Cl}^2 = 1$. While for the same molecule, as there is no [Pt] fragment, and hence $\sigma_{Pt}^1 = 0$. The type two-body interactions for this molecule can be extracted easily as well, for example, $\sigma_{CyCy}^2 = 3$. The graphical representative one-body interaction σ_A^1, two-body interaction σ_{BA}^2 and three-body interaction σ_{EBA}^3 are shown in Figure 18-1 as an example.

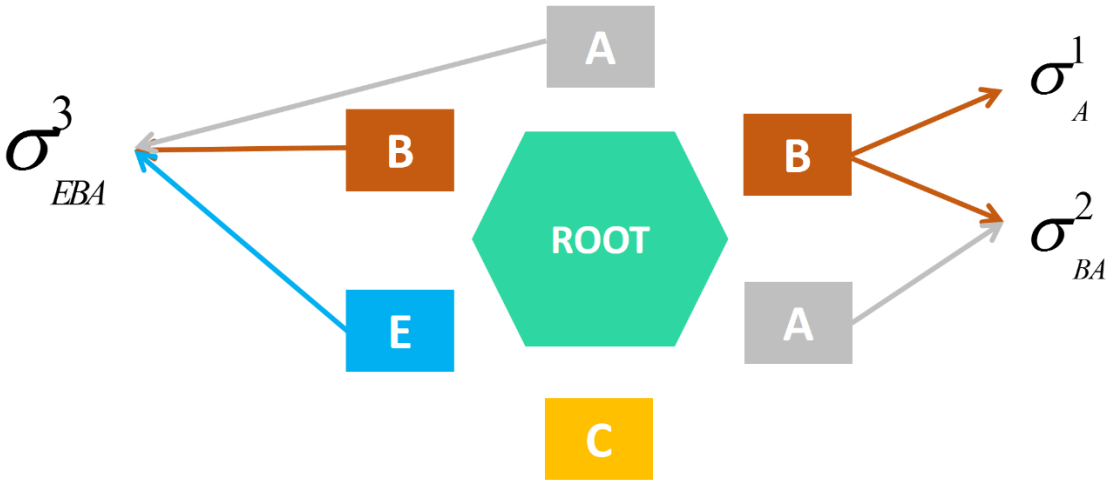


Figure 18-1. Graphical representation of a molecule with root and six fragments (A, B, A, C, E, B) linked to the root. The $1\sigma_B$, $2\sigma_{BA}$ and $3\sigma_{EBA}$ represent the one-body, two-body and three-body interactions.

With this geometrical many-body representation, every molecule in our molecular configuration space can be represented by a vector. For example, if we are considering one-body interaction only for simplicity, then, every molecule is represented as a vector with dimension of 16 (we include root here, 6 different roots plus 10 different fragments). If the vector is chosen to be represented in an order as $\sigma^1_{Bz}\sigma^1_{Pd}\sigma^1_{Pz}\sigma^1_{Fr}\sigma^1_{Tp}\sigma^1_{Pr}\sigma^1_{Cy}\sigma^1_{Ah}\sigma^1_{Ca}\sigma^1_{Fm}\sigma^1_{Dc}\sigma^1_{Cz}\sigma^1_{Pt}\sigma^1_{Cl}\sigma^1_{Ti}\sigma^1_{Id}$, then the previously discussed molecule [Bz[Cy][Cl][Cy][Cz][Cy][Cz]] can be written as the vector 1000003000020100. We will call this vector as many-body interaction vector in the following text. And hence each molecule is uniquely represented in this vector space spanned by the types of many-body interactions. It is noteworthy here that as long as we are considering truncated

expansion model (this will be discussed in detail in the next section), one molecule can be represented as a unique vector. However, one vector can represent multiple molecules in certain cases.

The chemical and physical properties, can be written as a functional of the geometrical many-body expansion. For example, the S_1-T_1 energy gap $E\left(S_1 - T_1\right)$ and S_1-S_0 transition dipole moment $D\left(S_1 - S_0\right)$ can be written as,

$$E\left(S_1 - T_1\right) = E\left[F\left(\sigma\right)\right] \tag{18-2}$$

$$D\left(S_1 - S_0\right) = D\left[F\left(\sigma\right)\right] \tag{18-3}$$

And hence, because of the structure-function relation, if we know the exact many-body interaction coefficients, for example, C_i^1, C_{ij}^2 and C_{ijk}^3, as well as the exact functional form, we will be able to obtain the same results as we calculated from *ab initio* quantum chemistry methods.

18.1.2 Expansion truncation

Because the maximum positions of a root in Figure 17-2 is six (Bz has 6 positions, while others, for example, Pd has 5, Fr as 4), we will not go beyond six-body interactions in order to completely represent such a molecule in the current molecular configuration space. However, even we just have maximum up to six-body interactions, the types of such interactions will be an astronomical number. For example, since we have 10 different choices of fragments (excluding roots), the number of two-body interactions will be 100 (as 10 times 10), the number of three-body interactions will be 1000, and hence

the number of six-body interactions will be one million. If a molecule is represented completely, then to fit this one million many-body interaction coefficients requires database contains the data of all the molecules in the molecular configuration space, which makes such effort of building a model with fast property evaluation step meaningless. Hence, a truncation of the geometrical many-body expansion is necessary. Then following question is where we should truncate? It is important to consider the size of the existing database. As has been discussed above, we need to know the form of functional and many-body interactions coefficients to calculate the chemical and physical properties of the molecule. With the database generated by quantum chemistry calculations, we will be able to fit those many-body interactions coefficients with a known functional form. And hence this becomes a supervised learning problem. To avoid overfitting, the input data should be a couple of times bigger than unknown parameters to fit.[13] In our case, there exist about 7000 calculations, and hence the unknown parameters should stay around 1000 or even less. Thus, with this consideration, we should truncate the many-body interactions up to three-body interaction, as there exist 1000 three-body interactions and 100 two-body interactions if we consider fragments only.

18.1.3 Generating the many-body interaction vector

When to build the many-body interaction vector, there are multiple ways to consider certain types of many-body interactions. Let us employ the molecule in Figure 18-1 as

an example. If we count nearest neighbor types of interactions only, we can have three many-body interactions, for example, ABA, CEB, etc. We will not have three-body interactions like ABC or AAC. The many-body interaction counter code is implemented in a general way such that we can define the allowed differences between the pair of positions for different many-body interactions. For example, if we are considering two-body interactions, then there are one pair of positions in two-body interactions. We can define a nearest neighbor counting way, which means the difference should be 1. We can define a way of counting that allows 2 position differences, for example, considering the molecule in Figure 18-1 again; this way of counting allows we count two-body interactions AB, AA, BA, BC, etc. While two-body interactions like BE, etc. will not be counted as the fragments B and E have 3 position difference. It is noteworthy here that we distinguish the interaction order, that is to say, many-body interaction BE and EB are different. For three-body interactions, there are two pairs of positions (the pair of first and second positions, and the pair of second and third positions). Considering the same molecule as an example, nearest neighbor allows three-body interactions, for instance, ABA, BAC, etc. While three-body interactions like ABC, AAC, etc. are not allowed.

In this chapter, we will show results regarding the full counting, which means we allow arbitrary differences between a pair of positions. This type of counting procedure provides a non-sparse many-body interaction vector, however, it fails to distinguish between some molecules. For example, considering the molecule

[[root][A][B][A][B][A][B]] and [[root][A][A][A][B][B][B]], the full counting procedure can not distinguish between these two molecules. While employing the nearest neighbor counting procedure can distinguish these two molecules. On the other side, nearest neighbor counting procedure increases the sparsity of the many-body interaction vector as well. This is because the dimension of the vector space is completely decided by the many-body interaction expansion truncation. For example, let us consider the molecule [Bz[Cy][Cl][Cy][Cz][Cy][Cz]]. If we employ nearest neighbor interactions only, then two-body interactions such as σ^2_{CyCy} will be 0. However, as long as we are considering two-body interactions, the length of the many-body interaction vector will not be shortened. And hence applying the nearest neighbor or other truncated way to count the types of many-body interactions will increase the sparsity of the vector. The nearest neighbor counting procedure will be an interesting topic to study in the future.

18.1.4 The Representation Model

As we have seen from section 18.1.1, the functional of geometrical many-body expansion used to calculate chemical and physical properties can be arbitrary complex to achieve the exact result. However, to simplify our calculation, we will employ a unit functional form here. Combined with the fact we should not go beyond three-body interaction, the S_1-T_1 energy gap $E\left(S_1 - T_1\right)$ and S_1-S_0 transition dipole moment $D\left(S_1 - S_0\right)$ can be represented as,

$$E\left(S_1 - T_1\right) = \sum_i C_i^{1,E}\sigma_i^1 + \sum_{i<j} C_{ij}^{2,E}\sigma_{ij}^2 + \sum_{i<j<k} C_{ijk}^{3,E}\sigma_{ijk}^3 \quad (18\text{-}4)$$

$$D\left(S_1 - S_0\right) = \sum_i C_i^{1,D}\sigma_i^1 + \sum_{i<j} C_{ij}^{2,D}\sigma_{ij}^2 + \sum_{i<j<k} C_{ijk}^{3,D}\sigma_{ijk}^3 \quad (18\text{-}5)$$

And both $E\left(S_1 - T_1\right)$ and $D\left(S_1 - S_0\right)$ reduces to a function of many-body interaction vector. The subscript *E* and *D* in eq.(18-4) and (18-5) are employed to distinguish the many-body interaction coefficients for S_1-T_1 energy gap and S_1-S_0 transition dipole moment. In our model, we consider one-body interactions with both root and fragments, and hence there exist 16 one-body interactions. The two-body and three-body interactions are considered among fragments only. Thus the total number of two-body and three-body interactions are 100 and 1000 respectively. If we consider a model up to two-body interactions, the molecules are represented in a vector space with 116 dimensions. Thus, the vector in three-body interaction model space will have 1116 dimensions.

18.2 Supervised Learning

18.2.1 Supervised linear learning problem

To obtain the many-body interaction coefficients, we need to fit our model with the existing database. And hence this transforms the problem to a linear regression problem. The linear regression problem is shown in the following two equations in the matrix form,

$$\begin{bmatrix} \boldsymbol{\sigma}(\mathbf{1})^{\mathbf{1,E}} & \boldsymbol{\sigma}(\mathbf{1})^{\mathbf{2,E}} & \boldsymbol{\sigma}(\mathbf{1})^{\mathbf{3,E}} \\ \boldsymbol{\sigma}(\mathbf{2})^{\mathbf{1,E}} & \boldsymbol{\sigma}(\mathbf{2})^{\mathbf{2,E}} & \boldsymbol{\sigma}(\mathbf{2})^{\mathbf{3,E}} \\ \vdots & \vdots & \vdots \\ \boldsymbol{\sigma}(\mathbf{n})^{\mathbf{1,E}} & \boldsymbol{\sigma}(\mathbf{n})^{\mathbf{2,E}} & \boldsymbol{\sigma}(\mathbf{n})^{\mathbf{3,E}} \end{bmatrix} * \begin{bmatrix} \mathbf{C}^{\mathbf{1,E}} \\ \mathbf{C}^{\mathbf{2,E}} \\ \mathbf{C}^{\mathbf{3,E}} \end{bmatrix} + \begin{bmatrix} \varepsilon(1)^{E} \\ \varepsilon(2)^{E} \\ \vdots \\ \varepsilon(n)^{E} \end{bmatrix} = \begin{bmatrix} E(S_1 - T_1)(1) \\ E(S_1 - T_1)(2) \\ \vdots \\ E(S_1 - T_1)(n) \end{bmatrix} \quad (18\text{-}6)$$

$$\begin{bmatrix} \boldsymbol{\sigma}(\mathbf{1})^{\mathbf{1,D}} & \boldsymbol{\sigma}(\mathbf{1})^{\mathbf{2,D}} & \boldsymbol{\sigma}(\mathbf{1})^{\mathbf{3,D}} \\ \boldsymbol{\sigma}(\mathbf{2})^{\mathbf{1,D}} & \boldsymbol{\sigma}(\mathbf{2})^{\mathbf{2,D}} & \boldsymbol{\sigma}(\mathbf{2})^{\mathbf{3,D}} \\ \vdots & \vdots & \vdots \\ \boldsymbol{\sigma}(\mathbf{n})^{\mathbf{1,D}} & \boldsymbol{\sigma}(\mathbf{n})^{\mathbf{2,D}} & \boldsymbol{\sigma}(\mathbf{n})^{\mathbf{3,D}} \end{bmatrix} * \begin{bmatrix} \mathbf{C}^{\mathbf{1,D}} \\ \mathbf{C}^{\mathbf{2,D}} \\ \mathbf{C}^{\mathbf{3,D}} \end{bmatrix} + \begin{bmatrix} \varepsilon(1)^{D} \\ \varepsilon(2)^{D} \\ \vdots \\ \varepsilon(n)^{D} \end{bmatrix} = \begin{bmatrix} D(S_1 - S_0)(1) \\ D(S_1 - S_0)(2) \\ \vdots \\ D(S_1 - S_0)(n) \end{bmatrix} \quad (18\text{-}7)$$

where superscript *E* and *D* denotes the corresponding fitting problem for S_1-T_1 energy gap and S_1-S_0 transition dipole moment respectively. The number in the parenthesis $1, 2, \ldots, n$ are the indices for input data. The $\mathcal{E}$ is the error term. $\boldsymbol{\sigma}^{\mathbf{1}}$, $\boldsymbol{\sigma}^{\mathbf{2}}$ and $\boldsymbol{\sigma}^{\mathbf{3}}$ represent one-body interaction, two-body interaction, and three-body interaction vectors respectively. $\mathbf{C}^{\mathbf{1}}$, $\mathbf{C}^{\mathbf{2}}$ and $\mathbf{C}^{\mathbf{3}}$ represent one-body, two-body and three-body interaction coefficients column vectors respectively. Once we have decided the many-body interaction expansion truncation, and build the many-body interaction vector for each molecule, with the database, we can fit the many-body interaction coefficients by minimizing the error.

18.2.2 Cross-Validation for two-body interaction model

Our goal is to test the accuracy of our model, without large amount of new quantum chemistry calculations. While the error for the whole training set (training set is the set of molecules where used to fit the many-body interaction coefficients) cannot guarantee the error will still be low when applied to new test set. With this consideration, we

applied cross-validation procedure[17] to test our geometrical many-body expansion model. Cross-validation procedure divides the existing database into a training set and a validation set. The training set is used to fit the coefficients, and the validation set will be used to test the error of the fitted coefficients. And hence with this cross-validation procedure, there is no need for building a new database. Figure 18-2 shows the root mean square error (RMSE) of the whole set (including both training and validation set) as a function of the size of the training set by considering a model up to two-body interactions. As we can see that the RMSE reduces as the size of the training set increases. At about training set equals 1000 molecules, the RMSE changes little when to increase the size of the training set. We can reduce the RMSE for S_1-T_1 energy gap bellow 0.01 Hartree and the RMSE for S_1-S_0 transition dipole moment is reduced to around 0.8 Debye. This is still a relatively large error, especially for transition dipole moment. This can be explained as the two-body interaction model is not enough to describe the symmetry of the molecule. Figure 18-3 shows two molecules with completely different symmetry and assume there exists different electron negativity for fragments A and B (and they should be in reality), these two molecules should have different dipole moments. However, in our two-body interaction model, considering the full counting procedure, they are represented by the same many-body interaction vector. This indicates the fact nearest neighbor counting may be important in describing the transition dipole moment. The supervised linear regression is performed with R package.[18]

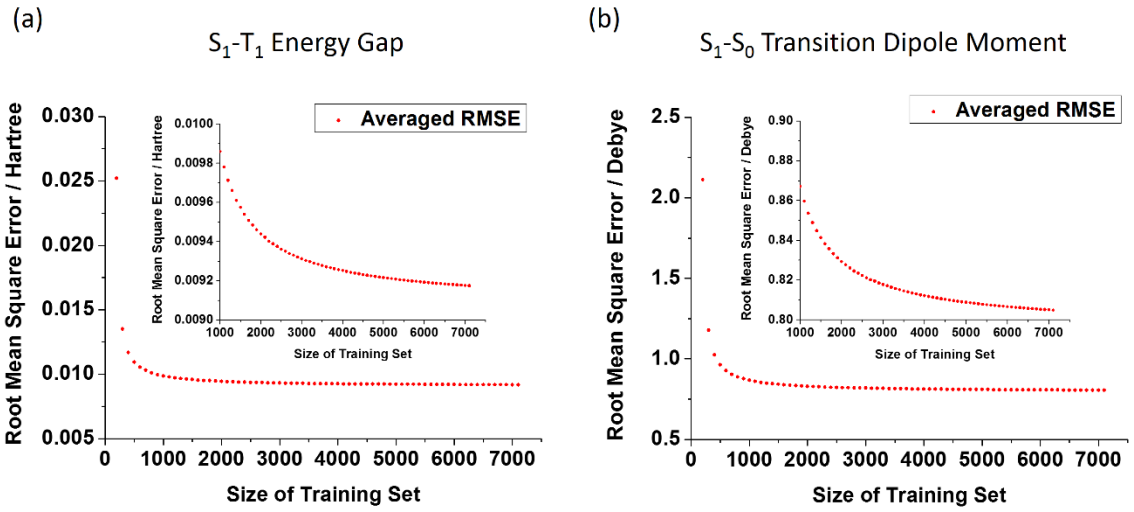


Figure 18-2. The root mean square error as a function of the size of the training set for (a) S_1-T_1 energy gap and (b) S_1-S_0 transition dipole moment with a model up to two-body interactions.

18.2.3 Three-body interaction model

Let us discuss how three-body interactions model behaves. Iit is unfortunate that we do not have enough data in the database to employ the cross-validation procedure to avoid the overfitting. And hence we utilize all our data to compare the fitted fitness values with the calculated fitness values. That is to say, we are only considering the training set here. The fitted fitness values are calculated according to the fitness function introduced in Chapter 17 with the S_1-T_1 energy gap and S_1-S_0 transition dipole moment based on the fitted many-body interaction coefficients. Figure 18-4 shows the scatter plot for fitted fitness values based on our many-body interaction coefficients with

the calculated fitness values. The whole set S_1-T_1 energy gap has RMSE as 0.008 Hartree with the similar order as those predicted by two-body. While the whole set RMSE for S_1-S_0 transition dipole moment reduces to 0.65 Debye as compared to that with two-body interaction model 0.8 Debye. A fitted linear function is shown in Figure 18-4 as well to indicate the relative accurate identity dependency of fitted and calculated fitness value. Figure 18-5 shows the distribution of the residual between calculated and fitted fitness values. It can be seen that most of the molecules can be predicted within 0.1 error in fitness value. Indicate the three-body interaction model should provide a qualitative good candidates for OLED emitters based on TADF mechanism. It is noteworthy that the evaluation of the fitness value of such a molecule with fitted many-body interaction coefficients much faster as compare to quantum chemistry calculations.

18.2.4 Model improvement

Results show that three-body interactions model still exist about 0.1 error in fitness value evaluation. This can be attributed to the following possible reasons. First, a nearest neighbor counting procedure is required to fully distinguish the two molecules shown in Figure 18-3. And hence the current three-body interactions model employing the full counting procedure lacks of such capability. Second, the unity functional form may be too simple to illustrate the relationship between geometry and chemical or physical properties. Third, the current database is generated by genetic algorithm and hence it has the tendency to compute more “good” molecules than “bad” molecules.

And hence a larger database is required to allow the variety in the database such that different fragments can achieve a relatively equal percentage.

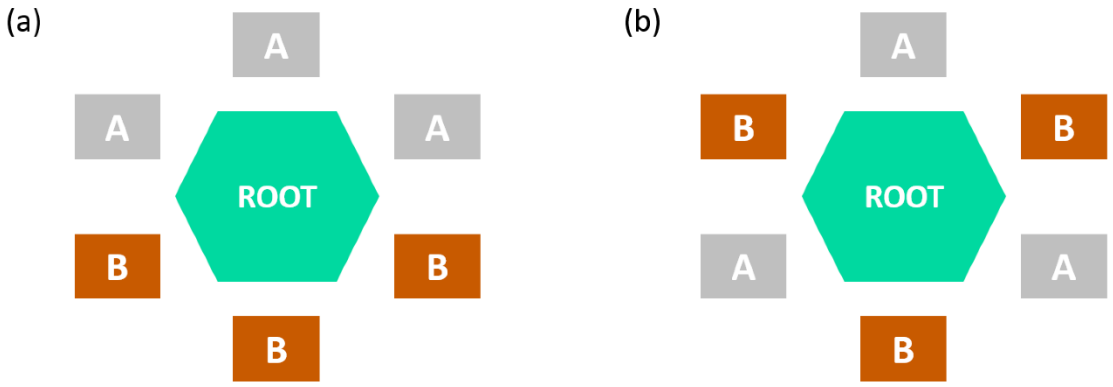


Figure 18-3. The graphical represented molecules with different symmetry (a) and (b). However, in the two-body interaction model, they are represented by the same many-body interaction vector.

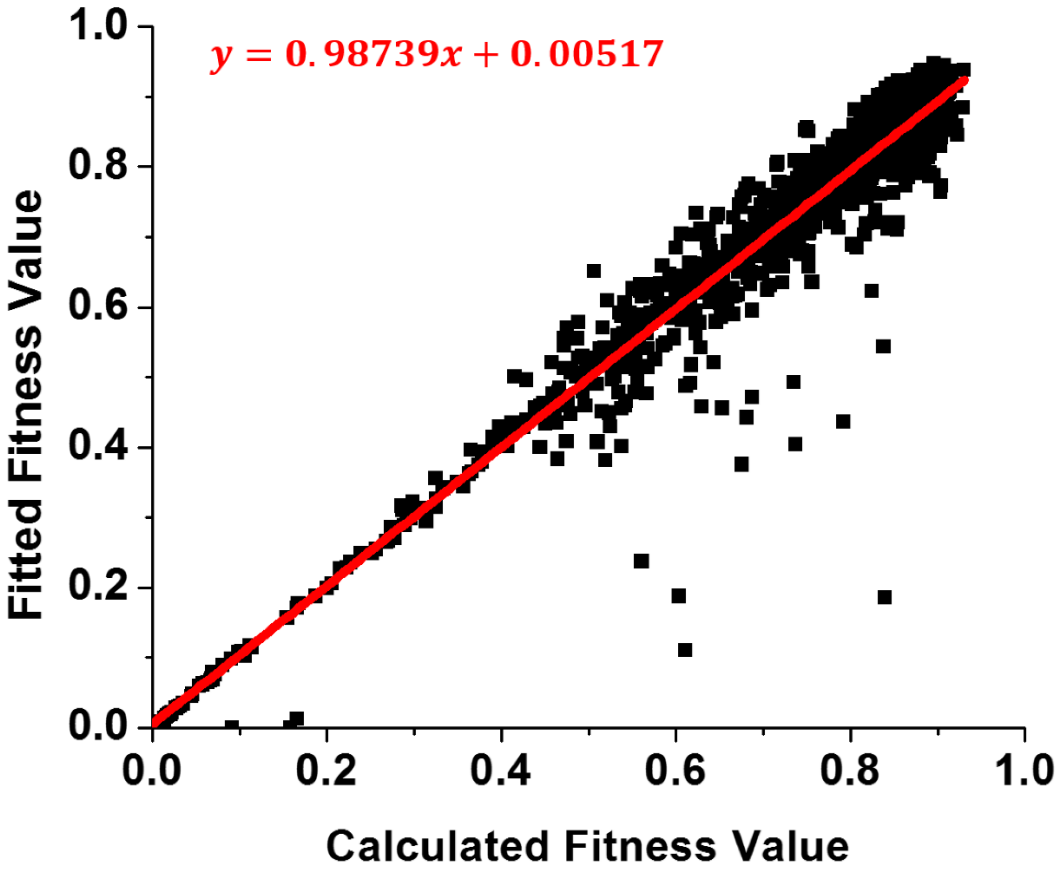


Figure 18-4. The scatter plot of fitted fitness value and calculated fitness value as shown in black dots for individual molecules. The red line is a fitted line with r^2=0.941.

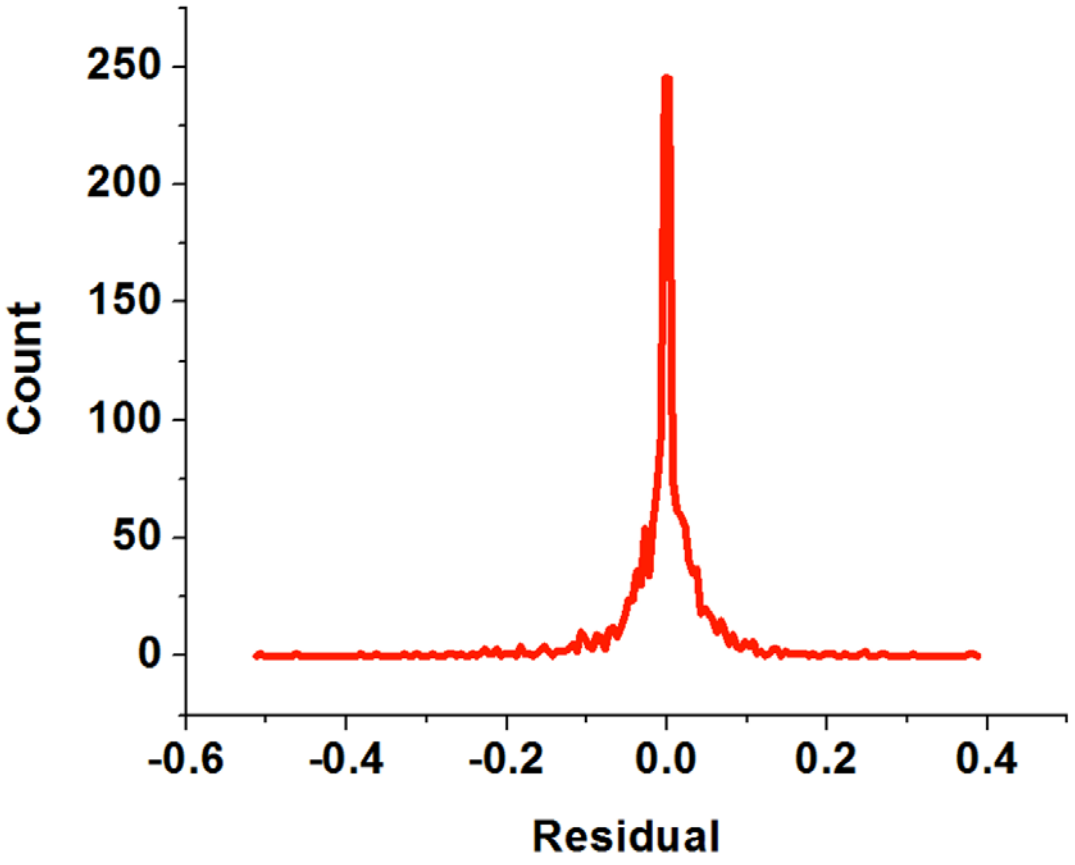


Figure 18-5. The distribution of the residual between fitted and calculated fitness values.

18.3 Visualization of The High Dimensional Data with The Geometrical Many-Body Expansion Model

The current database contains molecules that are hard to extract their structure similarities all at once through a graph such that a direct visualization of the structure of the data can be achieved. The visualization of high dimensional data is important in machine learning as it helps to extract the feature of the data.[19-21] In this section, I will show an example that with the help of the geometrical many-body expansion model, we

will be able to extract the molecular geometrical similarities by visualizing the data directly.

The geometrical many-body expansion model is able to project each molecule into a high dimensional space, and hence a metric space is built with the help of the model. The distance between the molecules can be measured directly. We computed the Euclidean distance between each pair of the molecules and generated the distance matrix. For the purpose of visualization, this high dimensionality has to be reduced. This is achieved by employing the principle component analysis.[22-24] We project the whole matrix onto the first two principle components. This is shown in Figure 18-6. Each dot in Figure 18-6 represents a molecule with the color indicates its fitness value. We can see that the molecules are separated into subgroups. We believe those molecules in the same group should share the similar structures as this is a visualization for distance matrix which represents the geometrical similarities in the geometrical many-body expansion model. Figure 18-7 shows the same graph with two representative zooms in part with corresponding randomly selected molecules and associated fitness values. The left hand side zoom in part shows that most molecules employ pyridine as root, and all the selected molecules have thieno[3,2,-b]indolyl fragment. The selected molecules in the right hand side zoom in part are almost identical, as can be seen from the list that they all employs benzene as root and are all composed of cyano, carbazolyl and indolyl groups. These subgroups are useful regarding many considerations, for example, 1) A much more accurate set of many-body interaction coefficients can be fitted by utilizing the molecules

in subgroups as they share the similar geometries. Hence, the fitted many-body interaction coefficients are believed to be superior in predicting the fitness values of the molecules nearby to that fitted from a whole set of molecules; 2) Each subgroup can be treated as a center, by developing models, for example, very simple linear extrapolation model, we are able to predict the fitness values of molecules that are not calculated between these centers.

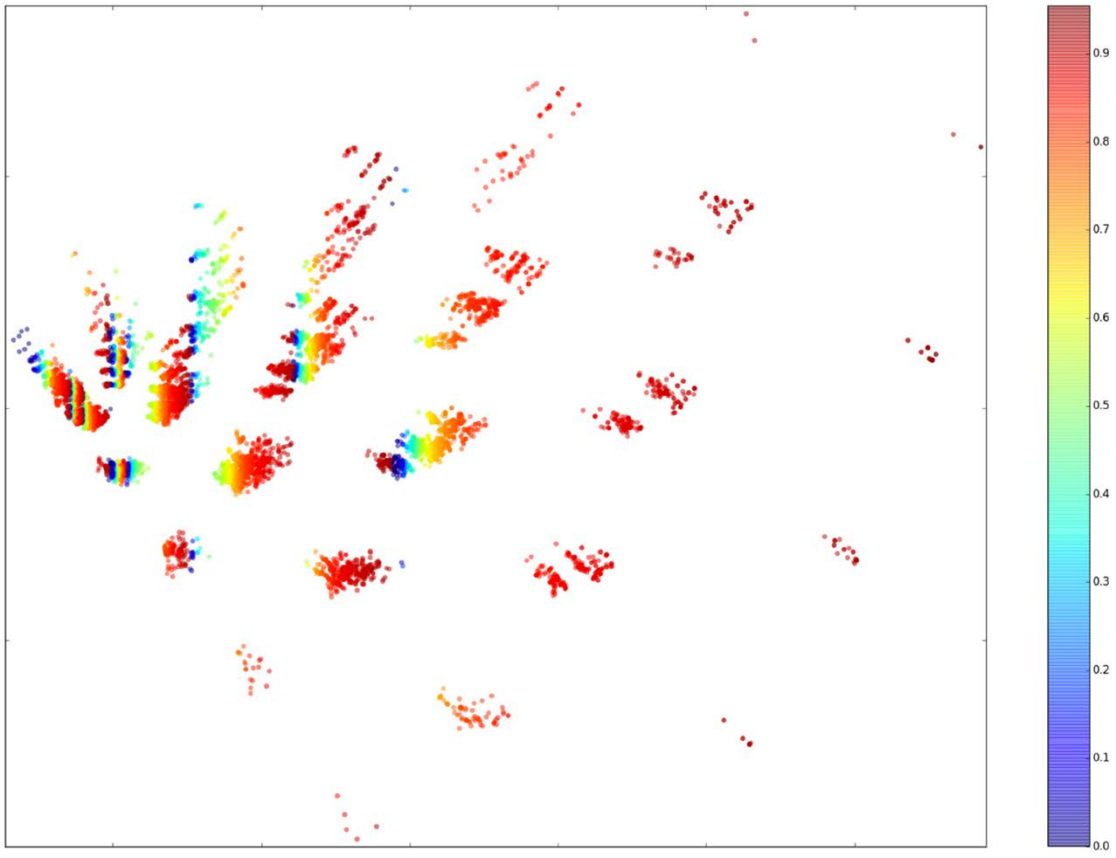
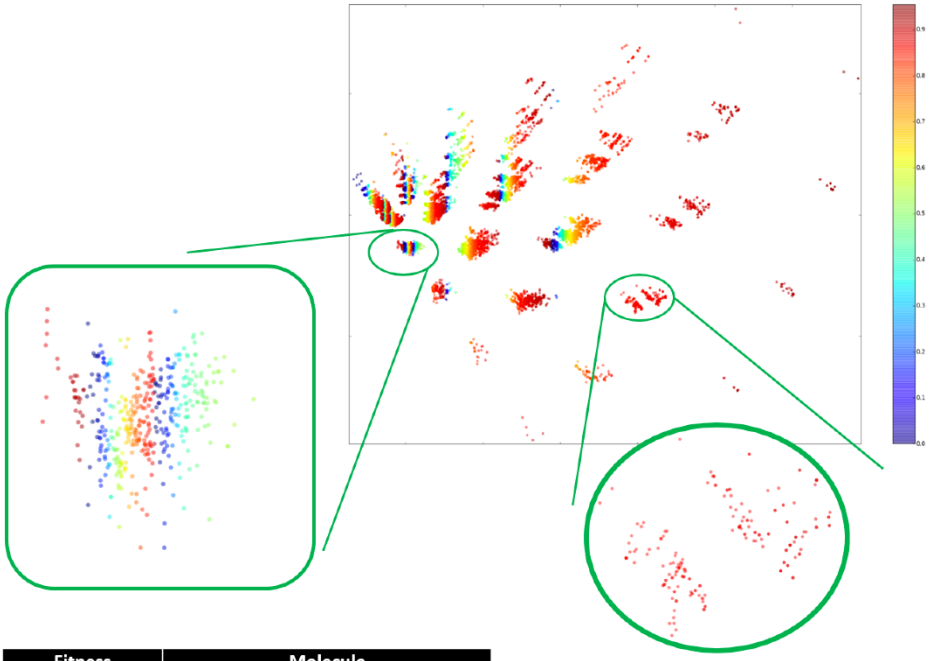


Figure 18-6. A projected distance matrix of the molecules on the first two principle components with the color represents the fitness value of the molecule.



Fitness	Molecule
0.839353	[Pd[Ti][Cl][Id][Cy][Cl]]
0.845593	[Pd[Cl][Ti][Cz][Ti][Cy]]
0.856707	[Pd[Cz][Cl][Id][Cy][Ti]]
0.893524	[Pd[Cz][Ti][Dc][Id][Cz]
0.025987	[Fr[Cz][Ti][Ah][Cz]

Fitness	Molecule
0.889724	[Bz[Cy][Cz][Id][Cz][Cy][Cz]]
0.894456	[Bz[Cy][Cz][Cy][Cz][Cy][Id]]
0.914773	[Bz[Cy][Id][Cz][Cy][Cz][Cz]]
0.889923	[Bz[Cy][Cz][Cz][Id][Cy][Cz]]
0.892489	[Bz[Id][Cy][Cz][Cz][Cz][Cy]]
0.893132	[Bz[Cz][Id][Cz][Cy][Cz][Cy]]
0.890635	[Bz[Cz][Id][Cy][Cz][Cy][Cz]]

Figure 18-7. A projected distance matrix of the molecules on the first two principle components with the color represents the fitness value of the molecule. The two zoom in part of the graph is shown with randomly selected molecules and their fitness values.

18.4 Conclusions

In this work, we have presented geometrical many-body expansion model that can be employed to uniquely project each of the molecules in our calculated database on to a high dimensional space. The position of the molecule in that high dimensional space

is decided by its molecular configurations, specifically, the root and the fragments composed of the molecules. Due to the structure function relationship, the molecular configuration of the molecule should have a definitive impact on the physical and chemical properties. And hence the physical and chemical properties of the molecules are decided by its position in that high dimensional space. With that in mind, we employed supervised learning algorithm to obtain the impact of each dimension on the properties of the molecules based on our database. The impact of each dimensional on the properties of the molecules is represented by the many-body interaction coefficients in our discussion and is a real-valued number. With the many-body interaction coefficients and the position of the molecule in that high dimensional space, the physical and chemical properties of the molecules can be calculated directly without *ab initio* quantum chemistry calculation. This means thousands of times speed up. We have shown that by employing a model up to three-body interaction, we will be able to reduce the error with 0.1 fitness value as compared with quantum chemistry results. However, the model can be improved by the nearest neighbor counting procedure, and is a direction of study in the future.

The geometrical many-body expansion model directly provides us a metric space for the molecules as well. With this metric space, we can find the structure similarities of the molecules by computing the distance matrix. Close distance indicates the similar structure. We also projected the distance matrix on to the first two principle component

and hence can be visualized directly. This further helps us to understand the structure of the data itself.

REFERENCES

REFERENCES

(1) Witten, I. H.; Frank, E. *Data Mining: Practical machine learning tools and techniques*; Morgan Kaufmann, 2005.

(2) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; von Lilienfeld, O. A. *J Chem Theory Comput* **2015**, *11*, 2087.

(3) Bubenik, P. *The Journal of Machine Learning Research* **2015**, *16*, 77.

(4) Carlsson, G. *Acta Numerica* **2014**, *23*, 289.

(5) Carlsson, G. *Bulletin of the American Mathematical Society* **2009**, *46*, 255.

(6) Ghrist, R. *Bulletin of the American Mathematical Society* **2008**, *45*, 61.

(7) Michalski, R. S.; Carbonell, J. G.; Mitchell, T. M. Machine learning: An artificial intelligence approach; Springer Science & Business Media, 2013.

(8) Sebastiani, F. *ACM computing surveys (CSUR)* **2002**, *34*, 1.

(9) Fayyad, U. M.; Piatetsky-Shapiro, G.; Smyth, P.; Uthurusamy, R. **1996**.

(10) Pomerleau, D. A. *Neural Computation* **1991**, *3*, 88.

(11) INSIDE GOOGLE'S QUEST TO POPULARIZE SELF-DRIVING CARS http://www.popsci.com/cars/article/2013-09/google-self-driving-car, 2013.

(12) News, B. Artificial intelligence: Google's AlphaGo beats Go master Lee Se-dol http://www.bbc.com/news/technology-35785875, 2016.

(13) Bishop, C. M. *Machine Learning* **2006**.

(14) James, G.; Witten, D.; Hastie, T.; Tibshirani, R. An introduction to statistical learning; Springer, 2013; Vol. 112.

(15) Friedman, J.; Hastie, T.; Tibshirani, R. The elements of statistical learning; Springer

series in statistics Springer, Berlin, 2001; Vol. 1.

(16) Anzai, Y. *Pattern Recognition & Machine Learning*; Elsevier, 2012.

(17) Kohavi, R. “A study of cross-validation and bootstrap for accuracy estimation and model selection”; Ijcai, 1995.

(18) Team, R. C. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2013.

(19) Roweis, S. T.; Saul, L. K. *Science* **2000**, *290*, 2323.

(20) Pudil, P.; Novovičová, J. Novel methods for feature subset selection with respect to problem knowledge. In Feature Extraction, Construction and Selection; Springer, 1998; pp 101.

(21) Lu, H.; Plataniotis, K. N.; Venetsanopoulos, A. N. *Pattern Recognition* **2011**, *44*, 1540.

(22) Wold, S.; Esbensen, K.; Geladi, P. Chemometrics and intelligent laboratory systems 1987, 2, 37.

(23) Abdi, H.; Williams, L. J. Wiley Interdisciplinary Reviews: Computational Statistics 2010, 2, 433.

(24) Moore, B. C. *Automatic Control, IEEE Transactions on* **1981**, *26*, 17.