

QUEEN MARY UNIVERSITY OF LONDON

DOCTORAL THESIS

**Computational Methods for
Photochemistry in Organic Molecular
Crystals**

Author:

Miguel RIVERA

Supervisor:

Dr. Rachel CRESPO-OTERO

*A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy*

in the

Department of Chemistry

June 3, 2020

Declaration of Authorship

I, Miguel Rivera, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signed: Miguel Rivera

Date: October 6, 2020

“We were working late on the time machine in the little makeshift lab upstairs. The moon was stuck like the whorl of a frozen fingerprint to the skylight. In the back alley, the breaths left behind by yowling toms converged into a fog slinking out along the streets. Try as we might, our measurements were repeatedly off. In one direction, we’d reached the border at which clairvoyants stand gazing into the future, and in the other we’d gone backward to the zone where the present turns ghostly with memory and yet resists quite becoming the past. We’d been advancing and retreating by smaller and smaller degrees until it had come to seem as if we were measuring the immeasurable. Of course, what we really needed was some new vocabulary of measurement. It was time for a break.”

Stuart Dybek, *Paper Lantern*

QUEEN MARY UNIVERSITY OF LONDON

Abstract

School of Biological and Chemical Sciences

Department of Chemistry

Doctor of Philosophy

Computational Methods for Photochemistry in Organic Molecular Crystals

by Miguel RIVERA

Organic molecular crystals hold great promise for photochemical applications such as photovoltaics, OLEDs or lasers. Understanding the excited states of these systems is particularly challenging due to the interconnectedness of the excited state character of the molecular wavefunctions and the environmental interactions within the condensed phase. The principal obstacle to this understanding is first and foremost the methodological gap between periodic simulation methods and molecular scale excited state formalisms. Multiscale cluster model methods are a viable extension of the latter, where they promise to recover the crystalline environmental interactions for excited state calculations whilst representing excitations as localised defects rather than periodic delocalised perturbations. This is pertinent to molecular crystals—periodic but non-covalently bonded, displaying quasilocal excitations—however, there is hitherto no software package dedicated their modelling as clusters.

The principal output of this project is the development and application of such a program. This thesis begins with a review of the theoretical background necessary to understand QM:QM' (Quantum Mechanics:Quantum Mechanics') cluster models for excited states, namely quantum electronic structure modelling methods, multi-scale embedding methods, and molecular photochemistry. Then, three results chapters are presented, each based on a publication carried out during the project. First, the cluster model method is described, where the electrostatic embedding of the QM region has been extended to model the Madelung potential. Then, the program (*fromage*, the FRamewOrk for Molecular AGgregate Excitations) which implements this method is presented, along with a suite of analysis tools and command line utilities which enable the reproducible investigation of cluster geometries in molecular crystals. Finally, *fromage* is used to model an array of luminescent materials and investigate their competing radiative and nonradiative excited state decay pathways.

Acknowledgements

This thesis summarises the work that I carried out between October 2016 and March 2020. Whilst it did cost an undeniably large amount of effort, CPU hours, and keystrokes, it was overall a pleasure and privilege, thanks to the support of a group of outstanding entities which I will proceed to list.

I give my thanks to the members of my panel at QMUL. Isaac Abrahams, for chairing meetings and Devis Di Tommaso for being my second supervisor, always supportive and available to talk about science or careers.

All of the workers of the Joseph Priestley building deserve a mention, for solidarity and good mood. Thank you for Friday drinks at the Senior Common Room on a Friday evening, coffee while waiting for a fire alarm to blow over, or miming conversations through the plexiglass that separates the wet lab and computational office.

Speaking of the latter, the Computational Chemists at QMUL were my closest colleagues for these past years, and I am grateful to all of them for their support. To Anna, Bob, Dimitri, Etienne, Fariha, Filip, Fu, Matteo, and Xiangwen thank you for conversations, lunches, teas, and compassion.

Now I reach the point of my research group, where individual mentions are in order. Scott and Warda stand out as a duo of pre-PhD students who exceeded every expectation but became friends and colleagues, despite my ruining their learning by trying to explain k -space to them. Sorry about that. Ljiljana has been a supportive and compassionate presence ever since she arrived in the UK, and will receive my unreserved thanks just as soon as she watches and returns my DVD of *Astérix et Obélix : Mission Cléopâtre*. Federico is only a recent addition, but has already proven himself a great deskmate, and a power user of fromage, for which I can only admire his taste. Amir has been an incredibly hard working and fast learning PhD-mate, and it has been a privilege to share some of my new expertise with him. Michael, thank you for providing an example of how to manage science and Rachel, and how to care only the right amount. Your bash-fu is unmatched amongst peers, as can only be evidenced by your program “cdjob”, which literally must have saved me about a day of my life so far in typing. Back of the net!

I would also like to thank the Thomas Young Centre organisers and community, for making me feel a part of a larger group of researchers, and allowing me to share hardships with so many peers. To Bethan, Matthias, Patrick, Sam, and many others, I would like to present my gratitude.

Here seems as good a point as any to thank the communities which made my technical life easy and rewarding during my PhD. The list of useful programs and forums is extensive, so I will only highlight my favourites: vim, Stack Overflow, vmd, Chemcraft, Jmol, Psi4, GNU/Linux, L^AT_EX, and the Python project.

Now for the extracurricular acknowledgments. I would like to extend my gratitude to my flatmates with whom I shared three-odd years in the Bethnal Green apartment; Dag, Devi, Jolien, and Merry (Emma, Lachlan and Yean get honorable mentions here). Your kindness and caring are only matched by the amount of disassembled mannequin parts scattered around your lair. More specifically, thank you: Dag for warmth, Devi for attentiveness, Jolien for care, and Merry for joy. Let the record show that the final stretch of this thesis is being written on Merry's antique laptop, which was delivered to me by Yean.

To my comrades in board games, roleplaying games, card games, and any other kind of games, thank you for late nights, laughs, camaraderie, and victories. Of those not yet mentioned, I add Dan, Emrys, Emy, Feargus, Hugo, Jasper, Oli, Ryan, and Sally.

Members of my musical sphere also deserve my gratitude, and I would like to thank for grooves, hummings, and wailings: Ben, Charlie, Harry, Jamie, Max, Mike, The Roundhouse Choir, and The Martinis.

Arthur and Matt must be singled out for being a part of all of the above adventures, and more. Thank you for everything. You know what you did.

I would like to thank Kasia, for her support and presence during some of the hardest parts of this story, as well as all of the other wonderful human beings that she allowed me to meet in Poland and elsewhere.

This thesis is written with calm and joy, thanks in no small part to Lucie, and her innate glow. You are wonderful, caring, and really good at eating apples. Thank you also, and by extension, to the members of the Hiron clan, for their giddiness and blissful chaos. The Grafton Road Noodles, also deserve my warmest thanks, for my

countless invasions. Marty, for cool moves, Jay for delight, and Élise for peace and for taking Lucie to a Brad Mehldau Concert in March 2019.

To my family in France, Switzerland, Colombia, Canada, Spain, London, and any other number of satellite countries, my absolute thanks. You are a well of warmth, kindness, and practicality, which has served me my whole life, but particularly in the last four years. Thank you to my parents for support, homegrown vegetables, and bash advice, and to my brother Alejandro, for dedication, conversations, and altruism.

There only remains to thank the true mastermind behind this thesis, Rachel. I don't think I have ever told you this, but when I am asked what my supervisor is like, I usually compare you to Yoda. You are cheerful and aloof, which fools unsuspecting Jedi into underestimating your vastly superior Force power (here, the Force, is excited state Computational Chemistry). With the hindsight of a few years, I can safely assert what I started to suspect only a few months after meeting you: when we disagree, you are usually right, and it's better to just follow your advice in the first place before opposing any sort of battle (you can show this sentence to future PhD students, just try to hide your smugness). I chose to pursue this project in great part because of your and Michael's characters, and you have proven that it was one of the best decisions of my life. Your combination of patience, joy, and rigour have brought me comfort and guidance at the worst lows of the project (e.g. after having to repeat months' worth of calculations because of a typo, or when my personal life got in the way). You have granted me confidence, expertise, and wisdom, and allowed me to preserve my life alongside an ambitious work project. For your support, presence, music recommendations, and pad thais at the Rusty Bike thank you. Anita and David are in good company.

This acknowledgments page is being written at a time of global crisis. In the spirit of camaraderie, and whatever the situation of the reader, I would like to extend my sincere sympathies for the hardships they may be encountering, however large or small, and solidarity in the obstacles to come.

Contents

Declaration of Authorship	3
Abstract	7
Acknowledgements	9
Introduction	29
I Background	33
1 Electronic Structure Methods	35
1.1 Wavefunctions Methods	35
1.1.1 The Born-Oppenheimer Approximation	36
1.1.2 Many-Electron Wavefunctions	36
1.1.3 The Hartree-Fock Method	37
1.1.4 Atomic Orbital Basis Sets	38
1.1.5 The Roothaan Equations	39
1.1.6 Electronic Correlation	40
1.1.7 Møller Plesset Perturbation Theory	41
1.1.8 Coupled Cluster	42
1.1.9 Multireference Methods	42
1.2 Density Functional Theory	44
1.2.1 The Hohenberg-Kohn Theorems	44
1.2.2 Kohn-Sham theory	45
1.2.3 The Kohn-Sham Potential	46
1.2.4 The Exchange-Correlation Functional	47
Local Density Approximation	47

	Generalised Gradient Approximation	47
	Hybrids	48
	Range Separated Hybrids	49
	Dispersion Correction	49
1.2.5	Density Functional Tight-Binding	49
1.3	Solid State Electronic Structure	50
1.3.1	The Periodic Lattice	51
1.3.2	Bloch's Theorem	51
1.3.3	Reciprocal Space	52
1.3.4	k-point Sampling	53
1.3.5	Plane Wave Basis Sets	53
1.3.6	The Ewald Potential	54
1.4	Excited States	55
1.4.1	Linear-Response	55
1.4.2	Time-Dependent Density Functional Theory	56
1.4.3	The Kohn-Sham Response Function	57
1.4.4	The Casida Formulation	58
1.4.5	Algebraic Diagrammatic Construction	59
1.4.6	Excited State Coupled Cluster	59
1.4.7	Multireference Methods (bis)	60
2	Embedding Methods	63
2.1	Introduction	63
2.2	Polarisable Continuum Models	63
2.3	Point Charge Embedding	64
2.3.1	Non-isotropic Media	64
2.4	General QM:MM	65
2.4.1	Cluster Models	65
2.4.2	Intermolecular Interactions	67
2.5	IMOMM	67
2.6	QM:MM Electrostatic Embedding	68
2.7	ONIOM QM:QM'	69

	15
2.8 Shortcomings of ONIOM Embedding Methods	69
2.8.1 Inclusion of Long-Range Effects	69
2.8.2 Point Charge Representation	70
2.8.3 Mutual Polarisation	71
2.8.4 Implementation	71
3 Photochemistry	73
3.1 Absorption	73
3.1.1 Principles of Photochemistry	73
3.1.2 Oscillator Strength	75
3.1.3 Natural Linewidth	75
3.1.4 Role of Vibrations	76
3.2 Emission	77
3.2.1 Fluorescence	77
3.2.2 Phosphorescence	77
3.3 Nonradiative Decay	78
3.3.1 Vibrational Internal Conversion	78
3.3.2 Conical Intersections	79
3.4 Excited States in Aggregates	80
3.4.1 Excitons	80
3.4.2 Kasha's Model	81
3.4.3 Types of Exciton	81
3.5 Quantum Yield of Fluorescence	82
3.6 Example of Application: Four-Level Lasers	83
II Results	85
4 Ewald Embedded Cluster	87
4.1 Introduction	87
4.2 Embedding Schemes	90
4.2.1 Point Charge Embedding	90
4.2.2 Ewald Embedded ONIOM QM:QM'	92
4.2.3 Origin of Point charges	96

4.3	Computational Details	98
4.4	Results	100
4.4.1	Localisation of the Excitation: Size of the QM Region	100
4.4.2	Point Charge Embedding: Electrostatic Effects in the Crystal	103
4.4.3	Embedded Cluster Models: Potential Energy Surfaces in the Crystal	108
4.5	Conclusion	113
5	Implementation	117
5.1	Introduction	117
5.2	Program Structure	119
5.2.1	Principal Classes	119
5.2.2	Dependencies	121
5.3	Features	123
5.3.1	Geometrical Analysis	123
	Voronoi Volumes	123
	Dimeric Arrangement	125
5.3.2	Exciton Analysis	128
	Exciton Classification	128
	Exciton Coupling Evaluation	131
5.3.3	ONIOM Calculations	135
	Example of Use	138
5.4	Conclusion	139
6	Application	141
6.1	Introduction	141
6.2	Computational Details	142
6.3	Results	147
6.3.1	Dimeric Arrangement	147
6.3.2	Exciton Coupling	149
6.3.3	Excited-state Relaxation by Series	153
6.4	Conclusions	165

	17
Conclusion	167
A Supporting Information for the Ewald Embedded Cluster	187
A.1 Crystal Structures	187
A.2 Absorption and Emission Spectra	187
A.3 Population Analysis for Aggregates	188
A.4 Vertical Excitation for Aggregate	190
A.5 Exciton Coupling in Dimers	191
A.6 Results with CC2 and CASPT2	192
B Supporting Information for the Applications	193
B.1 Geometries	193
B.2 Crystal Structures	193
B.3 Active Space	194
B.4 Full Normal Mode Analysis	194

List of Figures

- 1.1 Occupancy diagram of a closed shell Hartree-Fock wavefunction compared to a CASSCF(2,2) wavefunction. The occupancies C_μ are shown to be partial in the active space. 44
- 1.2 Schematic representation of the role of the unit cell in generating a lattice. In this 2-D example, \mathbf{a}_1 and \mathbf{a}_2 are lattice vectors, and \mathbf{R}_s is an atomic position in the unit cell. 51
- 2.1 Example of a cluster of benzene atoms taken from their lattice positions. The pink molecule denotes the central region. 66
- 3.1 Example of a potential energy surface demonstrating either a radiative decay *via* fluorescence, or a nonradiative decay *via* funnelling through a conical intersection. 74
- 3.2 Absorption spectrum of (2-hydroxyphenyl)propenone (HPP) in single crystal form.⁹⁷ 74
- 3.3 Diagram showing the manifold of transitions between two electronic states, each with four vibrational substates. The electronic wavefunctions are denoted ψ_n and the vibrational states ν_m 76
- 3.4 Visual representation of the three extreme types of excitons. 82
- 3.5 Diagram of the mechanism behind a four-level laser. Each step in the photocycle is labelled with its usual notation—Ground State (GS), Frank-Condon (FC), excited state minimum (S_1^*) and the corresponding ground state (S_1). 84

4.1	Molecular diagram of the 2'-hydroxychalcone derivatives, HC1 and HC2 . The access to conical intersections for these molecules is centred around the rotation of the blue group about the dihedral angle shown in bold. The notable atoms with large partial charge are labelled in pink.	89
4.2	Visual representation of the main energy equation for the Ewald Embedded Cluster model.	94
4.3	Flowchart of a calculation using the ONIOM Ewald Embedded Cluster (OEEC) and Self-Consistent ONIOM Ewald Embedded Cluster (SC-OEEC) models. The electronic program can be chosen by the user.	97
4.4	Selected tetramer configurations from both crystals. The molecule in pink is optimised using OEEC.	101
4.5	S_n-S_0 density differences obtained at TD- ω B97X-D/6-311++G(d,p) level of theory for the tetramer model (excited state density gain upon excitation is shown in orange, and the loss in blue). For the FC geometries, the tetramer's bright states were considered ($n=4$ for HC1 and $n=5$ for HC2); for K* , $n=1$. All configurations were obtained by optimising the geometry of the central molecule with OEEC.	102
4.6	Energies HC1 in vacuum and embedded in RESP and crystal Hirshfeld Ewald charges at different geometries. QM:MM energies were taken from Ref 132.	107
4.7	Deviation of the predicted emission energies of HC1 and HC2 with respect the experimental values by different embedding models. Reference experimental values were 1.8 eV for both HC1 and HC2 crystals	109
4.8	MECI geometries found with SA-2-CASSCF(12,11)/6-31G(d) and TD- ω B97X-D/6-311++G(d,p). Additionally, we include the configuration with least S_1-S_0 gap when scanning the β angle from the TDDFT geometry at MS-2-CASPT2(12,11)/6-31G(d) level. It is labelled CASPT2.	111

4.9	Relative energy diagrams showing the emission energy and accessibility of the MECI with multireference methods at the geometries shown in Figure 4.8. To minimise the CASPT2 S_1-S_0 gap at MECI configurations, a geometry scan was carried out as described in Figure 4.10. The newly optimised geometry is labelled “CASPT2 scan” and has a gap of 0.10 eV.	112
4.10	Plot of the S_1-S_0 gap, and the MECI accessibility as a function of the puckering angle β . The accessibility is defined as the average S_1-S_0 energy of the given geometry minus the FC bright state energy. The OEEC energy at the geometry with smallest S_1-S_0 gap is indicated by dashed lines.	113
4.11	Energy of the absorption, emission and conical intersection as the excited state region increases in size using the OEEC model. The region molecules are added to the region in the order displayed in Figure 4.4. The energies are evaluated with TD- ω B97X-D/6-31G(d).	114
5.1	Principal features of fromage.	120
5.2	Example of use of fromage as a Python library. The three dimensional structures are shown in orthographic projection.	121
5.3	a) Principal class diagram of fromage b) Description of the main callable modules.	122
5.4	a) Chemical structure of HC2 b) unit cell of the HC2 crystal with the asymmetric cell highlighted by colour c) Profiles of the Voronoi volumes for the four molecules of the asymmetric cell with their volume index.	124
5.5	Algorithm to generate principal axes from a monomer geometry with 2-fluoronaphthalene as an example.	126
5.6	Case study of DMAH for dimer geometry analysis a) Chemical structure b) Crystalline cluster c) Principal axes determination d) Heatmap of the angles between primary and secondary axes, respectively α and β	128

5.7	Transition density of two excited states of a perylene dimer taken from its crystal structure. The excited state energies are reported with respect to S_0 . The electron density migrates from the blue to the yellow areas upon excitation. The exciton classification indices are defined in equation 5.2 and are in units of e^-	130
5.8	a) Chemical structure of 4PV b) Top view of a cluster of 4PV molecules taken from their crystal positions c) Side view of the cluster with inequivalent dimers labelled by their exciton coupling value in meV. The values in parentheses are calculated as a trimer.	134
5.9	Generation of the bond order matrix for methanol.	137
5.10	Relative energies of critical geometries of DMAP evaluated with OEEC at critical points of the PES. The Franck–Condon point (FC), Enol (E^*) and Keto (K^*) excited state minima and Conical Intersection (CI) geometries were located.	139
6.1	Molecular structures of the studied systems.	143
6.2	a) Principal and secondary axes on the 3P monomer b) Top view of a 3P dimer c) Side view of the same dimer.	145
6.3	Illustration of two archetypal packing motifs in molecular crystals. . .	147
6.4	Distribution of the dimeric arrangements in the DCS series (14 dimers) and the rest of the crystals (14 dimers). The angle between principal axes is plotted against the angle of the secondary axes. Almost all DCS series dimers are perfectly parallel, the exceptions are shown. . .	148
6.5	(top) Exciton coupling (J) as a function of the centroid-to-centroid distance of the constituent monomers of each dimer. The values of every dimer are fitted to an inverse law $f(r) = a/r$ via least squares. The resulting function, with $a = 459$ is plotted in pink and has a standard deviation of 27, represented by the shaded area. A dashed line represents the same fit but using an aromatic carbon (bottom) as a reference for the distance calculation, denoted HBT*. In this case, $a = 480$ and the standard deviation is 23.	150

6.6	Transition densities of the bright states of HBT in monomer form and both dimers of least centroid-to-centroid distance. Upon excitation, the density is reorganised from the blue to the orange areas.	151
6.7	Exciton coupling values of irregularly packed DSB dimers, sampled from their vibrational phase space.	151
6.8	Exciton coupling as a function of half of the S_2-S_1 energy gap. The linear trend line, obtained <i>via</i> least squares, is $f(x) = 0.997x$	152
6.9	HBT energy at critical points of its excited state potential energy surface. The vacuum calculation used TD- ω B97X-D/6-31G(d) and the crystal calculation TD- ω B97X-D/6-31G(d):DFTB.	156
6.10	The energies of HOMO and LUMO orbitals of 3P, 4P, and 6P in the crystal computed at the TD- ω B97X-D/6-31G(d) level. The S_1 transition densities are represented as well.	157
6.11	Energies of S_0 and S_1 states of 3P and 4P at the FC point, S_1 minima, and S_1-S_0 MECI in vacuum (left) and crystal (right) computed at the RI-ADC(2)/def-SV(P) level. For the comparison, in the case of 3P, the corresponding CASPT2(10,10)/6-31G(d)//CASSCF(10,10)/6-31G(d) and CASPT2(8,8)/6-31G(d)//CASSCF(8,8)/6-31G(d) results are given	159
6.12	Energies of S_0 and S_1 states of 6P at the FC point, S_1 minima, and S_1-S_0 MECI in vacuum computed at the RI-ADC(2)/def-SV(P) level. The MECI structure is represented in the bottom.	160
6.13	α -DCS energy at critical points of its excited state potential energy surface. The vacuum and solvent calculation used TD- ω B97X-D/6-31G(d) and the crystal calculation TD- ω B97X-D/6-31G(d):DFTB.	162
A.1	Absorption and emission spectra for both molecules with experimental data ¹³⁴ for comparison. The experimental emission for HC2 is smoothed with a Savitzky-Golay filter.	188

- A.2 Deviation of charge on notable atoms from the explicitly calculated tetramer with the relaxing molecule in **FC** geometry. In the top row, the molecule is optimised with OEEC and in the bottom with SC-OEEC-S₁. We compare the charge obtained from the monomer population analysis of in vacuum S₀, S₁ and in EEC and SC-OEEC-S₁. The absolute average charge in the tetramer is indicated by horizontal ticks, positive values are in grey and negative ones in black. The atom labels are the ones found in Figure 1 of the main text. 189
- A.3 Deviation of charge on notable atoms from the explicitly calculated tetramer with the relaxing molecule in **K*** geometry. In the top row, the molecule is optimised with OEEC and in the bottom with SC-OEEC-S₁. We compare the charge obtained from the monomer population analysis of in vacuum S₀, S₁ and in EEC and SC-OEEC-S₁. The absolute average charge in the tetramer is indicated by horizontal ticks, positive values are in grey and negative ones in black. The atom labels are the ones found in Figure 1 of the main text. 189
- A.4 Largest differences in RESP charges between S₁ and S₀ for HC1 and HC2 in vacuum in units of 10⁻¹ e⁻¹. We only include atoms where the difference was more than 0.01 e⁻¹ in either molecule. The diameter of the circles is proportional to the charge difference. 190
- A.5 Density plots for the absorption of HC1 under OEEC and SC-OEEC. The for each excited state *n*, the difference in electronic density S_{*n*}-S₀ is plotted where positive values are shown in orange and negative values in blue. The bright state is highlighted in pink. 190
- A.6 Density plots for the absorption of HC2 under OEEC and SC-OEEC. For each excited state *n*, the difference in electronic density S_{*n*}-S₀ is plotted where positive values are shown in orange and negative values in blue. The bright state is highlighted in pink. 191
- B.1 Active spaces of 3P and HBT for multireference calculations. 195

List of Tables

4.1	Embedding models used in this study	95
4.2	Absorption, emission and K-MECI energies (in eV) of HC1 embedded in different types of Ewald point charge arrays. Unless specified the geometries were obtained at the ONIOM(TD- ω B97X-D/6-311++G(d,p):AMBER) level of theory. K-MECI energies are relative to the ground state energy of the Franck-Condon (FC) minimum. ¹³² [†] Optimised geometries within the PCE environment. ^{††} Geometry optimised in vacuum	104
4.3	Table of absorption and emission energies for both model systems after geometry optimisation with cluster models. The level of theory was TD- ω B97X-D/6-311++G(d,p). Energies are in eV. [†] Charges obtained for the K* form in S_1	108
5.1	Interfaced quantum methods and their availability in electronic structure programs	136
6.1	Photoactive molecular crystals considered. Φ_f : fluorescence quantum yield in crystal, V_i : steric volume index. ^a Largest dimeric exciton coupling in the crystal, ^b Herringbone (HB), Sheet (S) or other (-), ^c fraction of H (not J) dimers in the crystal, ^d tetrahydro-2-metehylfuran solvent, ^e cyclohexane solvent, ^f chloroform solvent, ^g powder, ^h film. .	152
6.2	Computed S_1 energies in eV ($E(S_1)$) with corresponding oscillator strength (f), radiative rates calculated with the Einstein relation and SB formula in ns^{-1} (k_r^{Ein} , k_r^{SB}), experimental luminescent efficiency (Φ_f), and reorganisation energies in eV (λ) for the $n\text{P}$ series. The superscripts indicate the medium, where "sol" is cyclohexane solvent, "vac" is vacuum, and "cr" is crystal. ^a Powder samples. ^b Single crystal.	157

6.3	Experimental radiative rates in ns^{-1} (k_r), nonradiative rates in ns^{-1} (k_{nr}), luminescent efficiency (Φ_f), and computed reorganisation energies for vibrational modes of less than 0.031 eV (250 cm^{-1}) in eV (λ_{low}) for the DCS series. The superscripts indicate the medium, where "sol" is chloroform solvent, "vac" is vacuum, and "cr" is single crystal	164
A.1	Exciton couplings (eV) for the different fragments involved in the tetramers computed with the diabatisation scheme. In parenthesis is half the energy splitting between S_1 and S_2 . The dimers A, B and C are shown in the Figure 6 of the main text	192
A.2	Table of absorption and emission energies with RI-CC2 and multireference methods for both model systems after geometry optimisation. The energies are in eV	192
B.1	Cambridge Crystallographic Data Centre (CCDC) IDs for the corresponding crystal structures.	193
B.2	Reorganisation energies for normal modes of the nP and DCS series, in vacuum and crystal, split between low and high energy normal modes.	194

List of Abbreviations

ADC2	Algebraic Diagrammatic Construction 2
AIE	Aggregation Induced Emission
BO	Born-Oppenheimer
CASPT2	Complete Active Space Perturbation Theory 2
CASSCF	Complete Active Space Self-Consistent Field
CC	Coupled Cluster
CCS	Coupled Cluster Singles
CCSD	Coupled Cluster Singles and Doubles
CGF	Contracted Gaussian Function
CI	Configuration Interaction
CSF	Configuration State Function
CT	Charge Transfer
DFT	Density Functional Theory
DFTB	Density Functional Tight-Binding
FC	Frank-Condon
FGR	Fermi's Golden Rule
fromage	FRamewOrk for MOlecular AGgregate Excitations
GGA	Generalised Gradient Approximation
GTO	Gaussian Type Orbital
HF	Hartree-Fock
IMOMM	Integrated Molecular Orbital Molecular Mechanics
IMOMO	Integrated Molecular Orbital Molecular Orbital
KS	Kohn-Sham
LDA	Local Density Approximation
MM	Molecular Mechanics
MP2	Møller Plesset 2

MRCI	Multireference Configuration Interaction
MRSCF	Multireference Self-Consistent Field
OLED	Organic Light Emitting Diode
ONIOM	Own N-layered Integrated molecular Orbital and molecular Mechanics for Quantum Mechanics
PBE	Perdew-Becke-Ernzerhof
PCM	Polarisable Continuum Model
QE	Quantum Efficiency
QM	Quantum Mechanics
RACI	Restricted Access to Conical Intersections
RIM	Restriction of Intramolecular Motions
SCF	Self-Consistent Field
SE	Shrödinger Equation
SLE	Solid state Luminescent Enhancement
TDDFT	Time-Dependent Density Functional Theory
TISE	Time Independent Shrödinger Equation

Introduction

The luminescent response of molecular crystals has found applications in several recent and impactful technologies such as biosensing devices, Organic Light Emitting Diodes (OLEDs), field-effect-transistors, lasers, and photovoltaics.¹⁻⁸ However, the theoretical investigation of the photochemistry of solids has not been as reliable an experimental aid as that of liquids or gases, due to the added complexity associated with modelling condensed phases. Indeed the calculation of a molecular excited states is often restricted to the low hundreds of atoms due to its prohibitive computational cost and theoretical inadequacy for extended correlated systems. On the other hand, solids are often defined by their long-range interactions, which can involve moles of collaborating atoms for periodic systems. Molecular condensed phases held together by van der Waals forces occupy a liminal space in this picture, where due to their non-bonded nature, they share properties both of crystals and single molecules, indicating that an overlap in methodologies could be used to investigate their excited states in a robust way.

In recent years, multiscale modelling methods such as "our Own N-layered Integrated molecular Orbital and molecular Mechanics for Quantum Mechanics:Quantum Mechanics'" (ONIOM QM:QM')^{9,10} have proven a useful asset in modelling the excited states of molecular crystals.¹¹⁻¹⁷ Cluster model methods represent excitations as defect perturbations rather than fully delocalised phenomena as in full unit-cell periodic excited state calculations. They also limit the amount of molecules to be calculated in the excited state in all cases but the ones with smallest unit cells. Finally, their use of local orbital basis sets avoids the increased computational cost of plane-wave Density Functional Theory (DFT) for highly accurate functionals, which remains a topic of research.¹⁸ Hybrid DFT and post-Hartree-Fock methods are therefore available within the cluster model paradigm.

However the conventional ONIOM QM:QM' method presents several weaknesses in modelling the condensed phase:

1. The system-environment interactions between the excitation and the crystal are truncated to include only the nearest neighbours.
2. The electrostatic interactions are modelled as point charge potentials whose charge values are not uniquely defined.
3. The response of the environment to the electronic reorganisation of the system is ignored.
4. They lack an implementation tailored to excited organic molecular crystals.

These shortcomings are elaborated upon in Section 2.8, after presenting the necessary theoretical framework to understand them.

The object of this thesis is to offer solutions to the above problems in QM:QM' modelling for excited states in organic molecular crystals, in the form of new modelling methods implemented in a software package, and to apply these methods in order to understand luminescent phenomena in systems previously difficult to model.

A principal outcome of the doctoral project, along with the publications cited throughout and this thesis, is the program `fromage`, the `FRamewOrk` for Molecular `AGgregate` Excitations. Readers interested in the practical applications of the topics discussed herein can peruse the software project at: <https://github.com/Crespo-Otero-group/fromage>.

The present work is split in two parts. Part I focuses on the required theoretical background to understand the use of cluster models to model photochemistry in molecular crystals. Chapter 1 introduces the quantum mechanical modelling methods which are used throughout, Chapter 2 presents the embedding schemes employed to allow these methods to account for a system's environment, and Chapter 3 introduces foundational topics in photochemistry which will help contextualise what it is that the cluster models are to simulate. Then, Part II presents the results of the thesis. Chapter 4 introduces new ONIOM QM:QM' cluster models tailored to model excited molecular crystals, Chapter 5 describes the program which implements them and other auxiliary tools for analysing these systems, and Chapter 6

investigates of the competing radiative and nonradiative mechanisms of an array of luminescent organic crystals, relying on the new methodological advances.

The following publications constitute the broader output of this doctorate:

- M. Dommett, M. Rivera and R. Crespo-Otero, How Inter- and Intramolecular Processes Dictate Aggregation-Induced Emission in Crystals Undergoing Excited-State Proton Transfer. *J. Phys. Chem. Lett.*, 2017, **8** (24), 6148–6153.
- M. Rivera, M. Dommett and R. Crespo-Otero, ONIOM(QM:QM') Electrostatic Embedding Schemes for Photochemistry in Molecular Crystals. *J. Chem. Theory Comput.*, 2019, **15** (4), 2504–2516. **Chapter 4**
- M. Dommett, M. Rivera, M. T. H. Smith and R. Crespo-Otero, Molecular and Crystalline Requirements for Solid State Fluorescence Exploiting Excited State Intramolecular Proton Transfer. *J. Mater. Chem. C*, 2020, **8** (7), 2558–2568.
- M. Rivera, M. Dommett, A. Sidat, W. Rahim and R. Crespo-Otero, fromage : A Library for the Study of Molecular Crystal Excited States at the Aggregate Scale. *J. Comput. Chem.*, 2020, **41** (10), 1045–1058. **Chapter 5**
- M. Rivera, L. Stojanović, R. Crespo-Otero, Competition between radiative and nonradiative excited state processes in photoluminescent organic molecular crystals. *in preparation*, 2020. **Chapter 6**
- V. Posilgua, D. Pandya, M. Rivera, R. Crespo-Otero, R. Grau-Crespo, Two-dimensional porphyrin-based metal organic frameworks for photocatalytic water splitting: a computational investigation. *in preparation*, 2020.

Part I

Background

Chapter 1

Electronic Structure Methods

NB: The notation in this chapter is largely borrowed from the following texts: *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* by Szabo and Ostlund,¹⁹ *Time-Dependent Density-Functional Theory: Concepts and Applications* by Ullrich,²⁰ and *Quantum Theory of the Solid State: An Introduction* by Kantorovich.²¹

1.1 Wavefunctions Methods

The state function of a nonrelativistic quantum mechanical system is wholly described by the Shrödinger equation (SE):

$$i\hbar \frac{d}{dt} |\Psi(t)\rangle = \hat{H} |\Psi(t)\rangle \quad (1.1)$$

Where the $|\Psi(t)\rangle$ is the time dependent wave function, and \hat{H} is the Hamiltonian operator. In a system of N electrons and M nuclei, \hat{H} has the form:

$$\hat{H} = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{r_{AB}} \quad (1.2)$$

Where ∇_i^2 and ∇_A^2 are the Laplacian operators acting on electrons and nuclei respectively, Z_A is the charge of nucleus A and r_{AB} is the distance between points A and B .

In an equilibrium, the quantum system becomes time independent by definition, and has a simpler equation of state, the time-independent Shrödinger equation

(TISE):²²

$$\hat{H}|\Psi\rangle = E|\Psi\rangle \quad (1.3)$$

Where $|\Psi\rangle$ is the time independent wave function and E its associated energy.

1.1.1 The Born-Oppenheimer Approximation

The SE is analytically unsolvable for any system with more than two particles. We must therefore employ approximations to model chemically meaningful situations.

Nuclei are several orders of magnitude heavier than electrons, suggesting that the Hamiltonian can be split into an electronic and a nuclear term. The electronic term treats nuclei as static potentials, and can yield an electronic wavefunction, uncorrelated with the nuclear one, which does not take into account the nuclear kinetic and internuclear potential terms of Equation 1.2:

$$\hat{H}_{el} = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} \quad (1.4)$$

$$\hat{H}_{el}\Phi = E_{el}\Phi \quad (1.5)$$

This is known as the Born-Oppenheimer (BO) approximation, and allows for the nuclear coordinates to be considered only parametrically when calculating the energy of a given nuclear geometry.²³

1.1.2 Many-Electron Wavefunctions

Electrons are fermions of spin 1/2. Therefore the wavefunction which describes an electron should contain both a spatial part ($\psi(\mathbf{r})$), and a spin part ($\alpha(\omega)$ or $\beta(\omega)$ for spin up or down). Single electron wavefunctions are therefore called spin orbitals and, for example for spin up, are written:

$$\begin{aligned} \chi(x) &= \psi(\mathbf{r})\alpha(\omega) \\ |\chi\rangle &= |\psi\rangle|\alpha\rangle \end{aligned} \quad (1.6)$$

Where the x coordinate contains both the position \mathbf{r} and the spin ω .

A many electron wavefunction should be constructed taking into account the properties of fermions that they are antisymmetric under exchange. To fulfil $\Psi(\mathbf{x}_1, \mathbf{x}_2) = -\Psi(\mathbf{x}_2, \mathbf{x}_1)$, we can exploit the property of the matrix determinant, which changes sign with the interchange of row or column. This kind of wavefunction is called the Slater determinant:²⁴

$$\Psi^{SD}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = (N!)^{-1/2} \begin{vmatrix} \chi_i(\mathbf{x}_1) & \chi_j(\mathbf{x}_1) & \cdots & \chi_k(\mathbf{x}_1) \\ \chi_i(\mathbf{x}_2) & \chi_j(\mathbf{x}_2) & \cdots & \chi_k(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \chi_i(\mathbf{x}_N) & \chi_j(\mathbf{x}_N) & \cdots & \chi_k(\mathbf{x}_N) \end{vmatrix} \quad (1.7)$$

In Dirac notation, we adopt the convention $\Psi^{SD}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = |\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_N\rangle$.

1.1.3 The Hartree-Fock Method

We now wish to find a method to compute all spin orbitals of the electronic system's Slater determinant to obtain the most accurate wavefunction possible. To do so, for each spin orbital, we write a one-electron secular equation:

$$f(1)\chi(\mathbf{x}_1) = \epsilon\chi(\mathbf{x}_1) \quad (1.8)$$

$f(1)$ is called the Fock operator, and is written:

$$f(1) = h(1) + v^{HF}(1) \quad (1.9)$$

Where $h(1)$ is the one-electron core Hamiltonian:

$$h(1) = -\frac{1}{2}\nabla_1^2 - \sum_{A=1}^M \frac{Z_A}{r_{1A}} \quad (1.10)$$

And $v^{HF}(1)$ is the Hartree-Fock (HF) potential, associated with the multi-electronic

interaction energy. It, in turn, is comprised of Coulomb and exchange electron-electron interactions, and is expressed:

$$v^{HF}(1) = \sum_b (\mathcal{J}_b(1) - \mathcal{K}_b(1)) \quad (1.11)$$

Where the Coulomb operator $\mathcal{J}_b(1)$ expresses the electrostatic potential on the electron 1 in orbital χ_a from electron 2 in orbital χ_b :

$$\mathcal{J}_b(1) = \int \frac{|\chi_b(2)|^2}{r_{12}} d\mathbf{x}_2 \quad (1.12)$$

The exchange operator, on the other hand, is a nonlocal operator; it depends on the value of χ_a at points of space different from \mathbf{x}_1 as follows:

$$\mathcal{K}_b(1)\chi_a(1) = \left[\int \chi_b(2)^* r_{ij}^{-1} \chi_a(2) d\mathbf{x}_2 \right] \chi_b(1) \quad (1.13)$$

In its full form, the Fock operator is therefore:²⁵⁻²⁷

$$f(1) = -\frac{1}{2}\nabla_1^2 - \sum_{A=1}^M \frac{Z_A}{r_{1A}} + \sum_b (\mathcal{J}_b(1) - \mathcal{K}_b(1)) \quad (1.14)$$

1.1.4 Atomic Orbital Basis Sets

The two above operators involve the integration of products of spatial orbitals. If we could decompose these orbitals in an infinite number of functions spanning all of space, this would yield an exact spin orbital within the constraints of Hartree-Fock theory. However, due to the computational constraints of our finite central processing units, we must settle for an easily integrable set of functions which are arranged in a physically meaningful way.

In the context of a molecular system, it makes most sense to employ sets of atom-centred functions which have the shapes of individual electrons about a particular nucleus. These are called atomic orbitals, and their linear combination makes up a molecular orbital. If our set has K basis functions, we write spatial orbital ψ_i :

$$\psi_i = \sum_{\mu=1}^K C_{\mu i} \phi_{\mu} \quad (1.15)$$

For ease of integration, we often use a solutions of the hydrogen atom multiplied by Gaussian radial component. This is called a Gaussian Type Orbital (GTO). However, the behaviour of Gaussian functions in the limit of large distance does not necessarily reflect the behaviour of atomic orbitals, which instead follow an inverse law, thus in practice we use linear combinations of several GTOs, whose coefficients are called contractions, explaining the name Contracted Gaussian Functions (CGFs).

Oftentimes, the valence electrons are assigned more than one CGF, to obtain a better resolution of the interatomic interactions. Basis sets with several CGFs per valence electron are called split valence, or Pople basis sets if they are written in the form $X\text{-}YZG$. For example, the basis set 6-31G has six GTOs per core electron basis function, and two CGFs for valence electrons, one with three GTOs and one with one GTO.

Additionally, the basis sets may be complemented by diffuse Gaussians to accurately describe the region of the atom far from the nucleus, or asymmetric functions to represent the polarisability of the atom. A 6-31G basis set complemented by diffuse and polarisable functions is written 6-31G(d,p).

1.1.5 The Roothaan Equations

Armed with our basis set, we are able to re-write the HF equation in matrix form by enclosing the Fock operator with basis functions. We write:

$$FC = SC\epsilon \quad (1.16)$$

Where the Fock matrix has elements $F_{\mu\nu} = \int \phi_{\mu}^*(1)f(1)\phi_{\nu}(1)dr_1$, S is the overlap matrix with elements $S_{\mu\nu} = \int \phi_{\mu}^*(1)\phi_{\nu}(1)dr_1$, which account for the nonorthogonality of the basis, C defines the basis coefficients of the spin orbitals in Equation 1.15, and ϵ is the diagonal matrix of orbital energies:

$$\epsilon = \begin{pmatrix} \epsilon_1 & & 0 \\ & \ddots & \\ 0 & & \epsilon_K \end{pmatrix} \quad (1.17)$$

The equations originating from the canonical Equation 1.16 are called the Roothaan equations. They can be manipulated to become orthogonal, and adopt a solvable form without the interfering overlap matrix. A particularity of the Roothaan equations is that since the Fock matrix contains the HF potential, it depends on the spin orbitals, which are themselves defined in \mathcal{C} . This indicates that the Roothaan equations must be solved iteratively, in order to find a self-consistent set of spin-orbitals. This procedure is known as the Self-Consistent Field (SCF) method.

The SCF method allows us to obtain a set of K spatial orbitals and thus $2K$ spin orbitals. If we construct a wavefunction which is the Slater determinant of the N lowest spin orbitals, we obtain the HF ground state:

$$|\Psi_0\rangle = |\chi_1\chi_2 \dots \chi_K\rangle \quad (1.18)$$

1.1.6 Electronic Correlation

The N spin orbitals which make up the HF ground state are called *occupied* orbitals, while all of the higher energy spin orbitals are *virtual* orbitals.

We can construct new Slater determinants by swapping out occupied orbitals for virtual orbitals from the HF ground state. These are called excited Slater determinants, and are characterised by the amount of virtual orbitals included at one time, for instance, single excitation, double excitation and so on. The real wavefunction has the conformational space to adopt any combination of excited determinants, therefore, ignoring the size of the basis set, the best possible description of the nonrelativistic BO wavefunction would be a linear combination of all HF Slater determinants.

$$|\Phi\rangle = c_0 |\Psi_0\rangle + \sum_{ra} c_a^r |\Psi_a^r\rangle + \sum_{\substack{a<b \\ r<s}} c_{ab}^{rs} |\Psi_{ab}^{rs}\rangle + \sum_{\substack{a<b<c \\ r<s<t}} c_{abc}^{rst} |\Psi_{abc}^{rst}\rangle + \dots \quad (1.19)$$

Where $|\Psi_{\{\alpha\beta\dots\}}^{\{\mu\nu\dots\}}\rangle$ is the Slater determinant with the excitations $\alpha \rightarrow \mu, \beta \rightarrow \nu \dots$ with the corresponding expansion coefficient $c_{\{\alpha\beta\dots\}}^{\{\mu\nu\dots\}}$. $|\Phi\rangle$ is called the full configuration interaction (CI) wavefunction.

Due to the explosively increasing number of terms of the full CI wavefunction, it is impractical to calculate for systems of more than a few electrons. However, it gives us a formal definition for the error in the HF ground state. We define the *correlation energy* as the difference between the energies of the HF ground state (E_0) and full CI wavefunctions (\mathcal{E}_0):

$$E_{\text{corr}} = \mathcal{E}_0 - E_0 \quad (1.20)$$

Several methods have been developed to recover correlation energy by including the effect of excited determinants in the HF ground state.

1.1.7 Møller Plesset Perturbation Theory

Rayleigh–Schrödinger perturbation theory is a numerical method in quantum mechanics which corrects small errors in approximate wavefunctions in the form of an asymptotic series. If we use the Fock operator, shifted by the sum of the orbital energies, as the the reference wavefunction, then the correlation energy can act as the perturbation.²⁸

The order of the asymptotic series increases with cost and accuracy of the method. At first order, this energy—called Møller Plesset 1 (MP1)—matches the HF ground state ($E_{\text{MP1}} = E_0$). At second order, the energy takes the form:

$$E_{\text{MP2}} = 2 \sum_{abrs}^{N/2} \frac{\langle ab|rs\rangle \langle rs|ab\rangle}{\epsilon_a + \epsilon_b - \epsilon_r - \epsilon_s} - \sum_{abrs}^{N/2} \frac{\langle ab|rs\rangle \langle rs|ba\rangle}{\epsilon_a + \epsilon_b - \epsilon_r - \epsilon_s} \quad (1.21)$$

Where we have adopted the shorthand:

$$\langle ij|kl\rangle = \langle \chi_i \chi_j | \chi_k \chi_l \rangle = \int \chi_i^*(\mathbf{x}_1) \chi_j^*(\mathbf{x}_2) r_{12}^{-1} \chi_k(\mathbf{x}_1) \chi_l(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 \quad (1.22)$$

And ϵ_i is the energy of orbital i . The MP3 and MP4 energies have a more complex form and are not as widely used as MP2 due to the increased computational cost with modest increase in accuracy. Furthermore the perturbative expansion of the HF wavefunction diverges when degenerate energy corrections occur. Olsen and Jørgensen showed that the energy F^- diverged starting at only third order.²⁹

1.1.8 Coupled Cluster

Coupled cluster techniques are an alternate method of recovering electron correlation, where excited determinants are directly used. We introduce the *cluster operator* $\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots$ such that \hat{T}_n is the operator of all of the excitations of n th order. The operation of \hat{T}_n on wavefunctions is computationally costly at high n , thus we often truncate the sum to first or second order. The order of the truncation defines the approximation, where the first order is the Coupled Cluster method for Singles (CCS) and the second order is for Singles and Doubles (CCSD).

The coupled cluster wavefunction is written:³⁰

$$|\Psi_{\text{CC}}\rangle = e^{\hat{T}} |\Psi_0\rangle \quad (1.23)$$

Operating on a wavefunction with $e^{\hat{T}}$ has useful properties due to the expansion of the exponential function:

$$e^{\hat{T}} = 1 + \hat{T} + \frac{1}{2!} \hat{T}^2 + \dots \quad (1.24)$$

Thus, in CCSD, even when $\hat{T} = \hat{T}_1 + \hat{T}_2$, we are still allowing for contributions from excitations of higher order than two due to the exponents appearing on \hat{T} in the expansion. As we employ higher order forms of coupled cluster such as CCSD(T), we allow a larger contribution to higher order excitations, which are naturally scaled down in the expansion above.

CCSD(T) considers triple excitations perturbatively, and is often used as a benchmark method. Indeed its accuracy has been shown to be in the range of 1 kcal/mol for small molecules.³¹

1.1.9 Multireference Methods

All of the methods introduced thus far have used the HF ground state as the reference wavefunction. In a closed-shell system with an even number of electrons, this wavefunction is composed of all of the occupied spin orbitals, which are weighted

equally in the Slater determinant. One could construct alternate Slater determinants with different expansion coefficients for the spin orbitals—called occupancies—which encompass some of the full CI wavefunction which was earlier rationalised as a collection of differently excited determinants. This paradigm treats electronic excitations of the wavefunction in a very different way than single-determinant based methods (which will be described in Section 1.4). Here, not only do occupancies change for different roots of the multireference secular equation, the molecular orbitals themselves would also be state-specific.

Our target wavefunction would be a combination of these partially occupied Slater determinants, except we must first impose an orthogonality condition to allow for this decomposition. We can produce a symmetry adapted linear combination of partially occupied Slater determinants, called a Configuration State Function (CSF). CSFs can then be, in turn, combined together variationally to produce a multireference CI (MRCI) wavefunction. The amount of CSFs determines the cost and accuracy of the MRCI procedure. Overall, the wavefunction will have the form:

$$|\Psi_{\text{MRCI}}\rangle = \sum_{\mu} C_{\mu} |\mu\rangle \quad (1.25)$$

Where C_{μ} are the occupancies and $|\mu\rangle$ are the CSFs.³²

We can reduce the amount of necessary CSFs by optimising the spin orbitals themselves as the procedure is carried out. This is known as multireference SCF (MRSCF). Furthermore, we usually limit the amount of CSFs to all of those arising from a certain number of electrons within a certain number of spin orbitals. These spin orbitals are called the *active space* and the method, Complete Active Space SCF (CASSCF).^{33–35} We use the notation CASSCF(a,b), indicating that we consider the CSFs which vary the occupation of a electrons in b orbitals. Figure 1.1 gives a visual representation of the CASSCF orbital arrangement, as compared to a closed shell HF arrangement.

CASSCF recovers a large amount of correlation due to near-degenerate configurations, which are ignored in HF. However it does not encompass all of the low excitation determinants, as is for example done in CCSD. To correct for this type of correlation, we can apply a perturbative correction of second order to the multistate

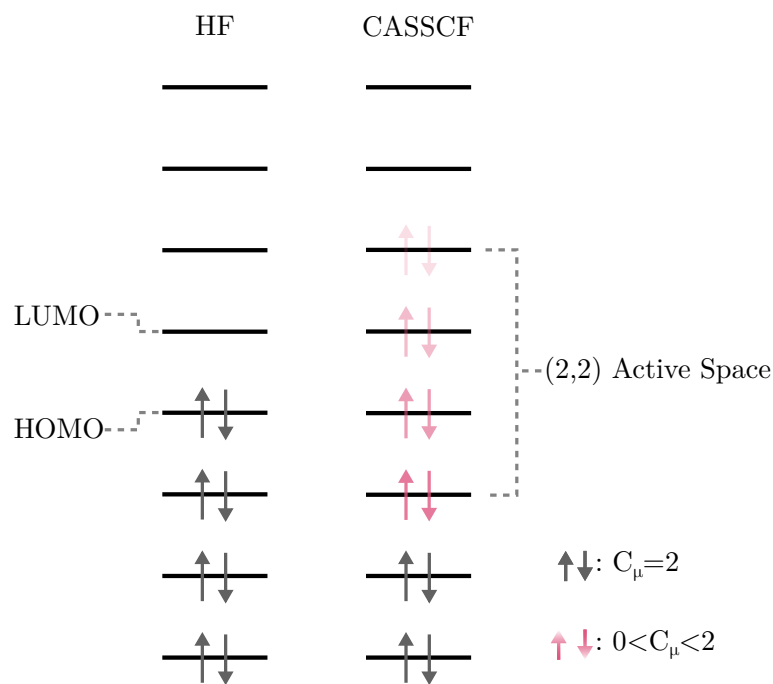


FIGURE 1.1: Occupancy diagram of a closed shell Hartree-Fock wavefunction compared to a CASSCF(2,2) wavefunction. The occupancies C_μ are shown to be partial in the active space.

CASSCF wavefunction. This produces the Complete Active Space Perturbation Theory 2 method, CASPT2.^{36,37} Alternate implementations, such as CASMP2 where the correlation from an MP2 calculation is used, are also possible.³⁸

1.2 Density Functional Theory

The previous section lists some of the most successful methods of calculating electronic properties of chemical systems by approximating their wavefunctions. However some of the most popular quantum chemistry methods today avoid calculating the wavefunction altogether, in favour of focusing on the electronic density as the principal state function.

1.2.1 The Hohenberg-Kohn Theorems

If we wish to express the energy of a system of N interacting electrons under some external potential v_{ext} , we can write:

$$E_{v_{\text{ext}}}[v_{\text{ext}}, n] = \int n(\mathbf{r})v_{\text{ext}}(\mathbf{r})d\mathbf{r} + F[n] \quad (1.26)$$

Where n is the electron density and $F[n]$ is the *universal functional* which maps n to the energy of the enclosed N interacting electron system.

Hohenberg and Kohn proved in the 1960s two fundamental theorems related to the electronic density n :³⁹

1. The external potential of a quantum system v_{ext} has a one-to-one correspondence with its electron density $n(\mathbf{r})$.
2. The density which minimises the total energy of the system is the real ground state density.

Theorem 1 implies that the energy of the system is a functional of its electron density, $E_{v_{\text{ext}}}[v_{\text{ext}}, n] \equiv E_{v_{\text{ext}}}[n]$. Theorem 2 states that we can find this density by evaluating its energy at each point, and varying it until we have found the minimum. This framework is known as Density Functional Theory (DFT). However, since the form of $F[n]$ is unknown, we must employ an approximation.

1.2.2 Kohn-Sham theory

The universal functional $F[n]$ is composed of the kinetic energy of the electron cloud $T[n]$ and the electron-electron interactions $W[n]$. The kinetic part is nontrivial to solve for electron densities, and the second one can be chosen to be the Coulomb interaction. In their wavefunction operator form, we would write:

$$\hat{F} = \hat{T} + \hat{W} \quad (1.27)$$

Where now the kinetic energy operator $\hat{T} = -\frac{\nabla^2}{2}$ is easy to solve for wavefunctions of a known form. We begin our model by imagining a set of N non-interacting electrons submitted to a fictitious $v_s(\mathbf{r})$. We can write their secular equations as:⁴⁰

$$(\hat{T} + v_s(\mathbf{r}))\varphi_i(\mathbf{r}) = \left(-\frac{\nabla^2}{2} + v_s(\mathbf{r})\right)\varphi_i(\mathbf{r}) = \epsilon_i\varphi_i(\mathbf{r}) \quad (1.28)$$

Where the eigenfunctions $\varphi_i(\mathbf{r})$ are known as the Kohn-Sham (KS) orbitals. The idea is to construct a KS wavefunction $\Psi_{\text{KS}}(x_1, \dots, x_N)$ from the Slater determinant of the $N/2$ lowest energy spatial KS orbitals. For this to work, we need $v_s(\mathbf{r})$, which

we call the KS potential, to accurately represent all of the electron-electron interactions which, based on the first Hohenberg-Kohn theorem, would yield a unique electron density:

$$n(\mathbf{r}) = \int |\Psi_{\text{KS}}(\mathbf{x}_1, \dots, \mathbf{x}_{N/2})|^2 d\mathbf{x}_1 \dots d\mathbf{x}_{N/2} = 2 \sum_j^{N/2} |\varphi_j(\mathbf{r})|^2 \quad (1.29)$$

The second Hohenberg-Kohn theorem requires this density to minimise the total energy functional. Fortunately, there is a systematic way of improving the accuracy of the wavefunction, and thus variationally reducing its energy. The external potential $v_s(\mathbf{r})$ must depend on the solutions of Equation 1.28. This means that they can be solved iteratively, in an SCF procedure analogous to that of HF. Equations 1.28 and 1.29 are known as the KS equations, and form the basis of KS DFT.

1.2.3 The Kohn-Sham Potential

We have yet to comment on the form of $v_s(\mathbf{r})$. Since it depends on the density, we write it as a functional:

$$v_s[n](\mathbf{r}) = v_{\text{ext}}(\mathbf{r}) + v_{\text{Har}}[n](\mathbf{r}) + v_{\text{XC}}[n](\mathbf{r}) \quad (1.30)$$

Where the first term is the external potential, including nuclear potentials under the BO approximation but also any external imposed potentials $v_{\text{field}}(\mathbf{r})$ to the system:

$$v_{\text{ext}}(\mathbf{r}) = v_{\text{field}}(\mathbf{r}) + \sum_i^{N_Z} \frac{Z_i}{|\mathbf{R}_i - \mathbf{r}|} \quad (1.31)$$

Where we have considered N_Z nuclei at positions \mathbf{R}_i with charge Z_i .

The second term of Equation 1.30 is the Hartree potential:

$$v_{\text{Har}}[n](\mathbf{r}) = \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (1.32)$$

Which represents the electron-electron Coulomb potential.

The final term of Equation 1.30 should represent all other electron-electron interactions, namely exchange and correlation. We write it as the derivative of the

exchange-correlation functional $E_{\text{XC}}[n]$:

$$v_{\text{XC}}[n](\mathbf{r}) = \frac{\delta E_{\text{XC}}[n]}{\delta n(\mathbf{r})} \quad (1.33)$$

We write it in this form because we will need the exchange-correlation functional when calculating the energy of the KS DFT system.

1.2.4 The Exchange-Correlation Functional

$E_{\text{XC}}[n]$ does not have a known exact solution, and we therefore adopt certain approximations in order to carry out DFT calculations.

Local Density Approximation

The simplest non-trivial exchange-correlation functional uses the analytical solution for the exchange energy density of a homogeneous electron gas $e_x^h(n(\mathbf{r}))$. The correlation part of the energy, $e_c^h(n(\mathbf{r}))$, has no known solution but one can parameterise it to match those of very accurate quantum Monte Carlo calculations.⁴¹

We can write the functional as the integral of the two energy densities:

$$E_{\text{XC}}^{\text{LDA}}[n] = \int e_x^h(n(\mathbf{r})) + e_c^h(n(\mathbf{r})) d\mathbf{r} \quad (1.34)$$

This functional is named the Local Density Approximation (LDA), since at each point, it only conveys the non-Coulombic many-electron interactions of a uniform electron gas with the density of the point. Surprisingly, this approximation still recovers the correct general structure of many bulk solids, for instance, though the error in energy is in the range of 1 eV per atom, which renders it poorly suited to quantitative prediction.⁴²

Generalised Gradient Approximation

LDA functionals can be improved by using the exchange and correlation from the fictitious system of the electron gas of constant gradient, instead of constant density. In the general terms, these Generalised Gradient Approximation (GGA) functionals

can be written:

$$E_{\text{XC}}^{\text{GGA}}[n(\mathbf{r})] = \int e_{\text{XC}}^{\text{GGA}}(n(\mathbf{r}), \nabla n(\mathbf{r})) d\mathbf{r} \quad (1.35)$$

The GGA exchange-correlation energy density $e_{\text{XC}}^{\text{GGA}}$ can take many forms. For example, one of the most popular GGA functionals, developed by Perdew, Becke and Ernzerhof⁴³ (PBE), was developed in order to be completely *ab initio* and satisfies as many exact properties of the KS wavefunction as possible. For instance, it is favoured by the community which simulates metal-organic frameworks for its affordability—given unit cells of up to hundreds of atoms—and accuracy in modelling bond angles with a precision greater than 0.1 Å.⁴⁴

One could also list the Becke 1988 exchange functional, B88⁴⁵, or the Lee-Yang-Parr correlation functional (LYP)⁴⁶ which combine into the BLYP functional.

Hybrids

A further improvement to consider is that of including exchange from a wavefunction-based calculation into our results. Indeed HF calculations have no correlation by definition but provide exact nonlocal exchange:

$$e_{\text{x}}^{\text{exact}} = -\frac{1}{2} \sum_{\sigma} \sum_{i,j=1}^{N_{\sigma}} \frac{\varphi_{i\sigma}^*(\mathbf{r}') \varphi_{j\sigma}(\mathbf{r}') \varphi_{\sigma i}(\mathbf{r}') \varphi_{j\sigma}^*(\mathbf{r}') d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} \quad (1.36)$$

Where the σ indices represent spin up or down.

By mixing this exact exchange with different proportions of LDA and GGA terms, one can construct hybrid exchange-correlation functionals. PBE0 combines PBE with exact exchange⁴⁷ for increased accuracy. The Heyd–Scuseria–Ernzerhof (HSE) functional screens the exchange part of PBE0 with an error function for added efficiency, calculating the long-range exchange with PBE.^{48,49} B3LYP is a three parameter exchange-correlation functional which mixes LDA and B88 exchange with LDA and LYP correlation.⁵⁰ B3LYP is currently the most used density functional⁵¹, and predicts bond lengths of bulk TiO₂ with an accuracy of 0.01 Å, compared to 0.1 Å for LDA.

Range Separated Hybrids

Other combinations of exchange and correlation terms are possible, and they can be screened by a distance parameter, in order to assign short-range energies to the best performing short-range functionals and *vice versa* for long-range ones. Here, we will only mention ω B97X-D,⁵² which uses the Becke 97 short-range GGA functional⁵³ and combines it with long-range HF exchange and a small portion of short-range H exchange. It reproduced CC2 excitation energies with a 0.2 eV accuracy on biochromophores, indicating a high reliability.⁵⁴

The long-range behaviour of exchange-correlation interactions is particularly relevant to intermolecular processes, since the interactions between monomers is beyond the distance of atomic bonds. In particular, excited state wavefunctions are typically more diffuse than their corresponding ground state, further justifying the use of range-separated functionals. Important delocalised phenomena such as charge transfer require such tools.

Dispersion Correction

Due to the parameterisation of the correlation functional at the root of LDA theory, KS DFT often omits the long-range correlation between electrons which results in dispersion (sometimes called van der Waals interaction), even when mixing HF exchange.

Grimme's D2 method⁵⁵ accounts for the dispersion of the system *a posteriori*, allowing for the correction of any of the functionals described above. In fact the *D* in ω B97X-D implies that this functional already uses Grimme's D2 correction.

1.2.5 Density Functional Tight-Binding

DFT is also the basis of a variety of semi-empirical methods, which aim to combine the efficiency of classical simulation with the accuracy of an electronic description. Prominently, the Density Functional Tight-Binding (DFTB) model eschews the determination of a KS density. Instead, it uses atomic KS orbitals, which have been previously parameterised through DFT calculations. The atomic orbitals are then combined in a Taylor expansion of the exchange-correlation functional.^{56,57}

The order of this Taylor expansion ranges in practice from one to three, where the first order energy expression is:

$$E^{\text{DFTB1}} = \sum_i n_i \epsilon_i + \frac{1}{2} V_{ab}^{\text{rep}} \quad (1.37)$$

Where n_i is the occupation of KS orbital i , ϵ_i is its orbital energy, and V_{ab}^{rep} are repulsive pair potentials either fitted from DFT calculations or from empirical data.⁵⁸

For organic systems, routine calculations usually use the second order, where density fluctuations are considered for all atoms. In this case, the atomic charge density fluctuations are modelled as exponentially decaying distributions with a decay parameter τ_a fitted to DFT calculations:

$$\delta\rho_a(\mathbf{r}) \approx \Delta q_a \frac{\tau_a^3}{8\pi} e^{-\tau_a|\mathbf{r}-\mathbf{R}_a|} \quad (1.38)$$

Where Δq_a is the fluctuation in partial charge on a given atom a , and \mathbf{R}_a is the position of the atomic nucleus.

The energy becomes:

$$E^{\text{DFTB2}} = E^{\text{DFTB1}} + \frac{1}{2} \sum_{ab} \Delta q_a \Delta q_b \gamma_{ab}(\tau_a, \tau_b, R_{ab}) \quad (1.39)$$

Where $\gamma_{ij}(\tau_i, \tau_j, R_{ij})$ is an analytical function resulting from evaluating the Hartree energy between the decaying charge densities and R_{ij} is the distance between two atomic centres.⁵⁶

Second order DFTB can be very accurate when applied to well-studied systems, since it is semi-empirical. Indeed it matches the accuracy of B3LYP in calculating geometrical parameters of TiO2 nanostructures.⁵⁹

1.3 Solid State Electronic Structure

Not only has DFT found success in molecular systems, it has also been formulated to address problems in systems with a natural periodicity such as crystals, surfaces or polymers.

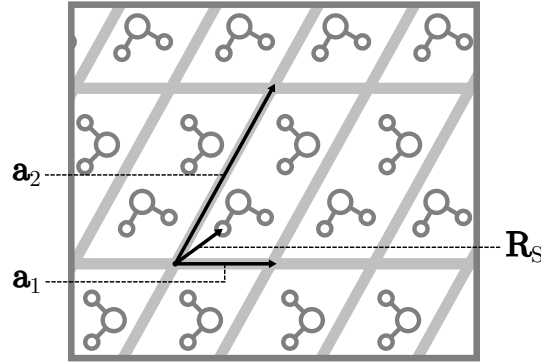


FIGURE 1.2: Schematic representation of the role of the unit cell in generating a lattice. In this 2-D example, a_1 and a_2 are lattice vectors, and R_s is an atomic position in the unit cell.

1.3.1 The Periodic Lattice

The defining feature of a crystal is that its structure has a given periodicity. We call the smallest repeating unit constituting the whole crystal by translations the *unit cell*. These translations determine the shape of the unit cell, and we therefore usually define a unit cell by its three primitive lattice vectors a_1 , a_2 , and a_3 .

Any atom within the unit cell is therefore located at:

$$\mathbf{R}_s = s_1 \mathbf{a}_1 + s_2 \mathbf{a}_2 + s_3 \mathbf{a}_3 \quad (1.40)$$

Where $0 \leq s_i < 1$. By definition, given an arbitrary number of transitions n_1 , n_2 and n_3 such that $n_i \in \mathbb{Z}$, there are atoms of the same type at all positions:

$$\mathbf{R}' = \mathbf{R} + n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \equiv \mathbf{R} + \mathbf{L}_{n_1, n_2, n_3} \quad (1.41)$$

Where we have defined $\mathbf{L}_{n_1, n_2, n_3}$ as an arbitrary lattice translation. By this process, an entire lattice can be constructed as shown in Figure 1.2.

1.3.2 Bloch's Theorem

We now wish to describe the wavefunction of electrons in a periodic array of nuclei, which is the application of the BO approximation to the crystal. The Schrödinger

equation is in this case:

$$\left(-\frac{\nabla^2}{2} + V(\mathbf{r}) \right) \psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad (1.42)$$

Where the periodic potential is $V(\mathbf{r}) = V(\mathbf{r} + \mathbf{L}_{n_1, n_2, n_3})$ for any lattice translation.

It can be shown that electrons in periodic potentials have wavefunctions in a basis of the form:

$$\psi_{nk}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{nk}(\mathbf{r}) \quad (1.43)$$

Where n is the integer energy level of the wavefunction, \mathbf{k} is called the *crystal wave vector* and has a dimension of inverse distance, and $u_{nk}(\mathbf{r})$ is a function with the periodicity of the potential. These basis functions are called *Bloch waves*.

1.3.3 Reciprocal Space

We now examine what would happen to the Bloch wave if \mathbf{k} is shifted by any vector \mathbf{b}_n such that $\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi\delta_{ij}$, where the set of \mathbf{a}_n vectors defines the periodicity of the potential.

$$\begin{aligned} \psi_{n(\mathbf{k}+\mathbf{b})}(\mathbf{r}) &= e^{i(\mathbf{k}+\mathbf{b})\cdot\mathbf{r}} u_{n(\mathbf{k}+\mathbf{b})}(\mathbf{r}) \\ &= e^{i\mathbf{k}\cdot\mathbf{r}} e^{i\mathbf{b}\cdot\mathbf{r}} u_{n(\mathbf{k}+\mathbf{b})}(\mathbf{r}) \\ &= e^{i\mathbf{k}\cdot\mathbf{r}} \tilde{u}_{n(\mathbf{k}+\mathbf{b})}(\mathbf{r}) \end{aligned} \quad (1.44)$$

For the cases where $\mathbf{r}' = \mathbf{L}_{n_1, n_2, n_3}$, we have $e^{i\mathbf{b}\cdot\mathbf{r}'} = e^{i2n\pi} = 1$, which means that the function $\tilde{u}_{n(\mathbf{k}+\mathbf{b})}(\mathbf{r}) = e^{i\mathbf{b}\cdot\mathbf{r}} u_{n(\mathbf{k}+\mathbf{b})}(\mathbf{r})$ has the same periodicity as $V(\mathbf{r})$. $\psi_{n(\mathbf{k}+\mathbf{b})}(\mathbf{r})$ can therefore be expressed as a Bloch wave with wave number \mathbf{k} , and presuming non-degenerate Bloch states, $\tilde{u}_{n(\mathbf{k}+\mathbf{b})}(\mathbf{r}) = u_{nk}(\mathbf{r})$ and $\psi_{n(\mathbf{k})}(\mathbf{r}) = \psi_{n(\mathbf{k}+\mathbf{b})}(\mathbf{r})$. This result can also be shown to hold for degenerate Bloch states. In other words, \mathbf{k} is periodic with respect to the vectors \mathbf{b}_n .

We call the inverse distance space of the lattice *reciprocal space*. It follows that the vectors \mathbf{b}_n are reciprocal primitive lattice vectors which form a space analogous to the primitive cell called the 1st Brillouin Zone (1BZ).

1.3.4 k-point Sampling

We have observed that electronic wavefunctions in crystals are expressed in the basis of Bloch waves. Therefore, setting aside the calculation of $u_{nk}(\mathbf{r})$, the representation of the wavefunction can be improved by increasing the size its basis, and therefore considering Bloch waves of more k values. The cost of a periodic DFT calculation is multiplied for each k-point considered, so it is important to calculate as few as necessary to represent the required properties of the system.

k-points can be understood as representing the periodicity of the material, therefore unit cells with a small lattice vector \mathbf{a}_1 should have a large amount of k-points in the \mathbf{b}_1 direction, and *vice versa*. In practice, we often use a grid of $N_1 \times N_2 \times N_3$ k-points, where the N_i values are chosen to be approximately inversely proportional to the lengths of the lattice vectors \mathbf{a}_1 , \mathbf{a}_2 , and \mathbf{a}_3 . This defines the Monkhorst-Pack grid.⁶⁰

1.3.5 Plane Wave Basis Sets

We now wish to calculate the $u_{nk}(\mathbf{r})$ term of Equation 1.43 using DFT. As in the case of molecular systems, periodic systems need to be decomposed in a basis in order to calculate the matrix elements of the external potential.

Since the wavefunction is periodic, its basis functions must also have a periodicity. We choose to use plane waves because their periodicity is easily tunable and their products can be integrated very efficiently using fast Fourier transforms.

$$u_{nk}(\mathbf{r}) = \sum_{\mathbf{G}} c_{\mathbf{G},k} e^{i\mathbf{G} \cdot \mathbf{r}} \quad (1.45)$$

Where \mathbf{G} are reciprocal lattice vectors. As $|\mathbf{G}|$ becomes larger, the expansion coefficients $c_{\mathbf{G},k}$ become smaller and therefore these terms of the expansion less important. Since we need to select a finite number of \mathbf{G} values for computation, we choose a cut-off value G_{cut} such that we only include plane waves with $|\mathbf{G}| \leq G_{\text{cut}}$ in the expansion.

This method works well for valence electrons, however since the energies of the electrons nearest to the nuclei are comparatively very high, we would need basis functions of a higher frequency to represent them and therefore a very high G_{cut} . To

avoid this, we usually avoid representing core electrons altogether, since they are unlikely to be involved in the chemical processes we are interested in. Instead, we represent the combined effect of the nucleus and the core electron's screening by functions called *pseudopotentials*.

1.3.6 The Ewald Potential

This subsection is adapted from *ONIOM(QM:QM') Electrostatic Embedding Schemes for Photochemistry in Molecular Crystals* by Miguel Rivera, Michael Dommett, and Rachel Crespo-Otero.⁶¹

An additional subtlety of performing DFT calculations in periodic systems comes from the new form of the nuclear potential. Indeed, the external potential of the KS equations has a component which is traditionally composed of atom centred point charges of the value of the atomic number. This is the second term in Equation 1.31. However in a periodic system, there are infinite nuclei involved, and the potential is therefore impossible to compute through direct summation.

A naïve approach would be to choose a large spherical region of the crystal, and, for each KS equation, consider only the nuclei of the crystal that fall within the chosen radius. Unfortunately, the Madelung sum—the electrostatic potential of the whole lattice—is known to be a slowly and conditionally convergent property.^{21,62} The former is due to the fact that Coulomb decay as an inverse law with distance, which is much further reaching than other forms of interatomic interactions like exchange. The latter can be understood from the fact that lattice has discrete sites in space which are not equally spaced from a particular point. Thus a radial truncation will necessarily be imbalanced towards some directions more than others.

The long-range Coulomb interactions with distant atomic centres are therefore traditionally evaluated using the Ewald summation technique^{63–65}. The expression for the Ewald potential at position r is:

$$V^{\text{Ewald}}(\mathbf{r}) = \sum_{L_s} q_s \frac{\text{erfc}(\gamma|\mathbf{r} - \mathbf{L} - \mathbf{R}_s|)}{|\mathbf{r} - \mathbf{L} - \mathbf{R}_s|} + \frac{4\pi}{v_c} \sum_{\mathbf{G} \neq \mathbf{0}} \frac{1}{G^2} e^{-G^2/4\gamma^2} \left[\sum_s q_s e^{i\mathbf{G}(\mathbf{r} - \mathbf{R}_s)} \right] \quad (1.46)$$

Where L and G are the real and reciprocal space lattice translations, q_s are the charges of each site s of the unit cell at positions \mathbf{R}_s , γ is the Ewald constant and v_c is the volume of the unit cell. Here, the direct sum electrostatic potential has been recast as a sum of two rapidly converging series. Short-range Coulomb terms are calculated in direct space and long-range interactions using a Fourier series in reciprocal space. To evaluate the Ewald potential on a lattice site \mathbf{R}_i , the self-potential of the charge must be subtracted to avoid a singularity which amounts to replacing the $L = 0$ and $s = i$ case of the first term of Equation 1.46 with $-\frac{2\gamma q_i}{\sqrt{\pi}}$.^{21,62,66,67}

1.4 Excited States

The methods outlined above focused on obtaining accurate and affordable ground state properties of chemical systems. However many processes of practical interest involve excited state systems which require specialised methods to model. We will first show how to extend DFT to the excited state and then focus on wavefunction based methods.

1.4.1 Linear-Response

We wish to obtain observable quantities, such as the density, from an optically excited system. This can be understood as investigating a system which starts in the ground state and is perturbed by a time-dependent potential $v(\mathbf{r}, t)$ infinitesimally later than a time t_0 , where this potential will have the same frequency as the incident photon.

The difference in density between a time t and the initial perturbation time t_0 is called the *response* of the density ($n(t) - n(t_0)$). This response can be expanded in powers of $v(\mathbf{r}, t)$:

$$n(t) - n(t_0) = n_1(t) + n_2(t) + \dots \quad (1.47)$$

The first and largest term of this expansion is called the *linear response* of the density. It is written:

$$n_1(\mathbf{r}, t) = \int_0^\infty dt' \int d\mathbf{r}' \chi_{nn}(\mathbf{r}, \mathbf{r}', t - t') v_1(\mathbf{r}, t) \quad (1.48)$$

Where $v_1(\mathbf{r}, t)$ is the linearised time-dependent potential and $\chi_{nm}(\mathbf{r}, \mathbf{r}', t - t')$ is the density-density response function defined as:

$$\chi_{nm}(\mathbf{r}, \mathbf{r}', t - t') = -i\theta(t - t') \langle \Psi_0 | [\hat{n}(\mathbf{r}', t - t'), \hat{n}(\mathbf{r})] | \Psi_0 \rangle \quad (1.49)$$

Where we have introduced the density operator $\hat{n}(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$ where the summation is over all electrons.

This expression can be recast into the frequency domain and shown to depend on the excitation energies of the system $\Omega_n = E_n - E_0$ where the energies are associated with the excited state wavefunctions Ψ_n :

$$\chi_{nm}(\mathbf{r}, \mathbf{r}', \omega) = \lim_{\eta \rightarrow +0} \sum_{n=1}^{\infty} \left\{ \frac{\langle \Psi_0 | \hat{n}(\mathbf{r}) | \Psi_n \rangle \langle \Psi_n | \hat{n}(\mathbf{r}') | \Psi_0 \rangle}{\omega - \Omega_n + i\eta} - \frac{\langle \Psi_0 | \hat{n}(\mathbf{r}') | \Psi_n \rangle \langle \Psi_n | \hat{n}(\mathbf{r}) | \Psi_0 \rangle}{\omega + \Omega_n + i\eta} \right\} \quad (1.50)$$

Where ω is the frequency of the perturbation, and η is a vanishingly small positive scalar chosen to avoid unphysical infinite response values. This is known as the Lehmann representation, and shows that if we can obtain an expression for $\chi_{nm}(\mathbf{r}, \mathbf{r}', \omega)$, we know that its poles will yield the excitation energies of the system.

1.4.2 Time-Dependent Density Functional Theory

We wish to use some of the pre-established formalisms of DFT in our efforts to define a DFT density-density response function, hopefully culminating in a time-dependent DFT (TDDFT). We begin with the fundamental theorem of TDDFT:

- The time-dependent external potential of a quantum system $v(\mathbf{r}, t)$ has a one-to-one correspondence with its time-dependent electron density $n(\mathbf{r}, t)$ up to a time-dependent phase.

This is known as the Runge-Gross theorem and it ensures that the correspondence established before t_0 by the first Hohenberg-Kohn theorem is maintained during the perturbation of the system. The second theorem can be cast into the time-dependent case by showing that the time-dependent density determines the action of the quantum system, which itself should be minimised by the variational theorem for the real wavefunction.

1.4.3 The Kohn-Sham Response Function

We are therefore justified in taking the steps of calculating the density-density response function in the non-interacting KS system. The principal difference will be that the linearised external potential has the form:

$$v_{s1} = v_1(\mathbf{r}, t) + \int \frac{n_1(\mathbf{r}, t)}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{xc1}(\mathbf{r}, t) \quad (1.51)$$

Where the first term is the linearised perturbation potential, the second is the Hartree potential, and the third the linearised exchange-correlation potential:

$$v_{xc1}(\mathbf{r}, t) = \iint \frac{\delta v_{xc}[n](\mathbf{r}, t)}{\delta n(\mathbf{r}', t)} \Big|_{n(\mathbf{r}, t_0)} n_1(\mathbf{r}', t') d\mathbf{r}' dt' \quad (1.52)$$

Where we use the shorthand:

$$f_{xc}(\mathbf{r}, t, \mathbf{r}', t') = \frac{\delta v_{xc}[n](\mathbf{r}, t)}{\delta n(\mathbf{r}', t)} \Big|_{n(\mathbf{r}, t_0)} \quad (1.53)$$

This is called the exchange-correlation kernel and is analogous to the exchange-correlation functional in ground state DFT, where the same approximations can be employed. In practice, the time dependence of this kernel implies that it depends on all past fluctuations of the density. An approximation to circumvent this tremendous requirement is to fix the kernel to its value at the beginning of the perturbation. This constitutes the *adiabatic local density approximation*.

A further improvement is to cast the kernel into the frequency domain, then written $f_{xc}(\mathbf{r}, \mathbf{r}', \omega)$, simplifying our task by using the periodicity of the perturbation. This is called the frequency dependent LDA.

Still in the frequency domain, we can construct a KS linear response of the density by means of a Fourier transform:

$$n_1(\mathbf{r}, \omega) = \int \chi_s(\mathbf{r}, \mathbf{r}', \omega) \left[v_1(\mathbf{r}', \omega) + \int \left(\frac{1}{|\mathbf{r}' - \mathbf{x}|} + f_{xc}(\mathbf{r}', \mathbf{x}, \omega) \right) n_1(\mathbf{x}, \omega) d\mathbf{x} \right] d\mathbf{r}' \quad (1.54)$$

Where the KS density-density response function is:

$$\chi_s(\mathbf{r}, \mathbf{r}', \omega) = \sum_{j,k=1}^{\infty} (f_k - f_j) \frac{\varphi_j^0(\mathbf{r})\varphi_k^{0*}(\mathbf{r})\varphi_j^{0*}(\mathbf{r}')\varphi_k^0(\mathbf{r}')}{\omega - \omega_{jk} + i\eta} \quad (1.55)$$

Where f_j and f_k are the occupancies of the KS orbitals at time t_0 , φ_j^0 and φ_k^0 , in the ground state and $\omega_{ij} = \epsilon_j - \epsilon_k$ where ϵ_j and ϵ_k are the orbital energies. Ideally, the KS original excitation energies ω_{ij} should match the real excitations Ω_n of Equation 1.50, relying upon the fact that the KS orbital transitions form a sufficient basis to represent this excitation. This hinges upon the form of the exchange-correlation functional, which explains its importance.

1.4.4 The Casida Formulation

We would like to formulate the above linear response problem into an eigenproblem, which is solvable in a direct way.

We first construct a coupling-matrix for the KS orbital transitions and add a spin index to each KS orbital:

$$K_{jk\sigma, j'k'\sigma'} = \iint \varphi_{j\sigma}^{0*}(\mathbf{r})\varphi_{k\sigma}^0(\mathbf{r}) \left(\frac{1}{|\mathbf{r} - \mathbf{r}'|} + f_{xc}(\mathbf{r}, \mathbf{r}', \Omega) \right) \varphi_{j'\sigma'}^{0*}(\mathbf{r}')\varphi_{k'\sigma'}^0(\mathbf{r}') d\mathbf{r}' d\mathbf{r} \quad (1.56)$$

We can write the tensors:

$$\begin{aligned} A_{ia\sigma, i'a'\sigma'}(\Omega) &= \delta_{ii'}\delta_{aa'}\delta_{\sigma\sigma'}\omega_{a'i'\sigma'} + K_{ia\sigma, i'a'\sigma'}(\Omega) \\ B_{ia\sigma, i'a'\sigma'}(\Omega) &= K_{ia\sigma, i'a'\sigma'}(\Omega) \end{aligned} \quad (1.57)$$

Each element of the $A_{ia\sigma, i'a'\sigma'}(\Omega)$ and $B_{ia\sigma, i'a'\sigma'}(\Omega)$ tensors represents an individual KS orbital transition. We can build an anti-Hermitian eigenvalue problem which will tell us the contributions from each of these transitions to a given excitation:

$$\begin{pmatrix} \mathbb{A} & \mathbb{B} \\ \mathbb{B} & \mathbb{A} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \Omega \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \quad (1.58)$$

Where the elements of \mathbf{X} and \mathbf{Y} can be shown to correspond to the orbital transition coefficients for the single excitation of energy Ω . These are called the Casida

equations,⁶⁸ and they enable us to efficiently calculate excitation energies, occupancies and oscillator strengths of excited states in TDDFT.

1.4.5 Algebraic Diagrammatic Construction

We aim to produce an excited state wavefunction of N electrons without using the density as the state function, which defines DFT. To do so, we first calculate all of the states resulting from the single excitations of a ground state wavefunction. Since they are not, in general, orthonormal, we employ successive Gram-Schmidt orthogonalizations to build a satisfactory basis of intermediate states $\{|\tilde{\Psi}_J^N\rangle\}$.⁶⁹

We now build the so-called Algebraic Diagrammatic Construction (ADC) matrix:

$$M_{IJ} = \langle \tilde{\Psi}_I^N | \hat{H} - E_0^N | \tilde{\Psi}_J^N \rangle \quad (1.59)$$

Where we have shifted the Hamiltonian by the ground state energy E_0^N . This matrix yields the eigenvalue equation:

$$\mathbf{MX} = \mathbf{X}\mathbf{\Omega} \quad (1.60)$$

Where $\mathbf{\Omega}$ is the diagonal matrix of eigenvalue excitation energies Ω_n and \mathbf{X} has values used to reconstruct the exact wavefunction from the basis of intermediate states like so:

$$|\Psi_n^N\rangle = \sum_J X_{nJ} |\tilde{\Psi}_J^N\rangle \quad (1.61)$$

To carry out this scheme, we need an accurate ground state wavefunction. A common choice is the MP2 wavefunction, which yields the ADC2 excited state wavefunction.^{70,71}

1.4.6 Excited State Coupled Cluster

CC methods can also be cast into an eigenvalue form for the excited states:⁷²

$$\mathbf{AR} = \mathbf{\Omega R} \quad (1.62)$$

Where the A matrix arises from linear response theory, much in the same spirit of Casida TDDFT:

$$A_{\mu_i\nu_j} = \langle \mu_i | e^{-\hat{T}} [\hat{H}, \hat{\tau}_{\mu_j}] | \Psi_{\text{CC}}^0 \rangle \quad (1.63)$$

Where $\hat{\tau}_{\mu_j}$ is the i -tuple excitation operator and μ_i the collection of orbitals involved in this excitation.

In this paradigm, one can employ the full CCSD expression for the ground state, and only use perturbative double excitations for the excited state. The resulting scheme is called CC2 and scales with a power of N^5 orbitals instead of the N^6 of CCSD.⁷³

1.4.7 Multireference Methods (bis)

Finally, multireference methods can also naturally be applied to the excited state, as can be understood from Section 1.1.9. CASSCF excited state calculations are usually calculated in a state-averaged way, in order to take into account the different excited states in the total density.

CASPT2 can apply its dynamic correlation correction to several state-averaged CASSCF wavefunctions simultaneously.³² This is called Multistate CASPT2 or MS- N -CASPT2 where N is the amount of excited states.⁷⁴ It is a method of state-of-the-art accuracy, for example predicting a vertical singlet excitation of 7.98 eV in ethene, compared to the best estimate of 7.8 eV, where CC2 predicts 8.40 eV.⁷⁵

Multireference methods are necessary to model highly nonadiabatic regions of the PES, such as bond breaking or conical intersections, which are discussed in Section 3.3.2. While ADC(2) and CC2 methods may be able to approximate the topology of certain conical intersections,⁷⁶ their pertinence is unpredictable, and CASSCF or CASPT2 are the most robust tools available for this task. In particular, TDDFT, a workhorse of excited state modelling, displays serious qualitative deficiencies in this domain.⁷⁷

Whilst the accounts of the accuracy of multireference methods for excited states are extensive,^{78–80} their formidable cost limits their use. Indeed whilst most single reference methods discussed herein scale their cost with power laws, multireference methods scale factorially.⁸¹ This imposes significant restrictions to the system

sizes and basis sets available to multireference methods. In particular, until recently, CASPT2 energy gradients could only be computed numerically,⁸² making geometry optimisation out of the capabilities of routine calculations on institutional clusters for practically sized molecules.

Chapter 2

Embedding Methods

2.1 Introduction

The electronic structure methods outlined in the previous chapter all succeed in achieving increased accuracy with greater computational cost. In the case of chemical processes in condensed phases, the amount of electrons representing every actor in the process becomes unmanageably large.

Fortunately, oftentimes, the process of interest is confined to a local area within the material. We may therefore adopt the strategy of selecting a molecular sample from the system, and performing calculations on this subsystem. Of course, the interaction of the environment with the subsystem can be determining in the process under study, and therefore often needs to be reflected in the computational simulation.

2.2 Polarisable Continuum Models

The first approximation to the environmental interactions between a solvent or crystal and a local site, is that of directly applying a model potential to the Hamiltonian of the quantum description of that site. This is in contrast to later discussed methods which explicitly represent the environment of the area of interest. As such, in the current case, the molecules under scrutiny are said to be placed in a *continuum*, and the large family of implicit solvation models are also called continuum models.⁸³

As a general rule, continuum models partition the simulation space between a cavity, which contains a large portion of the electron density of the simulated

molecule, and an infinite environment, meant to represent the medium. An effective method to reflect the electrostatic potential of the medium, is to reduce it to a potential on the surface of the cavity. The potential within the cavity is related to the surface charge $\sigma(\mathbf{s})$ as follows:⁸⁴

$$V_{\sigma}(\mathbf{r}) = \int_{\Gamma} \frac{\sigma(\mathbf{s})}{|\mathbf{r} - \mathbf{s}|} d\mathbf{s} \quad (2.1)$$

Where we are integrating over the surface of the cavity Γ . The correct form of $\sigma(\mathbf{s})$ determines the quality of this kind of continuum model. If we presume that the inner and outer regions of the cavity have different but isotropic dielectric permittivities, we can derive a surface charge:⁸⁵

$$\sigma(\mathbf{s}) = \frac{\epsilon - 1}{4\pi\epsilon} \frac{\partial}{\partial n} (V_M(\mathbf{s}) + V_{\sigma}(\mathbf{s})) \quad (2.2)$$

Where the ϵ is the medium's dielectric constant, $V_M(\sigma)$ and $V_{\sigma}(\sigma)$ are the electrostatic potentials emanating from the inner region charge density and the surface charge, and \mathbf{n} is the unit vector normal to the surface at point \mathbf{s} . The surface charge is a function of its own potential, making Equation 2.2 require self-consistent calculations to solve.

This method is known as the Dielectric Polarisable Continuum Model (DPCM, sometimes PCM). It has the advantage that it does not require additional explicit molecular calculations to evaluate the solvation properties of the sample, and the solvent is defined by a tunable dielectric constant. However it ignores non-electrostatic interactions, which are key to the structure of dense condensed phases. Moreover, it presumes an isotropy which, loses the ordered property of the Coulomb environment in crystal materials.

2.3 Point Charge Embedding

2.3.1 Non-isotropic Media

The electrostatic interactions between a molecular system and its environment can indeed be highly directional. In molecules with an asymmetric polarity, the periodic stacking of their crystal form means that one side of the molecule will be in

closer contact with electronegative atoms than the other. Thus, instead of using a continuum, it would be useful to locate the sources of electrostatic potential—the environment atoms—in space.

Given a certain partial charge q_i on an environment atom at position R_i , one can use atom centred point charges to represent the total electrostatic potential of the medium as such:

$$V^{\text{ES}}(\mathbf{r}) = \sum_i \frac{q_i}{|\mathbf{r} - \mathbf{R}_i|} \quad (2.3)$$

If we embed the system's Hamiltonian in the above potential, we can make its electron density respond to the environmental Coulomb potential.

Note that this scheme does not include any non-Coulombic forces, giving it a very poor description of short range interatomic interactions. Furthermore, while in the long range limit, the electrostatic potential from an atom does tend towards an inverse law, it does not have the infinite growth at $\mathbf{r} = \mathbf{R}_i$ which exists in Equation 2.3. It, instead, shows a finite cusp centred about the nucleus, called Kato's cusp, whose value depends on the charge of the nucleus.^{86,87}

Thus, we must be aware that the accuracy of the electrostatic potential emanating from point charges deteriorates as the distance from the nucleus becomes small. We call the error in electronic calculations caused by this discrepancy *overpolarisation*.⁸⁸ In such cases, with the point charge potential being infinitely attractive, the simulated charge density can be funnelled to the nuclear position. It is therefore important to choose a basis set which does not allow the wavefunction of the molecular system to extend onto nuclear positions in the environment.

2.4 General QM:MM

2.4.1 Cluster Models

An alternate method for modelling a multiscale system is by explicitly modelling the environment at a lower level of theory than the site of interest. This defines the family of *hybrid* methods, which combine different quantum or classical methods into one. At this point, it becomes important to concretise what would take the role

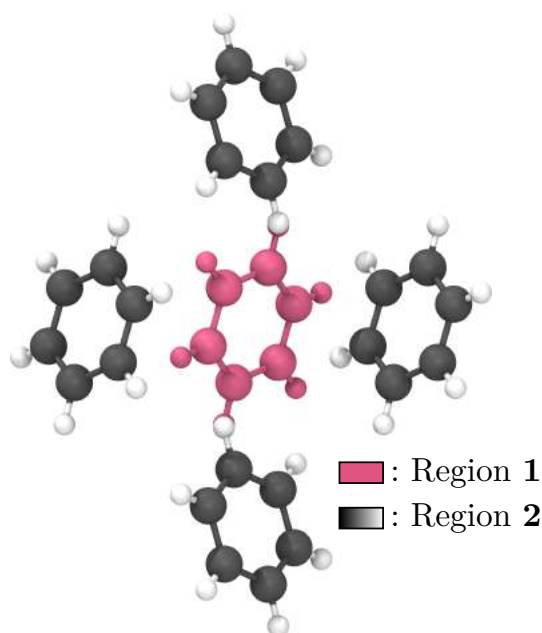


FIGURE 2.1: Example of a cluster of benzene atoms taken from their lattice positions. The pink molecule denotes the central region.

of the central region and environment in the case of molecular crystals. Thankfully, the non-covalent nature of these materials offers us an inherent partition.

If we care about a process localised on a few neighbouring molecules, then they become the high level of theory region (region 1), and a surrounding cluster of molecules taken from their crystal position becomes the low level of theory region (region 2). This partition is represented on Figure 2.1 for the case of a cluster of benzene molecules where the active site is one unit. The size and shape of this cluster is a topic of discussion, but the smallest useful one would include all of the nearest neighbours to the region 1 molecules, ensuring an isotropic coverage of the short range interactions.

As a side note, if we wanted to partition the cluster through a covalent bond, we would terminate the dangling bonds of region 1 and potentially 2 with appropriate fictitious monovalent atoms. This explains why in certain scenarios, region 1 is referred to as *model system*, in the literature. We forego this convention, since this discussion is centred around system-environment boundaries which do not cleave bonds.

2.4.2 Intermolecular Interactions

Our main goal now consists in modelling the energy of region $\mathbf{1} \cup \mathbf{2}$. We can write:

$$E_{\text{high:low}}(\mathbf{1} \cup \mathbf{2}) = E_{\text{high}}(\mathbf{1}) + E_{\text{low}}(\mathbf{2}) + E_{\text{high:low}}^{\text{int}}(\mathbf{1}, \mathbf{2}) \quad (2.4)$$

Where $E_l(n)$ is the energy of the isolated region n , at the level of theory l , and $E_{\text{high:low}}^{\text{int}}(\mathbf{1}, \mathbf{2})$ is the interaction energy between the two regions. The first developed hybrid methods, and still the most popular today, were those combining quantum mechanics methods with molecular mechanics methods (QM:MM). We can rewrite Equation 2.4 with QM:MM indices:⁸⁹

$$E_{\text{QM:MM}}(\mathbf{1} \cup \mathbf{2}) = E_{\text{QM}}(\mathbf{1}) + E_{\text{MM}}(\mathbf{2}) + E_{\text{QM:MM}}^{\text{int}}(\mathbf{1}, \mathbf{2}) \quad (2.5)$$

The crux of QM:MM techniques relies in finding an appropriate expression for $E_{\text{QM:MM}}^{\text{int}}(\mathbf{1}, \mathbf{2})$. The family of methods which focus on calculating the first two terms of Equation 2.5 with electronic structure and molecular mechanics packages, and then computing the interaction term explicitly are called *additive* QM:MM.

2.5 IMOMM

A parallel family of methods uses a different formulation, instead choosing to implicitly consider the interaction term *via* a subtraction. First, we consider the components of the energy of the whole system in at the MM level of theory:

$$E_{\text{MM}}(\mathbf{1} \cup \mathbf{2}) = E_{\text{MM}}(\mathbf{1}) + E_{\text{MM}}(\mathbf{2}) + E_{\text{MM}}^{\text{int}}(\mathbf{1}, \mathbf{2}) \quad (2.6)$$

We can reorder this as $E_{\text{MM}}(\mathbf{2}) + E_{\text{MM}}^{\text{int}}(\mathbf{1}, \mathbf{2}) = E_{\text{MM}}(\mathbf{1} \cup \mathbf{2}) - E_{\text{MM}}(\mathbf{1})$, and insert it in Equation 2.5, provided that we only require an MM level description of the interaction energy. The full Integrated Molecular Orbital Molecular Mechanics (IMOMM), energy is finally written:⁹⁰

$$E_{\text{QM:MM}}^{\text{ONIOM}}(\mathbf{1} \cup \mathbf{2}) = E_{\text{QM}}(\mathbf{1}) + E_{\text{MM}}(\mathbf{1} \cup \mathbf{2}) - E_{\text{MM}}(\mathbf{1}) \quad (2.7)$$

The meaning of the term “ONIOM”, is explained later. The energy gradients can naturally be calculated in the same way. For instance the force on region 1 atoms is the negative of the following:

$$\frac{dE_{\text{QM:MM}}^{\text{ONIOM}}(\mathbf{1} \cup \mathbf{2})}{d\mathbf{1}} = \frac{dE_{\text{QM}}(\mathbf{1})}{d\mathbf{1}} + \frac{dE_{\text{MM}}(\mathbf{1} \cup \mathbf{2})}{d\mathbf{1}} - \frac{dE_{\text{MM}}(\mathbf{1})}{d\mathbf{1}} \quad (2.8)$$

Let us carefully compare Equations 2.5 and 2.7. In the former, only one MM calculation is carried out, as opposed to two in the latter, including one necessarily larger one ($\mathbf{1} \cup \mathbf{2}$). In our case, this is not a significant issue, since we will wish to push the accuracy of the QM level of theory, making it, and not any MM calculations, the computational bottleneck of the method.

We have also gotten rid of any explicit calculation of the interaction term. This is a promising feature, since the components of the intermolecular interaction are difficult to express and disentangle, however we are now left with the inter-region interactions reflected only by their MM description.

2.6 QM:MM Electrostatic Embedding

One significant issue of the MM description of inter-region interactions is that they do not capture any electronic response from the region 1 atoms to the environment at a quantum level of theory. This is a particularly severe shortcoming for electrostatic interactions, which are far-reaching by nature, and can determine the electronic structure of the investigated ground or excited states.

To overcome this difficulty, the inter-region electrostatic energy in the second term of Equation 2.7 must be removed—this is a known quantity in MM calculations—and injected in the first term. This method is called *electrostatic embedding*, and hybrid methods without it are said to use *mechanical embedding*. The most common electrostatic embedding method complements the QM Hamiltonian of region 1 with point charges at the location of region 2 atoms, with charge values determined by the forcefield. This allows the electron density of region 1 to respond to the electrostatic potential of the environment, as is discussed in Section 2.3.

2.7 ONIOM QM:QM'

If we want to increase the level of accuracy of the inter-region interactions, while still using a subtractive scheme, we must use a better low level of theory than MM. The Integrated Molecular Orbital Molecular Orbital (IMOMO) method, does just this, yielding an energy equation of the form:⁹

$$E_{\text{QM:QM}'}^{\text{ONIOM}}(\mathbf{1} \cup \mathbf{2}) = E_{\text{QM}}(\mathbf{1}) + E_{\text{QM}'}(\mathbf{1} \cup \mathbf{2}) - E_{\text{QM}'}(\mathbf{1}) \quad (2.9)$$

Where QM' is a quantum level of theory lower in accuracy than QM. At this point, it is worth mentioning that these hybrid method schemes can be extended to include as many regions and methods as necessary. The general term for subtractive schemes, regardless of number of layers and type of modelling method, is *our own n-layered integrated molecular orbital and molecular mechanics* (ONIOM).¹⁰ ONIOM is a superset of IMOMM and IMOMO, and we therefore employ it in our energy expressions since specifying the levels of theory involved renders the distinction between IMOMM or IMOMO superfluous.

If we want to keep using electrostatic embedding for IMOMO, the implementation is more delicate than for IMOMM. Here, the inter-region Coulomb interaction is not a known quantity which could be removed from the second term of Equation 2.9. Instead, we must embed the third term in point charges, thus cancelling out the electrostatic interaction present in the second. A discussion of the choice of point charge can be found in Chapter 4.

2.8 Shortcomings of ONIOM Embedding Methods

This section serves as a motivation for the rest of the work presented in this thesis. As such, some of the problems of ONIOM methods are briefly discussed now, and are addressed in Chapter 4.

2.8.1 Inclusion of Long-Range Effects

In ONIOM schemes, region 1 is subject to the environmental interaction of region 2. A region 2 containing only the nearest neighbour molecules should suffice to cover

the non-electrostatic effect, owing to their short-range nature.⁹¹ However electrostatic effects decay as an inverse law, and therefore can have a significant impact on the electronic state of region 1.

It is unclear how many layers of molecules in region 2 would suffice to represent the screened electrostatic potential in an amorphous material, but the larger the region, the higher the cost to model computationally. Moreover, as discussed in Section 1.3.6, in a crystal, the accuracy of the potential does not systematically improve with a larger region, making traditional ONIOM a poor choice if long-range effects have a role in the system under study.

2.8.2 Point Charge Representation

The use of point charge embedding to represent electrostatic effects presumes a similarity between the Coulomb potential of a point charge and that of an atomic charge density. We have already discussed a discrepancy in the short-range in Section 2.3, but how can we ensure that the point charge values which we employ for each atom will have the correct behaviour in the mid-to-long-range?

Indeed, there is no absolute definition of an atom's partial charge, and the variety of population analysis schemes available to use suggests that some will be more appropriate than others in representing the electrostatic potential of the atom for the embedding of region 1 at high and low levels of theory. A successful scheme, combined with valid electronic structure methods, would reproduce accurate absorption and emission energies of molecules in their crystal environment, whilst fulfilling an appropriate cancellation of errors in the ONIOM equation.

A commonly used method for both is that of the Mulliken analysis,⁹² which essentially accounts for the contribution of individual atomic orbitals to the overall electron density, and then sums up this contribution between all the orbitals belonging to one atom.⁹³

Some known problems are the fact that the electron population being shared between two orbitals of different atoms will necessarily be split in half, not accounting for the potential discrepancy between the electron affinity of the nuclei. Furthermore, due to being directly defined by the basis set of the calculation, Mulliken

charges are highly basis set dependent. For instance the charge values on a water molecule can vary up to $0.7 e^-$ depending on the basis set.⁹⁴

It is therefore unclear whether the most commonly used partial charge scheme is appropriate for electrostatic embedding. We ought to seek out alternatives, and compare the effect of their embedding on the electronic structure of region **1** molecules.

2.8.3 Mutual Polarisation

The electrostatic embedding methods described above allow for the representation of the static Coulomb interaction of the crystal on the active site. This one-way relation is not an accurate representation of the real scenario where the regions **1** and **2** mutually polarise each other until they find an equilibrium at fixed nuclear coordinates, following the BO approximation.

Mutual polarisation plays an important role in certain excited state processes. For instance, the charge transfer states in LiF interacting with a Be atom were probed with different kinds of QM:MM implementations; and the non-polarisable one overestimated excited state energies by over 0.1 eV.⁹⁵ It is as of yet unclear how important these errors become in molecular crystals, and which methods should be employed to mitigate them.

2.8.4 Implementation

Finally, QM:QM' techniques face many issues in terms of their software implementation. On a technical level, the electronic structure programs which offer this feature are few and far between. Furthermore, each of those programs only implements a limited set of electronic structure methods, thus limiting the amount of QM and QM' method combinations available to the user. When modelling excited states, having as much flexibility as possible in the choice of method is crucial, due to the variety in character of many body excited state wavefunctions.

On a practical level, the task of selecting a region **1** and **2** combination, and assigning point charges to pertinent lattice positions is laborious, and difficult to incorporate in a reproducible workflow. For instance, *nearest neighbour molecules* is a term with no formal definition, and while they can often be recognised at a glance,

edge cases are certainly subjective to the researcher. An algorithmic method for the practical steps of setting up these calculations is therefore desirable.

Chapter 3

Photochemistry

With the arsenal of methods discussed in Chapters 1 and 2, we are in a position to model some photochemical systems in the excited state. This chapter outlines the theory behind photochemistry in the vacuum and aggregate phases so that we can better understand the phenomena we are trying to model.

3.1 Absorption

3.1.1 Principles of Photochemistry

Light-induced chemical processes follow two fundamental laws:⁹⁶

- **Grotthuss–Draper Law:** Only light which is absorbed by the chemical system can produce change.
- **Stark–Einstein Law:** The energy of each absorbed photon corresponds to the energy difference between the initial and newly excited states of the system.

The combination of these laws implies that in order for a photochemical process to occur, the incident photons to the chromophore need to be of an energy matching an allowed electronic transition. This gain in energy takes the wavefunction from its original ground state to a new, excited state one.

In the BO approximation, the energy of the molecular wavefunction depends parametrically on the nuclear coordinates. This implies that for a given electronic excitation state, there is a continuous hyperdimensional landscape which maps a set of nuclear coordinates to a potential energy. Such Potential Energy Surfaces (PESs) are represented in Figure 3.1, displaying different excited state mechanisms (see later subsections).

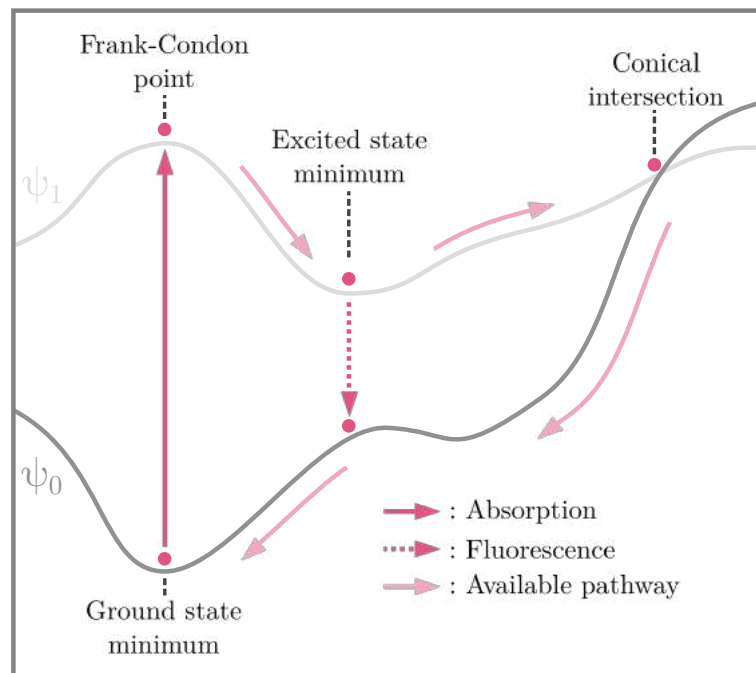


FIGURE 3.1: Example of a potential energy surface demonstrating either a radiative decay *via* fluorescence, or a nonradiative decay *via* funnelling through a conical intersection.

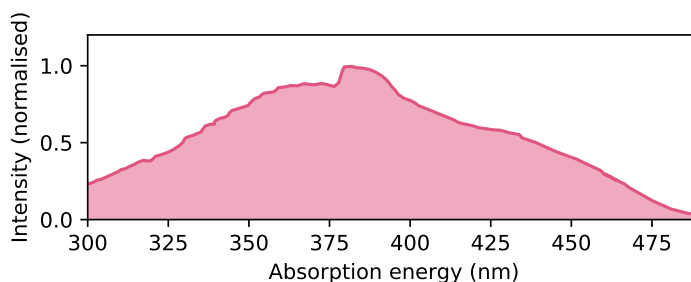


FIGURE 3.2: Absorption spectrum of (2-hydroxyphenyl)propanone (HPP) in single crystal form.⁹⁷

First, bringing our attention to the absorption process, we can see how the difference in energy between two potential energy surfaces at a fixed ground state geometry would correspond to the energy of one photon. Even considering all of the possible excitations present in the full CI expansion of Equation 1.19, electronically excited states only occupy discrete levels of energy, in this limited picture. In reality, molecular and crystalline systems absorb as continuous spectra across the electromagnetic range, as shown in Figure 3.2 for the 2-hydroxyphenyl)propanone single crystal. This material holds promise for solid state luminescent applications such as lasers or OLEDs.⁹⁷

Several effects contribute to the broadening and mixing of the excitation energies, shaping this absorption spectrum.

3.1.2 Oscillator Strength

First, we will investigate what the probability is of photons being absorbed for a given transition.

To answer this, we begin by examining the transition dipole moment (TDM) between an initial state i and final state j :

$$(\mathbf{T}_{ij})_x = \int \psi_j^*(\mathbf{r}) \hat{x} \psi_i(\mathbf{r}) d\mathbf{r} \quad (3.1)$$

Where $\psi_n(\mathbf{r})$ is the wavefunction of state n . This equation gives the x component of the total vector TDM, \mathbf{T}_{ij} . The probability of transition from one quantum state to another is worked out from the quantum electro-dynamical treatment of spontaneous emission. It is proportional to the square of the TDM, and to the energy gap between the states:⁹⁸

$$f_{ij} \propto (E_j - E_i) \left| \mathbf{T}_{ij} \right|^2 \quad (3.2)$$

Where E_n is the energy of state n , and f_{ij} is called the *oscillator strength* of the transition $i \rightarrow j$. The oscillator strength can therefore serve as a measure of the likelihood of population of different excited states, depending on the frequency of the incident light.

3.1.3 Natural Linewidth

For a given molecule, we therefore have several energy levels at which photons can be absorbed, with different probabilities. Additionally, we should consider the energy-time form of the uncertainty principle:

$$\Delta E \Delta t \geq \frac{\hbar}{2} \quad (3.3)$$

Which imposes a minimum on the product between the uncertainty in the excited state energy ΔE and the uncertainty in its lifetime Δt . There is therefore a so called

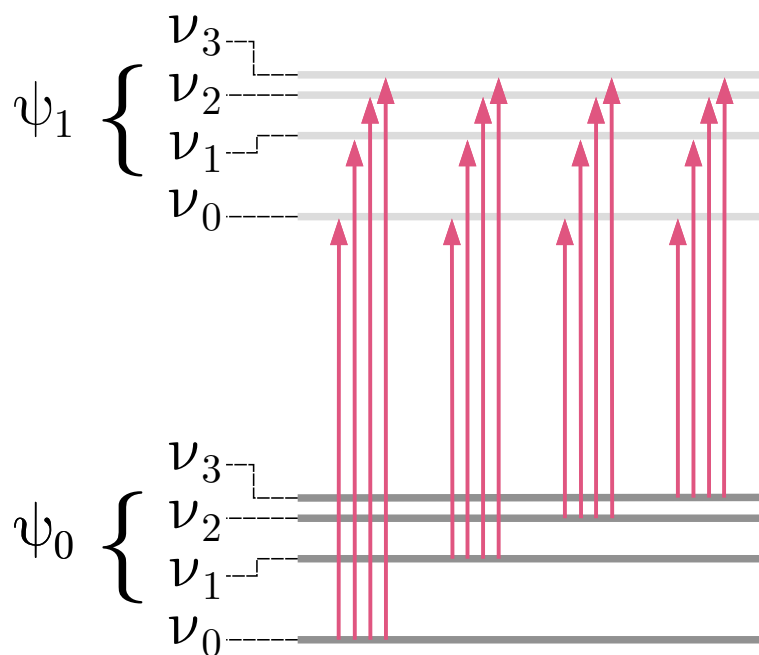


FIGURE 3.3: Diagram showing the manifold of transitions between two electronic states, each with four vibrational substates. The electronic wavefunctions are denoted ψ_n and the vibrational states ν_m .

natural linewidth which ensures that incident photons do not need to match the energy difference between states at infinite precision.

3.1.4 Role of Vibrations

Note that the wavefunctions in Equation 3.1 have both an electronic and a nuclear part. Working under the BO approximation, these parts can be factorised, and the transition dipole moment ends up accounting for both the change in dipole from the transition of the electron cloud, and the overlap between the nuclear vibrational states. Labelling the different potential states involved in a transition as ψ_{n,ν_m} , for the m th electronic excitation and M th vibrational substate, we end up with a large amount of possible combinations, depicted schematically in Figure 3.3.

Moreover, for periodic materials, we consider lattice vibrations rather than molecular ones, and bands of energy rather than discrete levels. This further contributes to the continuity of the spectrum in Figure 3.2

3.2 Emission

We now return to Figure 3.1 to understand what can happen to the molecule after excitation. If the molecular system is at rest at the time of absorption, its nuclear coordinates will be at the point of lowest energy, the ground state geometry. The corresponding excited state at the same geometry is reached after absorption, and is referred to as the *Frank-Condon point* (FC). Since the FC geometry is unlikely to be a minimum of the excited state potential energy surface, the molecule will have enough energy to travel downhill (or through tunnelling, or extra vibrational energy) to new nuclear configurations.

3.2.1 Fluorescence

When the excited state reaches a minimum geometry, the molecule naturally oscillates around this minimum. However excited states are by definition metastable, and therefore the energy difference between ground and excited state must be released in some way. Such transitions from states of same multiplicity are called *fluorescence*. When a molecule fluoresces, it emits a photon of the frequency corresponding to the difference in energy between excited and ground state.

The lifetime of this process is dependent on the overlap of the wavefunctions of initial and final state (see Section 3.1.2) and the relative populations of the two states within the chemical system. The result is usually lifetimes of the order of nanoseconds.

3.2.2 Phosphorescence

In certain cases, as a part of the excited state mechanism, the wavefunction may enter a state of different multiplicity than the ground state. In this case, after reaching the excited state minimum, the transition between states is classically forbidden due to the spin flip involved. The process may still occur, but the lifetime of phosphorescence can reach minutes or hours.

3.3 Nonradiative Decay

Electronic wavefunctions can also decay from the excited to the ground state nonradiatively. Understanding how to limit this behaviour is essential in designing efficient luminescent materials.

3.3.1 Vibrational Internal Conversion

Fermi's Golden Rule (FGR) gives an expression for the rate of transfer between quantum states under small perturbations.⁹⁹ The nonradiative transition between separate excited state of the same spin multiplicity in the adiabatic regime is called internal conversion. Derived from perturbation theory on the TDSE, FGR states:

$$k_{\text{IC}}^{\text{Fermi}} = \frac{2\pi}{\hbar} \sum_{v_i, v_f} P_{v_i}(T) \left| \sum_k \langle \Psi_{f, v_f} | \hat{P}_k | \Psi_{i, v_i} \rangle \right|^2 \delta(E_{i v_i} - E_{f v_f}) \quad (3.4)$$

Where $P_{v_i}(T)$ is the Boltzmann distribution at temperature T of the vibrational manifold at the initial electronic state, Ψ_{n, v_n} is the wavefunction at electronically excited state n and vibrational state v_n , and \hat{P}_k are the nuclear momentum operators for the k -th normal mode of vibration.

In the BO approximation, we can separate the wavefunction $|\Psi_{i v_i}\rangle = |\Phi_i \Theta_{i v_i}\rangle$ into a product $|\Psi_{i v_i}\rangle = |\Phi_i\rangle |\Theta_{i v_i}\rangle$, where Φ_i is the electronic wavefunction and $\Theta_{i v_i}$ is the vibrational wavefunction.

This allows us to rearrange the internal conversion rate as follows:^{100,101}

$$k_{\text{IC}}^{\text{Fermi}} = \frac{2\pi}{\hbar} \sum_{v_i, v_f} P_{v_i}(T) \left| \sum_k \langle \Phi_f | \hat{P}_k | \Phi_i \rangle \langle \Theta_{f v_f} | \hat{P}_k | \Theta_{i v_i} \rangle \right|^2 \delta(E_{i v_i} - E_{f v_f}) \quad (3.5)$$

All else being fixed, this allows us to express the rate of internal conversion in terms of the overlap between vibrational wavefunctions, which can be rationalised as the energy barrier between PESs being overcome by thermal vibrations.

Under FGR, the transition occurs between degenerate vibrational states rather than necessarily degenerate electronic ones. Because of this, the numerical methods for solving Equation 3.5 use approximations for the nuclear motion of the molecule,

such as a harmonic potential. This potentially leaves out points of degeneracy far from excited state minima.

3.3.2 Conical Intersections

In certain scenarios, however, the excited and ground electronic potential energy surfaces are prone to meeting, as is shown in the rightmost phenomenon of Figure 3.1. In this case, the time-dependent SE has two degenerate solutions. Additionally, the coupling between excited and ground states also has critical points, where the adiabatic separation between them breaks down. The points which combine these two conditions allow for the funnelling of the wavefunction from the excited to the ground state without emission at any point.

We can inspect these exceptional cases by enclosing the TISE in wavefunctions in the Born-Huang representation:

$$|\Psi\rangle = \sum_i^{\infty} |\Phi_i\rangle |\Theta_i\rangle \quad (3.6)$$

Where we have separated the electronic wavefunction Φ from the nuclear one Θ and labeled the excited state i .

The result is that sufficient condition for this particular degeneracy to occur between states i and j , is the following:

$$v_{ij} = \frac{\langle \Phi_i | \frac{\partial \hat{H}^{\text{el}}}{\partial \mathbf{R}} | \Phi_j \rangle}{E_j - E_i} \rightarrow \infty \quad (3.7)$$

$$g_{ij} = \frac{\partial}{\partial \mathbf{R}} \left(\langle \Phi_i | \hat{H}^{\text{el}} | \Phi_i \rangle - \langle \Phi_j | \hat{H}^{\text{el}} | \Phi_j \rangle \right) = 0 \quad (3.8)$$

Where Φ_n is the electronic wavefunction of state n , \hat{H}^{el} is the electronic Hamiltonian, E_n is the energy of state n , and \mathbf{R} are the nuclear coordinates. v_{ij} is called the nonadiabatic coupling vector and g_{ij} is the derivative coupling vector.¹⁰²

The intersections are *conical* because they are determined by two coordinates in hyperspace. They should be understood as a multidimensional seam in the potential energy surfaces.¹⁰³ This leads us to consider the point of lowest energy along that seam with special interest. The Minimal Energy Conical Intersection (MECI), is an

important geometry to characterise, since energy relative to that of the FC geometry dictates its accessibility, and thus the availability of a nonradiative decay channel.

We should clarify the distinction between this decay and the internal conversion through thermal vibrations. The former involves a seam of electronically degenerate states, attainable through thermal motion. The latter directly has to do with the overlap of vibrational wavefunctions, and provides a probability of transition. Both decays can be attained through vibrations, though they are distinct processes which can collaborate or not. Conical intersection decay is the dominant one in the strongly nonadiabatic regions of the PES and *vice versa* elsewhere.¹⁰⁴

3.4 Excited States in Aggregates

3.4.1 Excitons

In molecular aggregates, it is possible for the excited state process to delocalise over several molecules. A collection of thus correlated excited states is called an exciton, when speaking at the molecular scale. As a semantic aside, when considering a lattice, excitons usually refer to an electron-hole pair; however, for small groups of molecules, this charge separation is usually much less distinct, and we call exciton what a spectroscopist might call an "exciplex". Our convention is in line with the language used in the field of molecular photochemistry at the aggregate scale.^{91,105,106}

For clarity of explanation, we will consider the case of two molecules, but know that much of what is discussed is generalisable to the trimer, tetramer, and so on.

Two neighbouring molecules are said to be excitonically coupled if their excited states in isolation are different to those as a dimer. We can quantify this exciton coupling by writing the Hamiltonian of the molecular system in the diabatic basis:

$$H^D = \begin{pmatrix} E_1^D & J \\ J & E_2^D \end{pmatrix} \quad (3.9)$$

where E_n^D is the diabatic energy of state n and J is the exciton coupling.

3.4.2 Kasha's Model

Obtaining the above Hamiltonian matrix is not a trivial effort, therefore a common approximation is employed to characterise the excitonic states is to reduce each molecule to a point dipole and investigate their electrostatic interaction upon transition. In this case, the relative arrangement of the point dipoles determines the exciton coupling. The expression for the exciton coupling between two molecules i and j can be derived by the Coulomb energy of two point dipoles. It is, in atomic units:¹⁰⁷

$$J_{ij} = \frac{1}{r_{ij}^3} \left(\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j - 3 \frac{(\boldsymbol{\mu}_i \cdot \mathbf{r}_{ij})(\mathbf{r}_{ij} \cdot \boldsymbol{\mu}_j)}{|\mathbf{r}_{ij}|^2} \right) \quad (3.10)$$

Where $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_j$ are the transition dipole moments for each molecule and \mathbf{r}_{ij} is the distance between their centroids. This approximation, known as Kasha's exciton model, is used to rationalise the blue or red shifting of the excitation of molecules in aggregate due to collective excitation.

The exciton coupling causes a splitting of the multi-molecular excited state, thereby selecting the ordering between the symmetric and antisymmetric combination, and the allowed same-symmetry radiative decay transition.⁶ Dimers with a negative coupling have the $\frac{|i\rangle+|j\rangle}{\sqrt{2}}$ wavefunction lower in energy than the $\frac{|i\rangle-|j\rangle}{\sqrt{2}}$ wavefunction, and are denoted J aggregates. The opposite is the case for H aggregates.

Due to the dot products in Equation 3.10, the sign of J_{ij} , and therefore the classification of the dimer, solely hinges upon the angle between transition dipole moments, and the vector between the monomer's centroids. The transition occurs at an angle of 54.7°.

Following Kasha's rule, decay to the ground state overwhelmingly happens from the lowest energy excited state. Therefore, for H aggregates, radiative decay is quenched and blue shifted, while for J aggregates, it is red shifted.

3.4.3 Types of Exciton

Of course, in reality, the electron density of the two monomers does not behave like a static point dipole. Beyond the missing short-range interactions and the higher resolution of the Coulomb potential, one should also take into account the reorganisation of the density inter- and intramolecularly.⁹¹

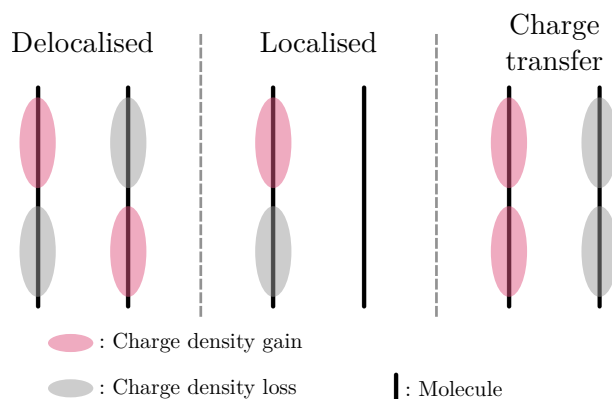


FIGURE 3.4: Visual representation of the three extreme types of excitons.

Indeed three extreme scenarios are possible, represented on Figure 3.4. Upon excitation, the electron density of each molecule can reorganise within and between them. The exciton is then said to be delocalised. In a different case, the reorganisation could be confined to only one molecule, making the exciton localised, occasionally called a Frenkel exciton. Finally, the electron density can migrate from one molecule to another, producing a charge transfer (CT) state.¹⁰⁸

CT states are of much relevance to photovoltaic applications due to their role in electron transport within a material. They also have a role in nonradiative decay mechanisms, where an excitation can split into a CT state and therefore dissipate the photon energy by releasing the electron-hole pair at the surfaces of the material.

3.5 Quantum Yield of Fluorescence

The net effect of fluorescent radiative and nonradiative mechanisms, is a ratio of emitted photons to absorbed photons. This is a measurable quantity, called the *quantum efficiency of fluorescence* (QEF), which is usually the principal result reported from experiments on luminescent materials:

$$\Phi_f = \frac{k_r}{k_r + k_{nr}} \quad (3.11)$$

Where k_r is the radiative decay rate, related to the oscillator strength, and k_{nr} is the non-radiative decay rate, principally related to the accessibility of conical intersections, the exciton hopping rate, and the coupling between vibrational states. Additional mechanisms such as intersystem crossing are possible, but not relevant to this work.

An accurate modelling of the excited state potentially energy landscape of a molecule informs us of the oscillator strength of the potential transitions, their energy gaps, and the relative energy of the MECI. Examining the vibrational phase space of the molecules and the electronic structure of the excited state aggregates completes the picture of the nonradiative part of the QEF.

Of course the challenge is modelling all of these within the crystal environment of the excited state process.

3.6 Example of Application: Four-Level Lasers

To further motivate the importance of accurately modelling photochemistry in crystals, this section briefly explains the mechanism behind a four-level laser, and how it relates to the concepts presented in this chapter.

The full mechanism is represented in Figure 3.5. In a solid state organic molecular crystal laser, the medium can be pumped by an incident photon, taking it to its FC state from the ground state (GS). The excited molecule(s) then reorganise to a minimum geometry S_1^* nonradiatively (where the asterisk indicates the excited state), from which they undergo fluorescent decay to S_1 (still the same geometry but now in the ground state). Finally, the system relaxes back to the GS geometry through thermal vibration.

The four-level structure is motivated by the expression for the Maxwell-Boltzmann thermodynamic equilibrium between two states:

$$\frac{n_2}{n_1} = e^{-\frac{E_2-E_1}{kT}} \quad (3.12)$$

Where k is the Boltzmann constant, E_i is the energy of state i , and T is the temperature. The left-hand side of Equation 3.12 is vanishingly small, indicating that at an equilibrium, $n_1 \gg n_2$. To encourage emission, we should invert this ratio, and take

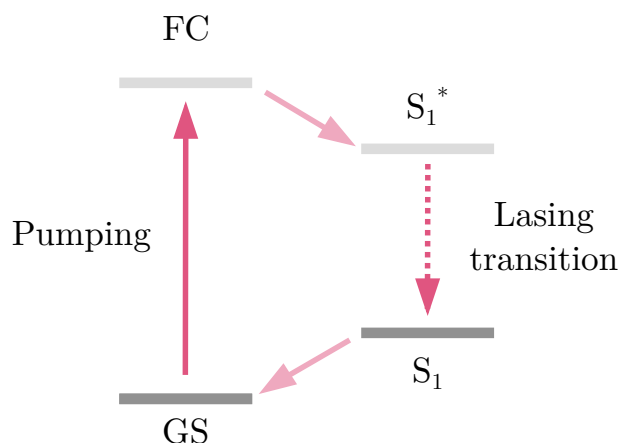


FIGURE 3.5: Diagram of the mechanism behind a four-level laser. Each step in the photocycle is labelled with its usual notation—Ground State (GS), Frank-Condon (FC), excited state minimum (S_1^*) and the corresponding ground state (S_1).

the system out of equilibrium. In our four-level laser example, the 1 state, labelled S_1 , is very depopulated compared to S_1^* due to the efficient vibrational decay which takes the ground state molecule from its S_1 geometry to FC. Thus we have $n_{S_1^*} \gg n_{S_1}$, enabling population inversion.

Additionally, the efficiency of the laser depends on the QEF of the lasing transition. A large Stokes shift is also desirable, to minimise reabsorption, therefore requiring a substantial reorganisation between the FC and S_1^* geometries. Furthermore, if the lasing transition is only accessible through S_0 – S_1 absorption, then the S_1 – S_2 energy gap should be large enough to discourage absorption to S_2 and higher states.

Such specific properties make organic single crystals ideal candidates for laser media. The conformational variety of their crystal packing makes their intermolecular properties tunable, which can principle be exploited to tailor optimal materials. Reference 6, reviews a large list of different organic single crystals which desirable emissive properties spanning across the visible spectrum.

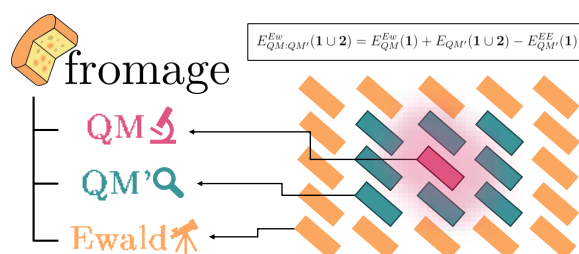
Part II

Results

Chapter 4

Ewald Embedded Cluster

This chapter is adapted from the published version of Reference 61.



4.1 Introduction

Excitations in organic molecular crystals have been found to localise on only a few constituent monomers of the crystal.¹⁰⁵ This poses a challenge for traditional electronic structure methods, which have been designed to describe either highly localised or periodic delocalised electronic states. In this context, embedding techniques, such as those discussed in Chapter 2, represent a viable option by combining higher quantum mechanical levels of theory to describe the excited region (QM) and more approximate methods for the crystal environment (QM' or MM).¹⁰⁹

Within the ONIOM scheme, the QM' method can be chosen to be plane-wave DFT^{11,12} for a natural description of the lattice periodicity, although this usually means sacrificing the electrostatic embedding. Correlated wavefunction-in-DFT periodic embedding approaches are a promising alternative^{110–112}. One of the most common approaches is to use cluster models to describe the periodic crystal.^{13–15} The cluster is extracted from the atomic lattice positions and provides an energetic description of the short-range interactions with the QM region.

In the case of ionic or highly polar crystals, long-range interactions can be of great importance since the electrostatic potential is slowly and conditionally convergent. This prompts the use of the Ewald summation, to evaluate long-range Coulomb interactions in the crystal (see Section 1.3.6).

When considering embedded finite cluster models, the electrostatic embedding can be modified to reflect the Ewald potential. In this case, the electrostatic interactions affecting the QM region extend beyond just the short-range and up to the infinitely large in a periodic system. Klintonberg *et al.* developed a methodology where a large array of point charges is fitted to reproduce the exact Ewald potential inside the QM region of a cluster model.^{63,113,114} This procedure has been used for the investigation of ionic crystals and the calculation of NMR parameters in organic crystals.^{63,115,116} Sokol *et al.* have implemented a related method in Chemshell to model defects in ionic materials.^{117,118} An alternative is the procedure proposed by Abrenkov and Sushko, where compensating point charges are added within unit cells to approach the Ewald potential.^{119,120}

Ewald embedding methods have been used with QM:MM and ONIOM approaches allowing the evaluation of the short-range non-Coulombic interactions.^{121–126} However a simpler variant is the Point Charge Embedding approach (PCE) where only the Coulomb interactions are considered, using point charges, and non-electrostatic interactions are neglected.^{63,109} The performance of these methods for the investigation of excited states PESs of molecular crystals is relatively unexplored. Recently, Ciofini and co-workers^{16,17,127} have implemented an Ewald PCE scheme based on the method proposed by Derenzo *et al.*¹¹⁴ In order to consider mutual polarisation effects of the crystal environment, a self-consistent algorithm was employed in the investigation of a crystal displaying aggregation-induced emission.¹⁶ Self-consistent schemes are typical tools used in QM:MM schemes when the polarisation of the environment is important.^{128–130}

In this chapter, we present different Ewald embedding approaches for the description of PESs of molecular crystals, with specific focus on the treatment of excited state minima and conical intersections. We show that due to the lack of short-range non-Coulombic interactions, geometry optimisation with the PCE method can be extremely problematic. As a solution, we implement an Ewald-embedded QM:QM'

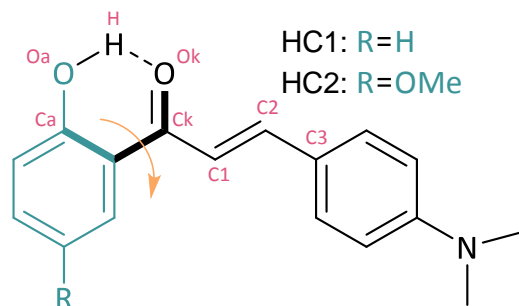


FIGURE 4.1: Molecular diagram of the 2'-hydroxychalcone derivatives, **HC1** and **HC2**. The access to conical intersections for these molecules is centred around the rotation of the blue group about the dihedral angle shown in bold. The notable atoms with large partial charge are labelled in pink.

cluster model that can be used to explore the PES of flexible molecules. We assess the efficacy of these schemes with two crystals based on 2'-hydroxychalcone (**HC1** and **HC2**, shown in Figure 4.1). These molecules undergo excited state intramolecular proton transfer (ESIPT), where the large changes in electronic structure in the excited state pose a challenge to embedding methods. The figure represents the enol (**E**) form. The keto form (**K**), where the proton is bonded to the other oxygen atom, is unstable in the ground state.

Our group has previously investigated **HC1** and **HC2** in the context of Solid State Luminescent Enhancement (SSLE).^{131,132} We opt for the term SSLE over Aggregation Induced Emission (AIE), which is more popular in the field, because it is more specific, as outlined in Reference 133. The population of the keto (**K***) and enol (**E***) excited states depends on the identity of the substituents and crystal packing.¹³² **HC1** displays emission in the crystal with promising properties to be used in solid state lasers¹³⁴ and predominantly forms herringbone-type aggregates. In contrast, **HC2**'s decay is mainly non-radiative and its crystal structure features mainly π -stack dimers. Their excited state PESs were found to be particularly sensitive to the electrostatic environment. The SLE character of **HC1** can be understood using the Restricted Access to Conical Intersections (RACI) model^{135,136} wherein upon aggregation the energy of the S_1 - S_0 conical intersections increases, thereby blocking nonradiative deactivation pathways and enhancing the emissive response.

The chapter is organised as followed. First, we present the different embedding models and the details of their implementation. Next, we define how to choose the

size of the high-level QM region, an important step in the division of the cluster regions. We then determine the effect of different point charge embedding schemes and assess their overall performance. In our conclusions, we suggest a protocol for researchers studying excited states in molecular crystals. The presented methodologies are implemented in a new open-source platform: `fromage` (FRameworkOrk for Molecular AGgregate Excitations).

4.2 Embedding Schemes

4.2.1 Point Charge Embedding

We consider two electrostatic Ewald embedding approaches to investigate excited states in molecular crystals: PCE and a two-level ONIOM(QM:QM') model. For the PCE approach, where only the Coulombic interactions are considered, we adopt a strategy similar to that proposed by Wilbraham *et al.*¹⁷. The atomic charges were obtained using the `Ewald` program from Derenzo *et al.*¹¹⁴ after being modified to allow non-integer charge values. The origin of the charges is discussed in Section 4.2.3. The effect of the polarisation of the environment was considered for both methods within a self-consistent embedding algorithm. These approaches were implemented in `fromage`, the source code and the documentation are available online.^{137,138}

In the `Ewald` program,¹¹⁴ an array of about 10^4 charges is generated from a supercell. Three zones are defined, the central region (zone **I**) is where the highest level of theory will be used. It is spherically surrounded by a buffer region (zone **II**) of approximately 500 point charges. Charges of both zone **I** and **II** are held constant. The rest of the charges (zone **III**) are altered to reproduce the Ewald potential in the central and buffer regions upon the direct summation of all point charges. The algorithm removes any artificial dipole moment introduced in the procedure. A detailed description of the method and the corresponding program can be found in Reference 114.

The implementation of PCE in `fromage` consists of electronic structure calculations at zone **I** atomic sites, embedded in the atomic charges of zones **II** and **III**. For clarity, we refer to zone **II** and **III** charges as Ewald charges. Excited state energies are obtained with TDDFT, CASSCF, CASPT2 and CC2 *via* interfaces with

Gaussian,¹³⁹ Molcas,¹⁴⁰ and Turbomole¹⁴¹. An interface with DFTB+¹⁴² is under development. The atomic charges can be obtained from molecular or periodic crystal calculations. We consider RESP, Mulliken and NBO from molecular calculations and RESP, Mulliken, Hirshfeld and AIM charges from periodic calculations. Currently, atomic charges can be read from Gaussian and CP2K.

fromage provides tools for the exploration of PESs of molecular crystals. The LBFGS minimisation algorithm is used to locate stationary points. A complete characterisation of excited state potential energy surfaces in molecular crystals require the description of conical intersections. We have implemented the penalty function method of Levine *et al.*¹⁴³ to optimise Minimal Energy Conical Intersection (MECI) geometries. In contrast with other methods,¹⁴⁴ this approach does not require nonadiabatic coupling vectors. A function of the averaged S_1 and S_0 energies (\bar{E}_{1-0}) and the S_1 - S_0 energy gap (ΔE) is minimised:

$$F = \bar{E}_{1-0} + \sigma \frac{\Delta E^2}{|\Delta E| + \alpha} \quad (4.1)$$

where σ is a Lagrangian multiplier and α is a parameter such that $\alpha \ll |\Delta E|$. The purpose of this functional form is to find a smooth minimum for the S_1 - S_0 energy difference, with a weight σ , whilst minimising the overall energy with a weight $1/\sigma$.

This algorithm is implemented in `fromage` for CASSCF, CC2 and TDDFT electronic methods. We would like to emphasise that even when multireference quantum methods are preferable for modelling S_1 - S_0 crossings,^{77,145} in many cases single-reference methods can provide a qualitative description of these regions of the PES. Conical intersections can approximatively be described with single-reference methods such as TDDFT.¹⁴⁵ Nonadiabatic dynamics simulations with these methods have shown for multiple systems that methods such as ADC(2) and CC2 can provide reasonable results.^{76,146} In the case of TDDFT, a careful selection of the functional is required.^{147,148} Considering the computational cost of multireference methods and the sensitivity of their active space, it can at times be necessary to resort to single-reference methods. However, their performance near S_1 - S_0 crossings should be carefully tested by comparison with multireference calculations.

4.2.2 Ewald Embedded ONIOM QM:QM'

Geometry optimisation and conical intersection search become problematic within the PCE scheme because of the lack of short-range non-Coulombic interactions which results in overpolarisation effects (see section 4.4.2). To overcome these limitations, we formulate an ONIOM¹⁴⁹ Ewald Embedded Cluster (OEEC) model. It is devised as an extension of the commonly used ONIOM Embedded Cluster model (OEC) which usually only includes electrostatic embedding up to the range of the cluster. We consider a QM:QM' scheme rather than QM:MM to avoid the need for specific parameterisation.

We wish to expand upon the electrostatically embedded expression for the ONIOM QM:QM' energy discussed in Chapter 2 by including long-range electrostatic effects. To do so, we begin by deriving the ONIOM equation. To define a total energy expression, we state the energy of the whole crystal divided into regions **1**, **2** and **3**, where **3** surrounds **2** which surrounds **1**. Regions **2** and **3** will be fixed in place whereas region **1** is allowed to reorganise. We denote inter-region interactions as $E^{\text{int}}(i \cup j)$ which can be split between electrostatic ($E^{\text{ES}}(i \cup j)$) and other ($E^{\text{X}}(i \cup j)$) interactions.

$$\begin{aligned}
 E(\mathbf{1} \cup \mathbf{2} \cup \mathbf{3}) &= E(\mathbf{1}) + E(\mathbf{2}) + E(\mathbf{3}) + E^{\text{int}}(\mathbf{1} \cup \mathbf{2}) + E^{\text{int}}(\mathbf{1} \cup \mathbf{3}) + E^{\text{int}}(\mathbf{2} \cup \mathbf{3}) + E^{\text{int}}(\mathbf{1} \cup \mathbf{2} \cup \mathbf{3}) \\
 &= E(\mathbf{1}) + E(\mathbf{2}) + E^{\text{int}}(\mathbf{1} \cup \mathbf{2}) + E^{\text{int}}(\mathbf{1} \cup \mathbf{3}) + E^{\text{int}}(\mathbf{1} \cup \mathbf{2} \cup \mathbf{3}) + \text{const.} \\
 &= E(\mathbf{1}) + E(\mathbf{2}) + E^{\text{ES}}(\mathbf{1} \cup \mathbf{2}) + E^{\text{X}}(\mathbf{1} \cup \mathbf{2}) + E^{\text{ES}}(\mathbf{1} \cup \mathbf{3}) + E^{\text{X}}(\mathbf{1} \cup \mathbf{3}) + E^{\text{int}}(\mathbf{1} \cup \mathbf{2} \cup \mathbf{3}) \\
 &\quad + \text{const.}
 \end{aligned}
 \tag{4.2}$$

We can neglect the $E^{\text{X}}(\mathbf{1} \cup \mathbf{3})$ term since Coulombic interactions dominate the long range regime. The nonadditive three-body term $E^{\text{int}}(\mathbf{1} \cup \mathbf{2} \cup \mathbf{3})$ is also neglected in this picture, as the freezing of **2** and **3** should drive it down significantly, and we wish to avoid modelling infinitely many molecules. By further removing the constant, and assigning different levels of theory to the various terms, we obtain an expression for the hybrid QM:QM' energy:

$$\tilde{E}_{\text{QM:QM}'}^{\text{OEEC}}(\mathbf{1} \cup \mathbf{2}) = \left[E_{\text{QM}}(\mathbf{1}) + E_{\text{QM}}^{\text{ES}}(\mathbf{1} \cup \mathbf{2}) + E_{\text{QM}}^{\text{ES}}(\mathbf{1} \cup \mathbf{3}) \right] + \left[E_{\text{QM}'}(\mathbf{2}) + E_{\text{QM}'}^{\text{X}}(\mathbf{1} \cup \mathbf{2}) \right] \quad (4.3)$$

Here, the first bracket can be approximated via a QM level calculation of the region $\mathbf{1}$ energy, electrostatic embedded with whichever point charges will best represent the electrostatic interactions between $\mathbf{1}$ and $\mathbf{2} \cup \mathbf{3}$. We use the notation:

$$E_{\text{QM}}(\mathbf{1}) + E_{\text{QM}}^{\text{ES}}(\mathbf{1} \cup \mathbf{2}) + E_{\text{QM}}^{\text{ES}}(\mathbf{1} \cup \mathbf{3}) \approx E_{\text{QM}}^{\text{Ew}}(\mathbf{1}) \quad (4.4)$$

For the second bracket, we can use a similar strategy:

$$\begin{aligned} E_{\text{QM}'}(\mathbf{2}) + E_{\text{QM}'}^{\text{X}}(\mathbf{1} \cup \mathbf{2}) &= E_{\text{QM}'}(\mathbf{1} \cup \mathbf{2}) - E_{\text{QM}'}(\mathbf{1}) - E_{\text{QM}'}^{\text{ES}}(\mathbf{1} \cup \mathbf{2}) \\ &\approx E_{\text{QM}'}(\mathbf{1} \cup \mathbf{2}) - E_{\text{QM}'}^{\text{EE}}(\mathbf{1}) \end{aligned} \quad (4.5)$$

Where $E_{\text{QM}'}^{\text{EE}}(\mathbf{1})$ denotes an energy calculation of region $\mathbf{1}$ embedded in QM' charges from region $\mathbf{2}$. We embed in QM' charges since the role of the embedding is to reflect the electrostatic interactions between regions $\mathbf{1}$ and $\mathbf{2}$ at the QM' level of theory.

By inserting approximations 4.4 and 4.5 in to equation 4.3, we obtain a total energy expression:

$$E_{\text{QM:QM}'}^{\text{OEEC}}(\mathbf{1} \cup \mathbf{2}) = E_{\text{QM}}^{\text{Ew}}(\mathbf{1}) + E_{\text{QM}'}(\mathbf{1} \cup \mathbf{2}) - E_{\text{QM}'}^{\text{EE}}(\mathbf{1}) \quad (4.6)$$

The hybrid gradients are defined accordingly.

To relate this equation to molecular clusters, a graphical representation of our OEEC model is shown in Figure 4.2. The OEEC model is comprised of two regions, the central region $\mathbf{1}$ (corresponding to zone I in the Ewald program) and nearest-neighbour molecules ($\mathbf{2}$). Region $\mathbf{2}$ should be large enough to include the most important short-range non-electrostatic interactions with the QM cluster. The buffer region defined for Ewald (zone II) does not necessarily correspond to region $\mathbf{2}$.

In the OEEC scheme, Coulombic interactions of any distance between region $\mathbf{1}$ and the crystal are described at the higher level of theory (QM). For excited state

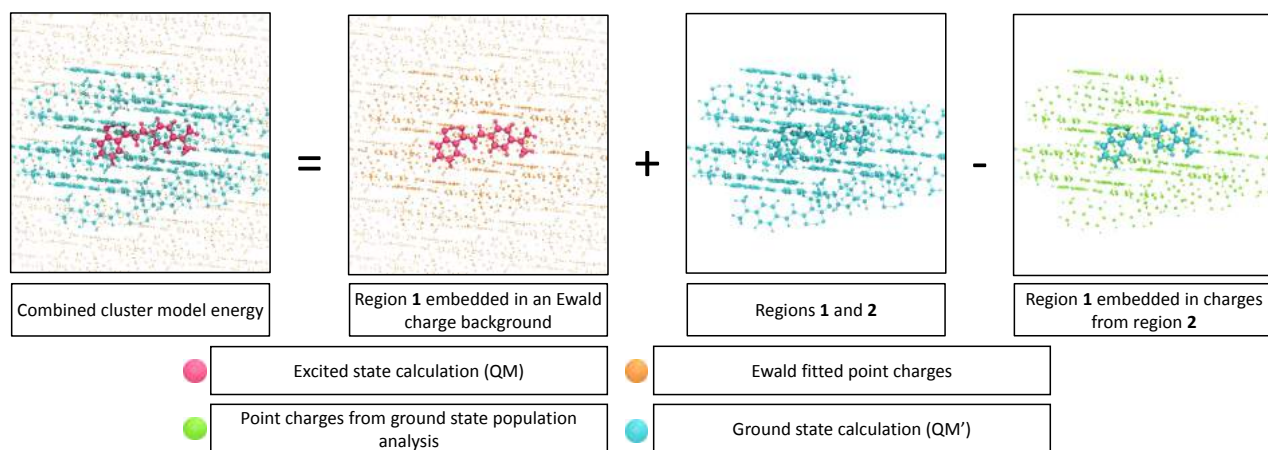


FIGURE 4.2: Visual representation of the main energy equation for the Ewald Embedded Cluster model.

calculations, this represents the interaction between an excited central region and the environment in the ground state, unless particular charges are considered in the Ewald algorithm (*vide infra*). In contrast, the short-range non-Coulombic interactions between **1** and **2** are considered at the QM' ground state level, which recovers some of the short-range contributions and improves the description provided by PCE. Since non-Coulombic interactions are considered in the ground state, for fixed geometries the energy gaps are equivalent to those obtained with the PCE. The selection of the QM' level of theory depends on the available computational resources. Previous studies on truncated cluster models have shown that low levels of theory such as HF/STO-3G achieve accurate results.^{11,14–17}

In order to consider the response of the environment to the excitation and recover mutual polarisation effects, we employ the extension of self-consistent Ewald embedding to excited states proposed by Wilbraham *et al.*¹⁷. Mutually polarising embedding methods have been applied to a number of ground state systems.^{63,116,150,151} In the self-consistent approach, a QM-level calculation is carried out on a quantum cluster. A population analysis is then applied and the charge values are re-assigned to the equivalent positions in the crystal. Those charges are then fitted using Ewald and another QM calculation is carried out. The loop between Ewald fitting and population analysis is repeated until convergence of the atomic charges. The new charge background is used for the electrostatic embedding of **1**. In fromage, the self-consistent approach is implemented for the PCE and the QM/QM' approaches

TABLE 4.1: Embedding models used in this study

Acronym	Full name	Description
PCE	Point Charge Embedding	Point charge embedding fitted to match the Ewald potential
SC-PCE-S ₁	Self-Consistent Point Charge Embedding	PCE computed self consistently in S ₁
SC-PCE-S ₀	Self-Consistent Point Charge Embedding	PCE computed self consistently in S ₀
OEC	ONIOM Embedded Cluster	QM:QM' ONIOM cluster model with the QM region embedded in charges from the QM' region
OEEC	ONIOM Ewald Embedded Cluster	OEC with the QM region embedded in charges from PCE
SC-OEEC-S ₁	Self-Consistent ONIOM Ewald Embedded Cluster S ₁	OEC with the QM region embedded in charges from SC-PCE-S ₁
SC-OEEC-S ₀	Self-Consistent ONIOM Ewald Embedded Cluster S ₀	OEC with the QM region embedded in charges from SC-PCE-S ₀

(SC-PCE and SC-OEEC). We consider two versions which may represent different physical situations in the crystal (discussed in section 4.4.1). The first, SC-PCE-S₁, closely corresponds to the embedding proposed by Wilbraham *et al.*; it uses excited state charges as an initial charge background and iterates with excited state population analyses. The second, SC-PCE-S₀, has a ground state initial charge background and performs ground state population analyses. We extend these methods to SC-OEEC-S₁ and SC-OEEC-S₀. For ease of reading, the embedding models are listed in Table 4.1. Note that it is impossible to provide a self-consistent charge environment where the central molecule is in the excited state and the surrounding ones are in the ground state, if we are to preserve the periodicity of the Ewald method.

For SC-PCE-S₁, the convergence of the charge values can be sped up by starting the loop from a ground state population analysis embedded in ground state Ewald charges. The final background was found to be very similar, with an RMSD of 10^{-5} e⁻ for atomic charges. Another alternative is to perform the loop on a molecule which has already been optimised in the excited state using OEEC. In this case, the equilibration of the charge background is made to match the excited state minimum,

however this implies assigning charges from an excited state minimum configuration to region 2 molecules which are in their ground state minimum geometry.

Figure 4.3 describes the structure of `fromage`. The charge background can be chosen to be computed self-consistently and the geometry optimisation can be set to search for ground and excited state minima or MECI. Currently, region 2 is fixed in place during geometry optimisation, although full cluster relaxation¹⁴⁴ is under development. For SC-OEEC, to recover point charges of the highest quality, the molecule of interest in the unit cell is first relaxed with OEEC. Furthermore the self-consistent charge background is computed only for the first step, at the ground state OEEC geometry, in order to maintain a consistent PES throughout the relaxation.

4.2.3 Origin of Point charges

The usual electrostatically embedded QM:QM' ONIOM model has the following energy expression:

$$E_{\text{QM:QM}'}^{\text{EE}}(\mathbf{1} \cup \mathbf{2}) = E_{\text{QM}}^{\text{EE}}(\mathbf{1}) + E_{\text{QM}'}(\mathbf{1} \cup \mathbf{2}) - E_{\text{QM}'}^{\text{EE}}(\mathbf{1}) \quad (4.7)$$

Traditional implementations use charges originating from QM' calculations in the embedding of the QM calculation in order to mitigate the overpolarisation arising when the QM wavefunction approaches the singular potential well of the point charge.⁹² The objective is to cancel out errors arising from the point charge approximation in the short range in the first and third terms of equation 4.7. However this cancellation becomes less exact when the region 1 QM charge density differs significantly from the QM' charge density as is the case when considering excited states. Furthermore, usual ONIOM schemes are designed to account for inter-region boundaries crossing a bond (and using link atoms to correct the dangling bond). In the cases discussed herein, the region boundary is defined inter- rather than intramolecularly and most intermolecular contacts are larger than 4 Å, making the usual choice of point charge less pertinent.

Moreover, in traditional ONIOM QM:QM', assuming a well behaved cancellation of point charge overpolarisation, the electrostatic potential of region 2 on 1 stems entirely from the $E_{\text{QM}'}(\mathbf{1} \cup \mathbf{2})$ term, thus causing the choice of point charges in the

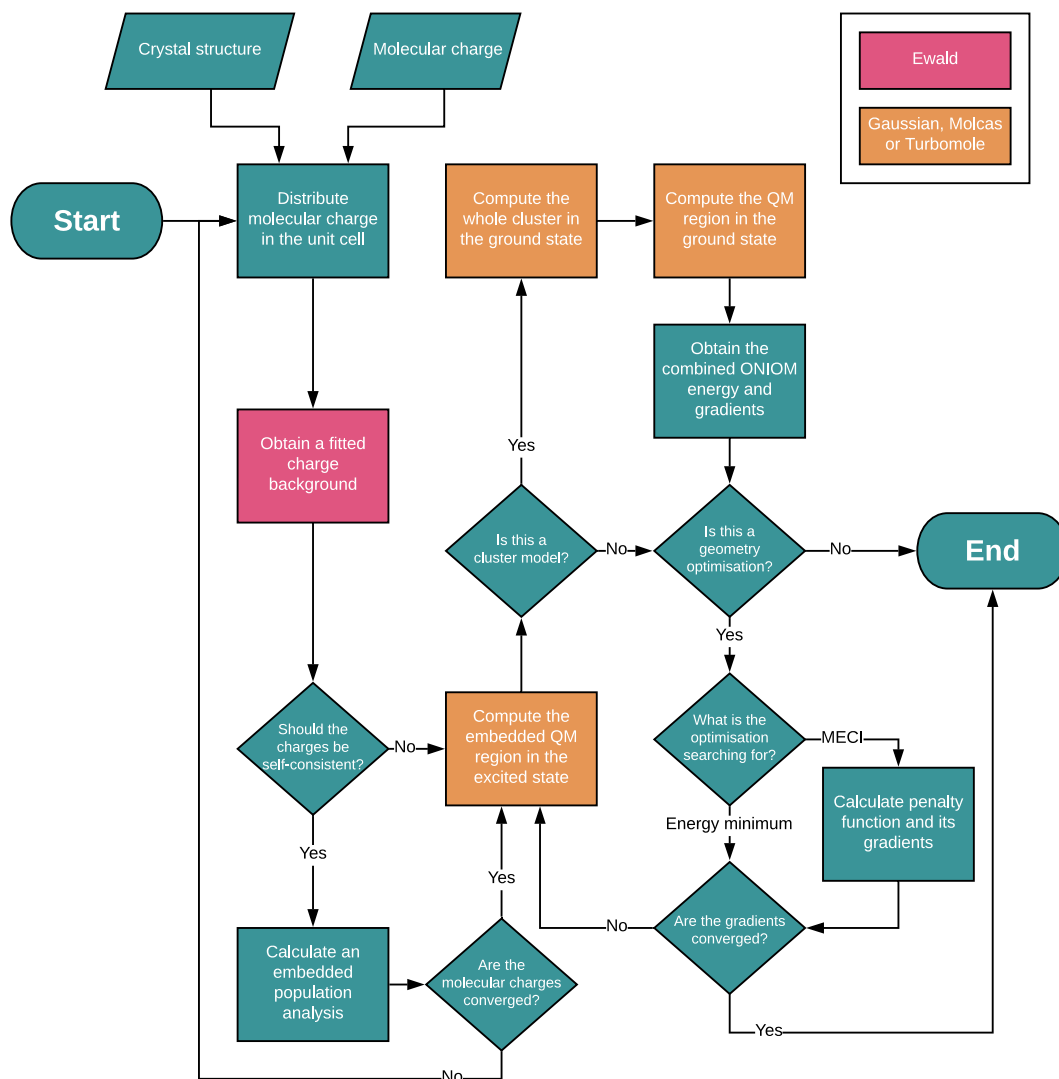


FIGURE 4.3: Flowchart of a calculation using the ONIOM Ewald Embedded Cluster (OEEC) and Self-Consistent ONIOM Ewald Embedded Cluster (SC-OEEC) models. The electronic program can be chosen by the user.

embeddings of $E_{\text{QM}}^{\text{EE}}(\mathbf{1})$ and $E_{\text{QM}'}^{\text{EE}}(\mathbf{1})$ to be of little overall impact as long as they match. In contrast, equation 4.6 aims to cancel out the potential between the second and third term, thus relying entirely on the embedding of the first term for a high quality electrostatic potential of regions **2** and **3** on **1**.

Therefore in the implementation discussed herein, we principally use ground state QM charges in the embedding of the first term and ground state QM' charges for the final term. However, all other embedding combinations are available in fromage. Further extensions of these methods can be implemented to reduce artificial polarisation⁸⁸ and make the methods useful for more dense systems.

The alternative scheme where Ewald charges are used for the QM' calculations should provide a marginally worse compensation of the inter-region Coulombic interactions. Nevertheless, the results obtained with this embedding scheme are similar to those obtained with the cluster charges for **HC1**. The absorption energies are only deviated by 0.01 eV from those obtained with the original scheme. In the case of the emission energies from the **K** form, the value obtained with this version of OEEC is 2.24 eV which is in relative good agreement with the results obtained with other schemes (Table 4.3). Indeed by extending the QM' point charges to into region **3** the only non-constant term which is artificially added to the energy is the QM' electrostatic interaction between regions **1** and **3**, which should be small compared to the overall energy differences in the PES, and does not affect excitation energies for a given geometry.

4.3 Computational Details

The crystal structures of **HC1** and **HC2** were optimised using PBE-D2 as implemented in Quantum Espresso.¹⁵² The plane wave cutoff was 30 Ry and the k-point meshes were 2x3x2 and 2x2x1 respectively, in accordance with the shapes of the unit cells. Subsequently, a single point PBE-D2/DZVP calculation was carried out using CP2K¹⁵³ to extract RESP, Hirshfeld and Mulliken periodic charges. For AIM charges, an external program developed by Henkelman *et al.* was used to process the Quantum Espresso DFT charge density¹⁵⁴⁻¹⁵⁷. Molecular RESP charges were first calculated at HF/3-21G(p) level for comparison with our previous ONIOM (QM:AMBER)

calculations.¹³¹ Every other molecular population analysis (NBO, Mulliken, RESP for OEEC and SC-OEEC models) used ω B97X-D/6-311++G(d,p) as implemented in Gaussian.

For the seven charge schemes, 1000 validation points were randomly sampled in the quantum cluster (in order to measure the accuracy of the fit to the Ewald potential) and 500 points had their value fixed to create a buffer region. The total charge background was comprised of 64 unit cells for **HC1** and 32 for **HC2**. These numbers were chosen so as to create a sufficient amount of point charges¹¹⁴—at least 10000—while keeping an isotropic distribution in accordance with the shape and size of each unit cell.

Both molecular crystals were then investigated using a hierarchy of models. First, PCE was used with all of the charge types described above on a single QM-level monomer. When possible, the excited state geometries were optimised with TD- ω B97X-D/6-311++G(d,p). Next, the cluster models were introduced, using RESP charges from ω B97X-D/6-311++G(d,p) in the embedding of $E_{\text{Ew}}^{\text{QM}}(\mathbf{1})$. OEC, OEEC and SC-OEEC were all employed on a single monomer of the crystal embedded in a cluster of 21 molecules for **HC1** and 16 molecules for **HC2**. The excited state minima and S_1 - S_0 MECI were found using *fromage*. For the location of S_1 - S_0 MECI, the parameters in Equation 4.1 were initially set to 0.02 Hartree for α , to provide a smooth minimum of the S_1 - S_0 energy difference and 3.5 for σ . σ was then increased if the gap was found to be insufficiently small after optimisation of F .

For the comparison of different points along the potential energy surface, we use the same charge background throughout. This avoids varying classical energy contributions due to charge-charge interactions and different Ewald constants.²¹ In this article, we used the charge background obtained for the **FC** conformation, although for crystals with significant Frenkel exciton occurrences, S_1 self-consistent charges could provide a better description of the excited states. All backgrounds are available in *fromage*, leaving the choice up to the user.

To probe for excitonic effects, the optimised monomer geometries were inserted into tetramers at the lattice positions, and their excited states were calculated using TD- ω B97X-D/6-311++G(d,p) and OEEC.

Overall, the QM methods employed were TD- ω B97X-D/6-311++G(d,p) using

Gaussian, RI-CC2/TZVP and RI-CC2/SV(P) using Turbomole and SA-2-CASSCF(12,11)/6-31G(d) and MS-2-CASPT2(12,11)/6-31G(d) using Molcas; all with PCE, OEEC and SC-OEEC. The QM' method was HF/STO-3G using Gaussian and the low level embedding charges of $E_{QM'}^{EE}(\mathbf{1})$ were accordingly chosen to be from RESP calculations at the same level of theory. For the self-consistent population analysis procedure, a convergence criterion of $0.001 e^-$ for the mean deviation of charge values between subsequent steps was chosen.

In certain cases, convergence issues occurred in the self-consistent loop such as divergence or oscillation. To address this, under-relaxation was employed with a damping factor of 0.75, where each new charge value adopts a weighted average value of $q_{n+1}^{\text{damped}} = 0.25q_{n+1} + 0.75q_n^{\text{damped}}$ where q_n is the charge value at the previous step, and q_{n+1} the new charge value obtained in the population analysis.

For excited state self consistent backgrounds, using initial charges from an isolated excited state molecule or an Ewald embedded ground state molecule yielded the same final background although the latter method converged in fewer steps.

For comparison, single monomers were also optimised in the ground and excited states using TD- ω B97X-D/6-311++G(d,p) in vacuum and using Polarizable Continuum Models (PCM) and Self-Consistent PCM (SC-PCM) with a dichloromethane (DCM) solvent as implemented in Gaussian. Exciton couplings were computed using the diabatisation scheme proposed by Arag3 and Troisi, which considers short and long-range contributions, further discussed in Chapter 5.¹⁰⁵

4.4 Results

4.4.1 Localisation of the Excitation: Size of the QM Region

The use of embedding techniques for excited states calculations in molecular crystals presumes the localisation of the excitation over a few molecular units. However, the degree of localisation is often unclear and unpredictable, conflicting with the intrinsic truncation of a cluster model. Therein lies the necessity for different kinds of embedding techniques which represent different physical situations. Before comparing the effect of these techniques, we wish to clarify how they relate to exciton localisation in our model systems.

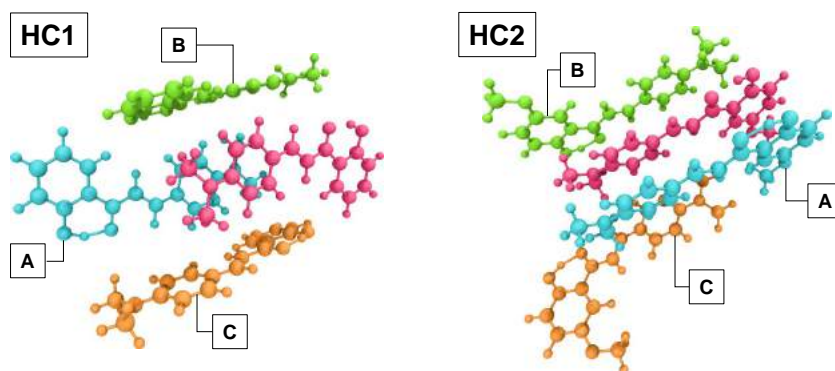


FIGURE 4.4: Selected tetramer configurations from both crystals. The molecule in pink is optimised using OEEC.

In the case of OEEC, the Ewald charges arise from a ground state population analysis. Consequently, this approach represents a localised excitation in region 1 before the environment has responded to the change in electronic density. To instead represent the extreme situation where all molecules are excited simultaneously and are mutually responsive, charges from excited state calculations can be used. We call this scheme SC-OEEC- S_1 (or, in general, SC-OEEC- S_n). If the molecules in the QM' region are considered to be in the ground state and the S_0 charges are self-consistently updated and alternative SC-OEEC- S_0 scheme can be defined. It is expected that the SC-OEEC- S_1 scheme will perform better in systems where excitation is highly delocalised and the S_1 electron density is significantly different from the ground state. We have implemented all these schemes in fromage so that the user can select the most suitable scheme for the system under investigation. The degree of localisation of the excitation in a molecular crystal will depend on the exciton coupling with neighbouring molecules and the experimental conditions for absorption.

In order to investigate the excitonic features of the excited state electron densities of the **HC1** and **HC2** crystals, we consider a tetramer (Figure 4.4) embedded in ground state Ewald charges as a reference. This model includes the short-range Coulomb interactions between the central and three surrounding molecules explicitly and thus should provide a benchmark to evaluate the ability of the different embedding schemes to describe the excited states considering a smaller QM region. Note that in contrast with the monomer, where the bright state is S_1 , for the tetramer the bright states are S_4 and S_5 for **HC1** and **HC2** respectively.

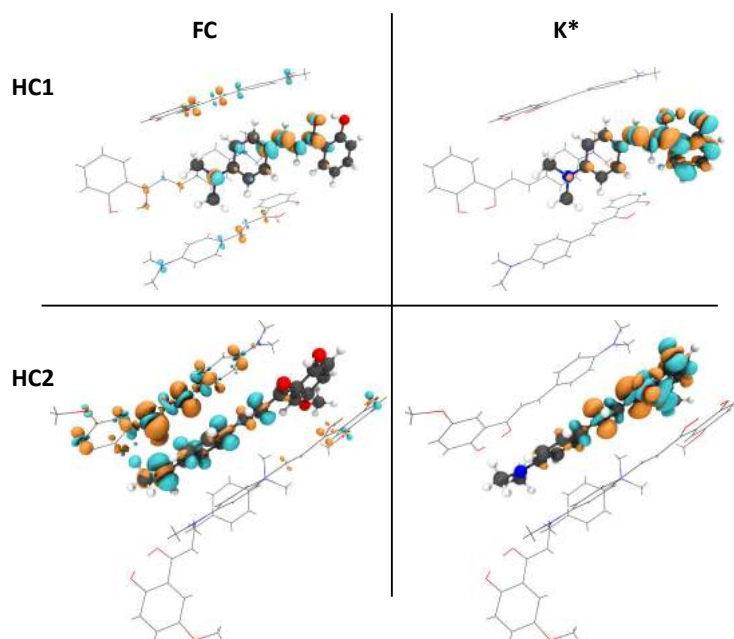


FIGURE 4.5: S_n-S_0 density differences obtained at TD- ω B97X-D/6-311++G(d,p) level of theory for the tetramer model (excited state density gain upon excitation is shown in orange, and the loss in blue). For the FC geometries, the tetramer's bright states were considered ($n=4$ for **HC1** and $n=5$ for **HC2**); for **K***, $n=1$. All configurations were obtained by optimising the geometry of the central molecule with OEEC.

Figure 4.5 shows the S_n-S_0 density differences obtained for the bright state at Franck-Condon geometry (FC) and the \mathbf{K}^* S_1 excited state minimum geometries. The plots for the first five excited states can be found in Appendix A.4. When considering a full tetramer in the excited state, the bright states of **HC1** and **HC2** are S_4 and S_5 respectively, whereas for single monomers, they are both S_1 . An important degree of localisation is observed on monomers and dimers for both crystals, despite four molecules being included in the QM region. Consequently, we expect that with embedding charges of sufficient quality, a QM region of one or two molecules would obtain accurate excited state energies.

Excitations in **HC1** are more localised than in **HC2**, which correlates with the larger exciton couplings obtained for the latter.¹³² In the case of **HC1**, only the coupling with molecule B is larger than 0.1 eV (Appendix A.4). For both crystals, in the \mathbf{K}^* minimum, the excitation is clearly localised in the central molecule, which suggests that schemes such as OEEC and SC-OEEC- S_0 could be best suited to describe this kind of situations (see discussion in the next sections). Additionally, the QM region with only one monomer should be able to describe emission from the \mathbf{K}^* form, which is confirmed by the evolution of the energies with the size of the region (see section 4.4.3).

4.4.2 Point Charge Embedding: Electrostatic Effects in the Crystal

Having now observed the localisation of the excitation upon fluorescence, we can begin assessing the performance of different embedding methods with only one monomer in the QM region. In this section, we analyse the performance of the PCE model and the effect of using different charges for the description of excited states in the **HC1** crystal. Our analysis is based on the results obtained with a monomer in the QM region.

The experimental absorption in the solid state shows two bands which have previously been attributed to absorption from the **E** and **K** forms.^{134,158} Our calculations show that neither the crystal composed of **K** molecules nor the one with **K** surrounded by **E** molecules are stable in the solid state. Additionally, the experimental crystal structure does not seem to be consistent with a significant population of the **K** form in the ground state.¹⁵⁸ Taking this into account, the presence of **K** in

TABLE 4.2: Absorption, emission and **K-MECI** energies (in eV) of **HC1** embedded in different types of Ewald point charge arrays. Unless specified the geometries were obtained at the ONIOM(TD- ω B97X-D/6-311++G(d,p):AMBER) level of theory. **K-MECI** energies are relative to the ground state energy of the Franck-Condon (FC) minimum.¹³² †Optimised geometries within the PCE environment. ††Geometry optimised in vacuum

Method	Charges		Absorption		Emission		S ₁ -S ₀
	Type	Basis	FC(E)	E*	K*	K-MECI	
Molecular							
TD- ω B97X-D/ 6-311++G(d,p)	NBO		3.28	3.10	2.67	4.35	
	RESP		3.30	3.12	2.66	4.41	
	RESP (SC-PCE-S ₁)	6-311++G(d,p)	3.09	2.96	2.65	4.72	
	RESP [†]		3.37	-	2.69	4.37	
	Mulliken		1.56	1.51	1.47	4.42	
	Mulliken	3-21G(d)	3.29	3.11	2.70	4.76	
	Mulliken	6-31G(d)	3.35	3.16	2.70	4.68	
RI-CC2/SV(P)	RESP	6-311++G(d,p)	3.11	2.95	2.35	3.82	
RI-CC2/TZVP	RESP		2.98	2.82	2.29	3.56	
Crystal							
TD- ω B97X-D/ 6-311++G(d,p)	RESP		3.33	3.15	2.68	4.32	
	AIM		3.35	3.16	2.68	4.30	
	Hirshfeld	DZVP	3.43	3.23	2.68	4.56	
	Hirshfeld [†]		3.50	3.20	2.28	2.89	
	Mulliken		2.95	2.85	2.64	4.20	
No charges							
TD- ω B97X-D/ 6-311++G(d,p)	-	-	3.52	3.31	2.67	4.76	
	- ^{††}	-	3.65	3.28	0.36	2.84	
-	Experimental ^{134,158}	-	2.9, 3.3	-	1.7–1.9	-	

the ground state seems to be associated with dynamic processes activated in the experimental conditions. For example, at room temperature, large amplitude motions of the proton along the H-bonded bridge can reduce the S_1 - S_0 energy gap to 2.76 eV, considering vibrational broadening. Additionally, given the ultrafast nature of the proton transfer in the solid state (3 ps¹⁵⁸), fast absorption from **K** forms generated in the excited state could be also possible. The dynamic nature of these processes is in line with the broad structure of the low energy band. Our focus is the analysis of the higher energy band which corresponds to the absorption in the **E** form.

To estimate the effect of vibrational broadening on the position of the absorption maximum, we use the nuclear ensemble method¹⁵⁹ as implemented in Newton-X¹⁶⁰ with TD- ω B97X-D/6-311++G(d,p) embedded in RESP charges. In this method, a normal mode calculation of the molecule is carried out in the ground state. Geometries are then sampled from these normal modes using a Wigner distribution, and vertical absorptions are calculated at each sampled point. The position of the **E** absorption maximum (3.21 eV, 0.1 eV shift with respect to the vertical excitation with the same method) is in excellent agreement with the experimental value (Appendix A.2).

To directly evaluate the effect of charges of different origin, we compare absorption, emission (from **E*** and **K*** forms), and S_1 - S_0 MECI energies with Ewald embedding. Given that the MECI associated with the enol pathway was consistently found to be at least 4 eV higher in energy than its keto counterpart, we will focus on the **K-MECI** deactivation pathway. The results are summarised in Table 4.2. In order to directly compare the impact of different charge partition schemes, we use the same geometry throughout, obtained at the ONIOM(TD- ω B97X-D/6-311++G(d,p):AMBER) level of theory.¹³²

Excited state calculations with PCE using non-Mulliken charges predict the maximum of absorption with close agreement to the experimental value of 3.3 eV. Overall, the effect of the embedding is to shift absorption to the red with respect to the energy obtained in vacuum (3.65 eV). For calculations at fixed geometries, there is no significant dependence on whether the charges are obtained from molecular or crystal calculations. In particular, the energies obtained using RESP charges are consistent between the molecular and the crystal descriptions. In the context of the

molecular organic crystals, this is not surprising as RESP charges are designed to match the electrostatic potential, and crystal packing has but a small effect on the electronic structure of these molecules due to their non-bonded nature.

In contrast, calculations using Mulliken charges strongly depend on the choice of basis set (both the size and the type). With these charges, results with smaller basis sets are closer to the experimental value. They provide reasonable energies with 3-21G(p) and 6-31G(d), but fail to reproduce sensible values if a larger basis set is used (6-311++G(d,p)). This is in line with the well-known sensitivity of the Mulliken method to the basis set.

The experimental emission in the solid state has been attributed to the \mathbf{K}^* form.¹³² Regardless of the higher stability of \mathbf{K}^* in the excited state, emission from the \mathbf{E}^* form is expected to be close to the initial absorption and consequently reabsorbed. Accordingly, our TDDFT calculations predict emission from \mathbf{E}^* in the range of 3.1–3.3 eV. Interestingly, emission from the \mathbf{K}^* form (~ 2.7 eV) is significantly deviated from the experimental values (1.7–1.9 eV).^{134,158} This is not improved by using self-consistent point charges in the SC-PCE- S_1 method, which is to be expected due to the localisation of the excited state to one molecule (Figure 4.5). When the emission is calculated using RI-CC2, the energy is improved but is still deviated by more than 0.5 eV from the experiments.

The most significant factor is the geometry itself, obtained at QM:MM level. We show later that a better emission energy is obtained when optimisation is done using the OEEC and SC-OEEC methods. When it comes to the optimisation of excited state minima and S_1 – S_0 MECI, the PCE approach was unsuccessful in locating stable minima for most charges types, due to the lack of non-Coulombic short-range interactions and ensuing overpolarisation effects. Only Hirshfeld and, in certain cases, crystal RESP charges were overall small enough in magnitude to allow for the determination of local minima, with the largest charge on an atom of $0.5 e^-$. As such optimisation with PCE in general is not recommended for systems with a high degree of conformational flexibility.

As for S_1 – S_0 MECI energies, all TDDFT results for the QM:MM geometries are more than 1 eV above the FC bright state energy (Table 4.2). The conical intersections are thus rendered inaccessible as expected since $\mathbf{HC1}$ displays aggregation-induced

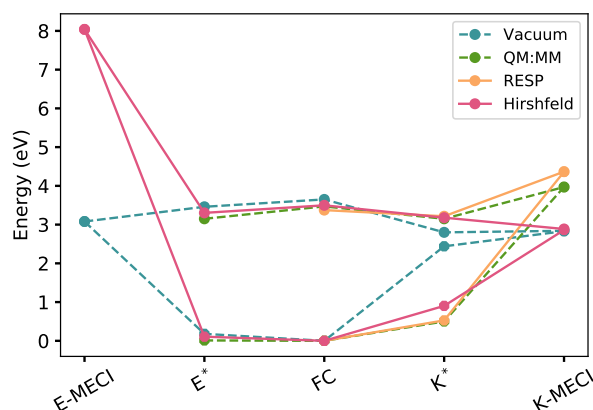


FIGURE 4.6: Energies **HC1** in vacuum and embedded in RESP and crystal Hirshfeld Ewald charges at different geometries. QM:MM energies were taken from Ref 132.

emission. The energies obtained with RI-CC2 are in the range of 0.6–0.7 eV above their corresponding excitation energies which also makes them inaccessible. These results are consistent within the RACI model but are overestimated with respect to the value of 3.97 eV obtained with QM:MM with a dimer in the QM region.¹³²

In the case of RESP charges, optimisation within the PCE model does not significantly change the energetics previously evaluated with single point calculations. Indeed the structures are close to those reported at the QM:MM level of theory. The resulting relative energies are shown in Figure 4.6. For Hirshfeld charges, the effect of optimisation is more significant reducing the **K*** emission energies to 2.28 eV and making the S_1 – S_0 MECI accessible.

Comparing the calculations in PCE and vacuum at the same geometry highlights the main effects of the crystal electrostatic environment in the excited state. The excited states are overall stabilised which reduces both the vertical excitation and conical intersection energies. However the accessibility of the latter depends on the former, netting no clear difference in emissive behaviour. A more substantial relative stabilisation of the MECI is observable when the molecule is fully optimised in vacuum. It reaches a highly distorted geometry which would be inaccessible in the solid due to short-range interactions of the closely packed neighbouring molecules.

Our simulations show some of the drawbacks of the PCE method, in particular for its use in geometry optimisation. Because of the effects of overpolarisation and the lack of short-range non-Coulombic interactions, electrostatic forces can become

too large and some nuclear configurations become unstable. Consequently, we do not recommend the use of PCE for geometry optimisations. While the method is effective in some cases,^{16,161} it is unpredictable whether it will provide reliable geometries for all regions of the PES. To mitigate these problems we implemented a two level embedded cluster model.

4.4.3 Embedded Cluster Models: Potential Energy Surfaces in the Crystal

We obtained the geometries of notable regions of the PES for **HC1** and **HC2** crystals using the OEEC and its self-consistent variant SC-OEEC methods. Table 4.3 shows the absorption and emission energies obtained after optimisation of **FC** and **K*** forms using these models.

TABLE 4.3: Table of absorption and emission energies for both model systems after geometry optimisation with cluster models. The level of theory was TD- ω B97X-D/6-311++G(d,p). Energies are in eV.[†]Charges obtained for the **K*** form in S_1

Cluster model	HC1		HC2	
	FC	K*	FC	K*
OEEC	3.37	2.06	3.41	2.07
SC-OEEC- S_1	3.08	2.60 (2.12) [†]	3.34	2.15
SC-OEEC- S_0	3.27	2.19	3.40	2.07
OEC	3.27	2.40	3.42	2.03
PCM	3.32	2.36	3.72	2.44
SC-PCM	3.00	2.66	3.01	2.21
ONIOM QM:MM (molecule) ¹³²	3.32	2.72	3.50	2.17
ONIOM QM:MM (dimer) ¹³²	3.27	2.61	3.29	2.19
Experimental ^{134,158}	2.9,3.3	1.7 - 1.9	-	1.8

Geometry optimisation using the embedded cluster method produces ground state geometries very similar geometries to QM:MM, consequently the absorption energies are not significantly altered between PCE (Table 4.2) and OEEC or SC-PCE- S_1 and SC-OEEC- S_1 provided that RESP charges are used throughout (Table 4.2). For

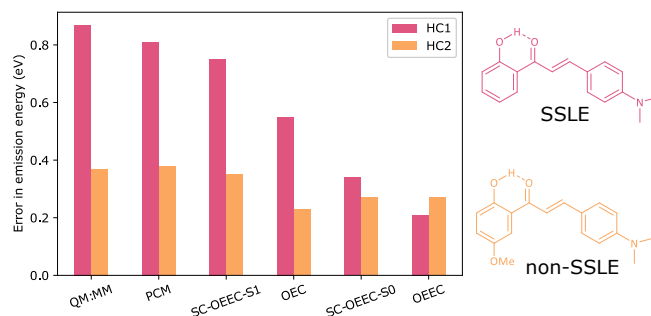


FIGURE 4.7: Deviation of the predicted emission energies of **HC1** and **HC2** with respect the experimental values by different embedding models. Reference experimental values were 1.8 eV for both **HC1** and **HC2** crystals

comparison, we have added the results obtained with ONIOM (QM:MM) including one and two molecules in the QM region and with the OEC model. Moreover, we present results in PCM and its self-consistent variant, both in DCM solvent, since continuum models are a common method employed to reflect the electrostatic environment in molecular condensed matter.¹⁶²

For all cluster models, the effect of the polarised response of the environment is to reduce the vertical excitation. This is also observed in the comparison between the PCM and SC-PCM models and when the size of the QM region increases from a molecule to dimer. The results obtained with the SC-OEEC procedure depend on the level of excitation in the self-consistent loop. If ground state charges are used, as expected, the energies are similar to those obtained with OEEC (3.27 and 3.37 eV for **HC1** and 3.40 and 3.41 for **HC2**). The emission energies obtained with different methods strongly depend on the rotation angle (Figure 4.1). In vacuum, the excited state minima show a significant deviation from their ground state planar structure. In the solid state, the \mathbf{K}^* geometries obtained with QM:MM and SC-OEEC-S₁ are more planar (**HC1**: 6° and 9°, **HC2**: 16° and 12° respectively for the angle depicted in Figure 4.1) than those obtained with OEEC (**HC1**: 32°, **HC2**: 18°).

The optimisation of the \mathbf{K}^* form with the OEEC model significantly improves the emission energies with respect to those obtained using QM:MM geometries. Figure 4.7 summarises the deviation of the calculated emission energies with respect to the experimental data.

For **HC1**, the result is 2.06 eV (TD- ω B97X-D/6-311++G(d,p)), which is in close

agreement with the experimental value. As with the PCE method, the self-consistent background based on the excited state charges at the FC state does not improve the results (2.6 eV). If ground state E charges are employed in the self-consistent loop, the emission energies remain in better agreement with the experimental values. Indeed, given the level of localisation of the excitation in these systems (Figure 4.5), the use of ground state charges seems to be more appropriate (SC-OEEC- S_0) with emission energy of 2.19 eV. An alternate version of SC-OEEC- S_1 is also employed where the self-consistent loop is carried out on a molecule in its OEEC optimised K^* geometry, which brings the emission energy to 2.12 eV. However, this charge background represents the situation where all molecules in the crystal exhibit charges from the keto form, which we have not found to be a stable minimum in the crystal.

Comparison of the OEEC with the OEC model indicates that long-range interactions account for more than 0.3 eV in the K^* emission energy of **HC1**. Interestingly, while the energies for **HC1** strongly depend on the charge background, the values for **HC2** are less affected. In the case of **HC2**, SC-OEEC, OEEC and OEC all provided very similar results (2.15, 2.07 and 2.07 eV), suggesting that the most important Coulomb effects are recovered at the short-range. This is linked to the difference between the ground and excited state charges of these molecules. For **HC1**, excitation significantly alters the charges of carbon atoms in the bridge, since it concentrates a large fraction of the molecule's π orbitals (Appendix A.3). In the case of **HC2** only the charge of the carbonyl carbon (C_k) changes more than $0.1 e^-$ upon excitation, since the electronic reorganisation is concentrated in the back ring (Figure 4.5). Consequently, the S_1 - S_0 energy gaps for **HC1** are far more dependent on the electrostatic environment. Indeed improving the description of the short-range intermolecular interactions does not significantly alter the energy gaps, as illustrated by the behaviour of the absorption and emission energies with the size of the QM region (see end of section). For both, **HC1** and **HC2**, the emission from K^* is fairly well reproduced with only one molecule in the QM region, which is in line with the localised nature of the K^* (Figure 4.5).

Conical intersections play a key role in photophenomena, providing a radiationless decay funnel for the excited state. One of the features implemented in fromage is the searching of crossing geometries using the penalty function method of Levine

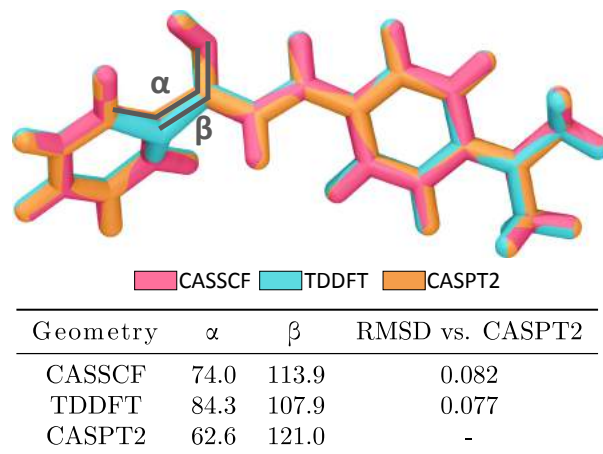


FIGURE 4.8: MECI geometries found with SA-2-CASSCF(12,11)/6-31G(d) and TD- ω B97X-D/6-311++G(d,p). Additionally, we include the configuration with least S_1 - S_0 gap when scanning the β angle from the TDDFT geometry at MS-2-CASPT2(12,11)/6-31G(d) level. It is labelled CASPT2.

*et al.*¹⁴³ The molecules considered here can deactivate to the ground state in solution via conical intersections associated with intramolecular rotation.^{131,132} We optimise the S_1 - S_0 MECI with the SA-2-CASSCF(12,11)/6-31G(d) and TD- ω B97X-D/6-311++G(d,p) methods within the OEEC scheme (Figure 4.8). For these systems, the geometries obtained with both levels of theory are in very good agreement.

In the crystal, the lowest energy conical intersection combines intramolecular rotation and a significant pyramidalisation of the carbonyl carbon.¹³² In the gas phase, the lowest energy conical intersection only involves intramolecular rotation while the one also involving pyramidalisation is higher in energy. Therefore one of the effects of the crystal environment is to modify the stability of the lowest energy conical intersections, which is consistent with the results obtained with QM:MM calculations. This confirms that the effect of short-range interaction is essential in determining the geometry while the long-range interactions modulate the total energy. However, the net effect of the embedding on the total energies is highly system dependent.

Figure 4.9 shows the PES obtained with multireference methods. The vertical excitation obtained with SA-2-CASSCF(12,11)/6-31G(d) is significantly deviated from the experimental value (4.25 eV compared to 3.3 eV). Including dynamic electron correlation with MS-2-CASPT2(12,11)/6-31G(d) shifted the value to the red in much better agreement (3.53 eV). The energy gap at the CASSCF conical intersection is too

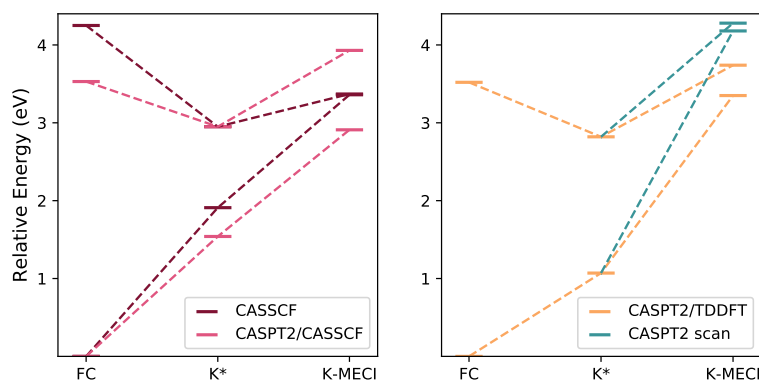


FIGURE 4.9: Relative energy diagrams showing the emission energy and accessibility of the MECI with multireference methods at the geometries shown in Figure 4.8. To minimise the CASPT2 S_1 - S_0 gap at MECI configurations, a geometry scan was carried out as described in Figure 4.10. The newly optimised geometry is labelled “CASPT2 scan” and has a gap of 0.10 eV.

large with PT2 (1.02 eV), but using TDDFT geometries as reference can significantly narrow the gap. These are common challenges found in multireference calculations and not due to the embedding approach.^{163–165} In order to further narrow the S_1 - S_0 gap, the aromatic C was systematically displaced *via* the β angle (Figure 4.8). Figure 4.10 shows how this scan locates a conical intersection at an energy 0.8 eV above the FC energy. These examples show that the Ewald embedding methods can provide all the information required to fully characterise the PES in molecular crystals. All of these methods are available in fromage.

Given that the Ewald embedding methods describe the effect of the electrostatics of the whole crystal, they represent unique schemes to analyse the convergence of properties with the size of the QM region. Exploring these effects is essential in systems with significant excitonic effects. We consider the behaviour of the energies and the accessibility of the S_1 - S_0 MECI with the size of the QM region (Figure 4.11). We employ the TD- ω B97X-D/6-31G(d) level of theory, which provides a good description of different regions of the PES. For **HC1**, the energies of the bright state for FC and of the emission from **K*** converge relatively quickly. On the other hand, the energy of the S_1 - S_0 MECI increases with the size of the QM region for **HC1** and decreases for **HC2** respectively becoming less and more accessible. This is in line with

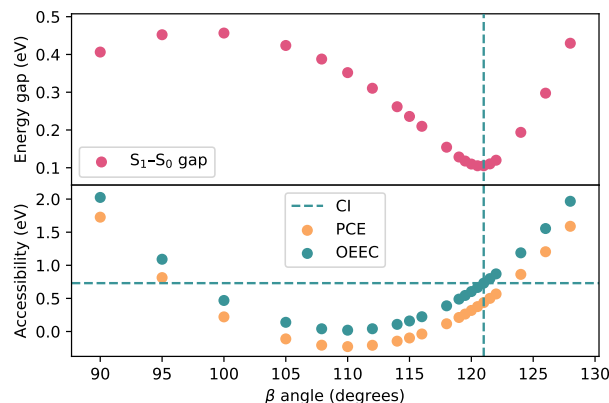


FIGURE 4.10: Plot of the S_1-S_0 gap, and the MECI accessibility as a function of the puckering angle β . The accessibility is defined as the average S_1-S_0 energy of the given geometry minus the FC bright state energy. The OEEC energy at the geometry with smallest S_1-S_0 gap is indicated by dashed lines.

the experimental behaviour of both crystals.

4.5 Conclusion

In this chapter, we analysed the behaviour of different Ewald embedding schemes for the description of excited states in molecular crystals. With focus on the exploration of potential energy surfaces, we have implemented these methods in the Python open-source platform `fromage`, which we make readily available. This program enables users to easily combine electronic structure codes of their choice for geometry optimisation using OEC, OEEC and SC-OEEC. The current implementation includes interfaces to popular quantum chemistry programs such as DFTB+, Turbomole, Gaussian and Molcas. Additional interfaces can be easily implemented, provided that the new codes allow for gradient calculations with point charge embedding.

We have shown that the PCE method is poorly suited to optimising the geometry of flexible molecules in the crystal form. Consequently, the photochemical conclusions that arise from PCE calculations of such molecules have the potential to be quantitatively and qualitatively erroneous. To overcome this problem, a series of two-level ONIOM(QM:QM') cluster models with Ewald embedding were formulated. They are suitable for geometry optimisation of excited state minima and

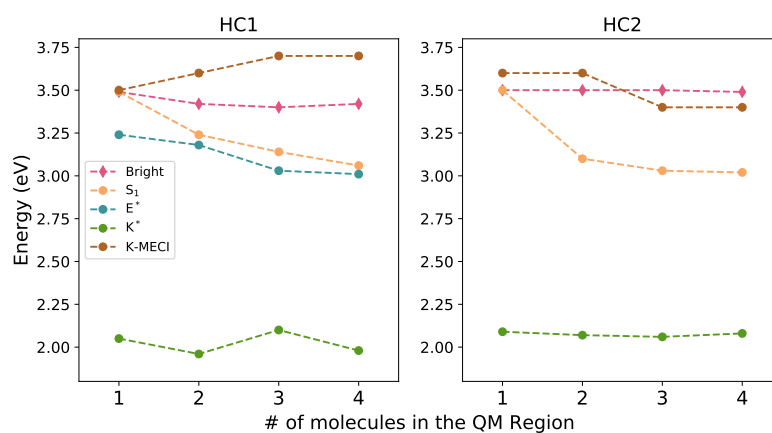


FIGURE 4.11: Energy of the absorption, emission and conical intersection as the excited state region increases in size using the OEEC model. The region molecules are added to the region in the order displayed in Figure 4.4. The energies are evaluated with TD- ω B97X-D/6-31G(d).

conical intersections, and include long range electrostatic system-environment interactions within the crystal, thus addressing the fault in traditional cluster models outlined in Section 2.8.1.

The potential of these tools was illustrated by applying them to the excited states of two model crystals, **HC1** and **HC2**, displaying excited state intramolecular proton transfer. **HC1** displays aggregation induced emission whilst **HC2** shows no emission in solution or solid state. For both systems, the excitations are clearly localised in one or two molecules which allowed the emission energies to converge with only a monomer or a dimer in the QM region. We found that using charges originating from molecular or crystal calculations did not significantly impact the results. The emission energy was progressively improved with a hierarchy of embedding models, ranging from a deviation from the experimental emission peak of 0.8 eV with ONIOM QM:MM to 0.2 eV with OEEC. Due to the flexibility of the molecules, this increase in accuracy could only be achieved by carrying out geometry optimisation with each embedding model.

Self-consistent procedures help to model the mutual polarisation between the excited state region and the environment, a known problem of ONIOM models presented in Section 2.8.3. In particular, if the self-consistent loop is carried out in the excited state, these procedures may help reflect the delocalisation of an excitation

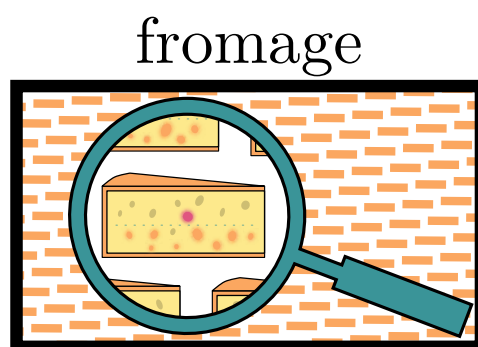
despite explicitly modelling fewer excited fragments than are involved in the delocalisation. For the systems considered in this study, the degree of localisation in both the absorption and emission processes made excited state self-consistent embedding unsuitable. The application of the self-consistent procedure to the ground state did not significantly alter the results obtained from the corresponding non-self-consistent procedure which suggests that the electronic structure of the isolated ground state molecule is not particularly altered by crystal packing.

With these conclusions in mind, we can suggest optimal embedding methods for the study of the photochemistry of different molecular crystals. If the molecule is certain to be structurally rigid and the exciton couplings are small, the PCE scheme can be appropriate. Otherwise, cluster models are preferred since they allow for exploration of the nuclear configuration space. In either case, long range electrostatic interactions can account for a significant contribution to the excited state energy. Comparison between excitation energies obtain with truncated cluster models and single point calculations with Ewald embedding methods can help decide whether these methods are required.

The size of the QM region should be motivated by the locality of the excitation at the noteworthy points of the PES which can be estimated by the calculation of exciton couplings between molecular fragments in their lattice positions. These coupling values can, in turn, be estimated as half of the S_2 - S_1 energy gap for the dimer¹⁶⁶ or by using more sophisticated methods (see Section 5.3.2). For localised excitations, ground state background charges can be a good choice (either OEEC or SC-OEEC- S_0). If the excitation is delocalised and the QM region becomes impractically large, the SC-OEEC- S_1 method can provide a better description of the excited states. We believe the use of these methods will contribute to a better understanding of complex photochemical processes in the crystal environment, impacting a broad range of applications.

Chapter 5

Implementation



5.1 Introduction

Chapter 4 should provide enough evidence that studying the photochemistry of some molecular crystals at the finite scale is a sensible idea. In fact, the structural features which are relevant to these systems are of the scale of the aggregate, ranging from monomer conformation and dimer packing all the way to molecular cluster shape.

The methodology already existing to model finite size molecular aggregates can therefore in part be repurposed to model periodic molecular crystals when dealing with excited state processes. The commonplace approach to investigating molecular crystals—using periodic electronic structure codes—should thus be complemented by adequately extending molecular aggregate methods for the excited state.

The generation and analysis of these aggregate nuclear configurations is often separated from the codes which produce their corresponding electronic structure. Whilst the former tasks are typically less computationally demanding than the latter one and therefore might be added as an auxiliary feature to a larger code (e.g.

Crystal17¹⁶⁷), they often require a degree of flexibility which situates them more comfortably in the realm of the programming library. Indeed the bridging of scales between the periodic crystal, finite cluster, dimer, and monomer is challenging to generalise due to the formidable conformational variety of organic molecules.

An optimal strategy is therefore to provide modular tools for the investigator to tailor to their system, accompanied by ready-made scripts which compile those tools for use in ubiquitous cases. In this way, non-expert users are able to use the program's principal features with a reliable degree of robustness, whilst more comfortable users can repurpose and extend the code to better suit fringe cases.

There exists a variety of computational chemistry scripting libraries for different tasks, with the ecosystem in rapid development. Their promise is to optimise and standardise the research workflow for increasingly specialised operations thanks to robust, tailored tools. The Cambridge Structural Database (CSD) Python API¹⁶⁸ focuses on crystallographic property analysis, accessing its associated database. The Atomic Simulation Environment (ASE)¹⁶⁹ specialises in interfacing with numerous electronic structure codes and communicates with them *via* Python scripting or a GUI. RDKit¹⁷⁰ provides programming tools for general purpose chemoinformatics and is itself used in numerous child libraries. Chemshell provides additive QM:MM interfaces between electronic structure and forcefield codes.¹⁷¹ Libra is a library designed for the development of quantum and classical dynamics.¹⁷² To our knowledge, there still lacks a library dedicated to the examination of photochemistry in molecular aggregates and crystals, exploiting the overlap in methodology between the two materials.

To address this, we offer the FRamewOrk for Molecular AGgregate Excitations (fromage). `fromage` is a standalone Python library, accompanied by ready-to-use command line scripts destined to facilitate the study of molecular aggregates in the excited state. They are summarised in Figure 5.1. The program is tested for Python 2.7 and 3.6, though the authors do not guarantee backwards compatibility with Python 2 in future releases. For geometry manipulation routines, the program only relies on the unit cell's Cartesian structure and lattice vectors. From this information, the user can obtain unique dimer configurations, molecular clusters and general structural information.

All of the electronic structure calculations are performed by popular quantum chemistry programs. Currently, interfaces are provided to run calculations in DFTB+,¹⁴² Gaussian,¹³⁹ Molcas,¹⁴⁰ and Turbomole.¹⁴¹ By delegating these calculations to different programs, their results can be combined into hybrid energy expressions. Thus fromage can perform ONIOM calculation—including those discussed in Chapter 4—whilst taking advantage of the diversity of modelling methods of several programs instead of only one.⁶¹

fromage has been used to study various QM:QM' electrostatic embedding schemes for applications in photoactive molecular crystals,⁶¹ the aggregation induced emission process in propeller-shaped molecules,¹⁷³ and the design principles for proton transfer luminescent materials.¹⁷⁴

Users interested in using the program should refer primarily to the full documentation, available at <https://fromage.readthedocs.io/>. It contains tutorials, and technical instructions for all primary operations. This chapter presents an overview of the program's capabilities, though readers wishing to use them should refer primarily to the documentation. First we describe the program structure, illustrated by basic usage examples. Then we enumerate its principal features, namely geometric analysis, exciton coupling evaluation and ONIOM QM:QM' models. This final point addresses the problems of inadequate implementation for cluster models in organic molecular crystal photochemistry discussed in Section 2.8.4.

5.2 Program Structure

5.2.1 Principal Classes

fromage makes use of two main classes in most of its operations, `Atom` and `Mol`. An `Atom` object is defined as a point in Cartesian space with associated physical properties. If the point is to represent an atom, supplying its element string provides standard properties such as covalent radius or atomic mass. The partial charge can also be specified, which can become useful in representing both atoms and point charges.

In practice, the user has little direct interaction with `Atom` objects. Instead they manipulate `Mol` objects, which contain a list of the former. `Mol` extends common

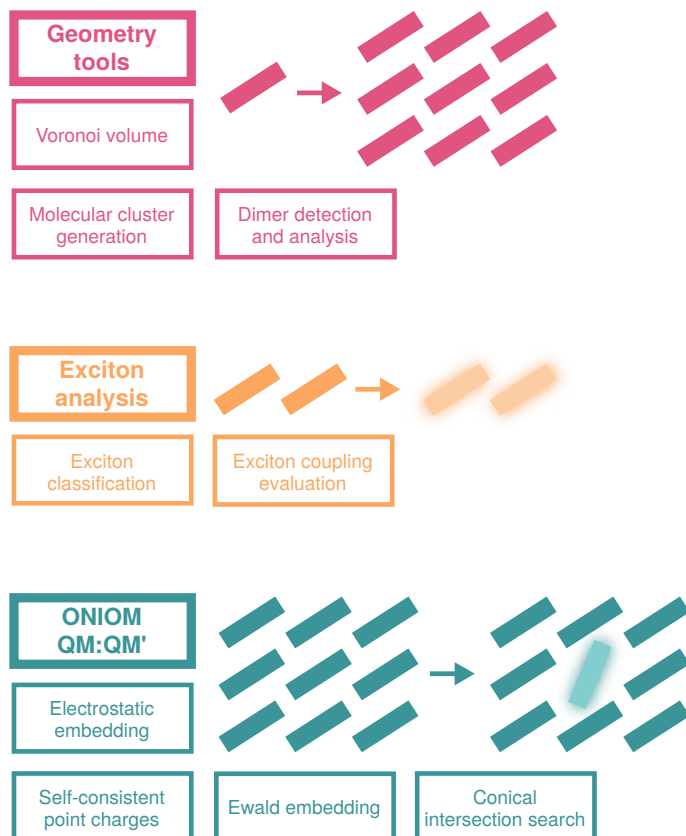


FIGURE 5.1: Principal features of fromage.

Python list methods *via* composition, allowing for intuitive appending, indexing, iterating etc. but also provides methods to manipulate molecular aggregate geometries and unit cells. `Mo1` objects can be created explicitly or generated from a typical geometry file such as `.xyz`. For instance, one may generate a molecular cluster and a supercell from unit cell information as illustrated in Figure 5.2.

Apart from the constituent atoms and lattice vectors, `Mo1` also has attributes pertaining to the definition of a bond within the collection of atoms. Two atoms are said to be bonded when their distance falls below a certain threshold `mo1.thresh`. This distance can be measured from nucleus to nucleus but also from the edge of the spheres of van der Waals radius or covalent radius. The method is selected by varying the `mo1.bonding` attribute. Having this flexibility is required for highly distorted molecular geometries or diverse element combinations. Armed with the definition of a bond, the `Mo1` class can single out covalently bonded complexes from an aggregate, generate molecular clusters from a single crystal and detect atomic connectivity.

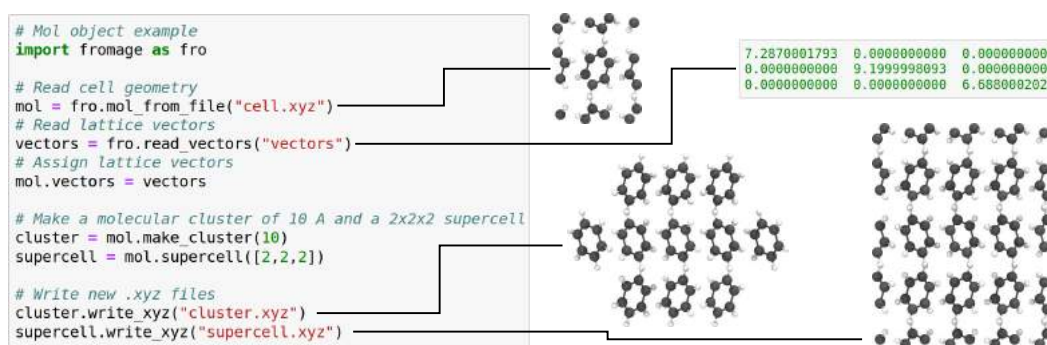


FIGURE 5.2: Example of use of `fromage` as a Python library. The three dimensional structures are shown in orthographic projection.

These tools are in and of themselves useful as a library for the Python literate user. However, several ready-made scripts are supplied for more complicated procedures and frequently required operations. Of the most practical use is perhaps `fro_uc_tools.py`, a command line script which performs operations related to unit cells. It is essential for comfortably communicating between periodic and finite systems which is a central concept in `fromage`. For instance, given a unit cell geometry file `cell.xyz` and a text file with the lattice vectors `vectors`, the line:

```
fro_uc_tools.py cell.xyz vectors -r 15
```

will produce a file `cluster_out.xyz` containing the geometry of a cluster of whole molecules where all atoms lie within a radius of 15 Å from the origin.

The other scripts are used to operate on unit cells, clusters, dimers and monomers. They are discussed in section 5.3 and represented in Figure 5.3 along with the core class structure.

5.2.2 Dependencies

The most common calculation performed by the geometry manipulation routines is the evaluation of interatomic distances. Therefore, for an overall speedup, these distances are calculated in C++ and wrapped as a Python function using SWIG. Some other more involved operations are sped up by using the `numpy`¹⁷⁵ library. Geometry optimisation is carried out with the BFGS implementation in `scipy`.¹⁷⁶

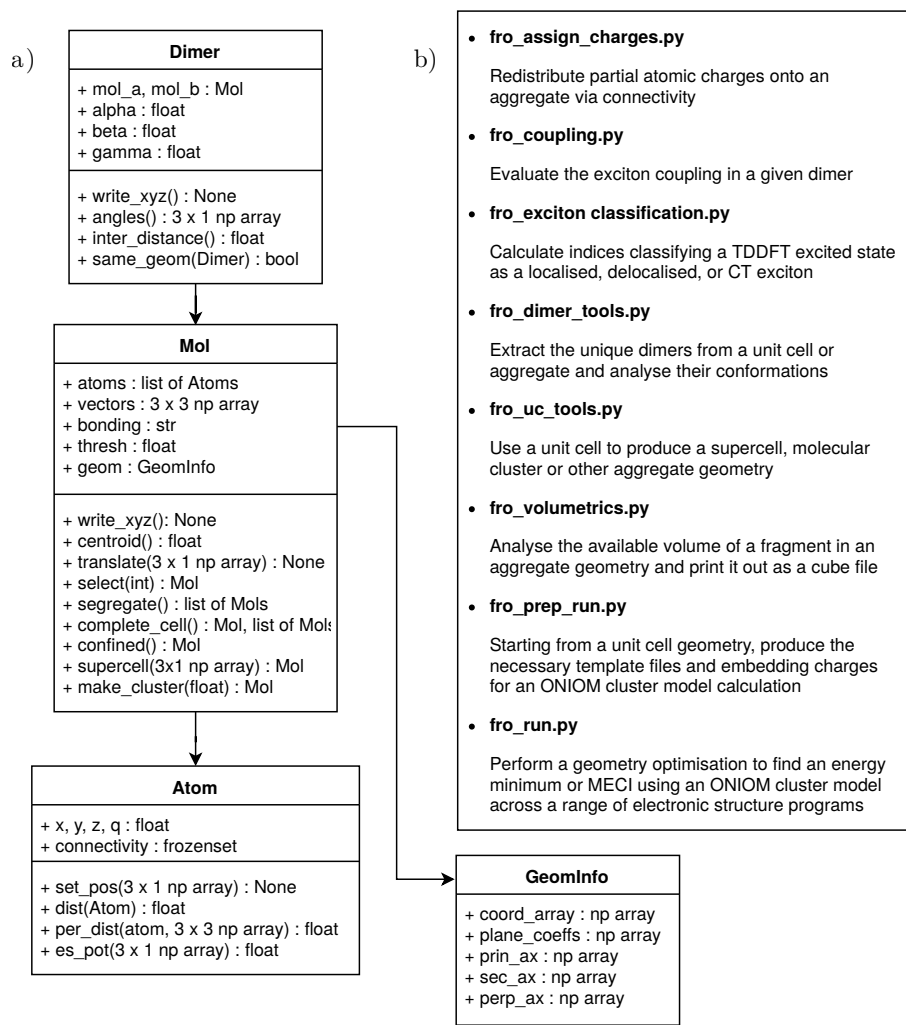


FIGURE 5.3: a) Principal class diagram of fromage b) Description of the main callable modules.

The program Ewald by Derenzo *et al.*^{113,114} is used for the fitting of point charges to the Ewald potential. It is modified to allow the use of partial charges and redistribution with permission from the authors here <https://github.com/Crespo-Oterogroup/Ewald>.

5.3 Features

5.3.1 Geometrical Analysis

Voronoi Volumes

The conformational freedom of a molecule in the gas phase is only limited by its structural features and how they relate to the potential energy surfaces (PES) of particular reaction coordinates. The electronic excitations within a molecule can in general be rationalised by scrutinising and comparing its different PESs. In contrast, in condensed phases, a molecule's freedom of nuclear reorganisation is hindered by the close packing imposed by its environment. The study of the effect of this close packing on excited state PESs and photochemical behaviour is a vast topic and involves an accumulation of inter-related factors both electronic and nuclear in nature which are often hidden behind the deceptively concise term of *steric hindrance*.^{132,135}

It is however appropriate to begin such an investigation by obtaining computationally inexpensive and easily interpretable features of the packing of the aggregate. A routine approach for crystals is dividing the unit cell volume by the amount of molecules in the cell to find the average volume V_c assigned to the molecule within the packing pattern in order estimate the tightness of packing. However this volume is difficult to compare between crystals composed of different molecules and is not visualisable. The determination of Voronoi volumes for molecules in aggregate is a promising alternative.

In an aggregate of atoms a point belongs to the Voronoi volume of a given molecule if the atom it is nearest to belongs to that molecule. The application of Voronoi cells¹⁷⁷ to molecular systems has successfully been used to characterise the geometry of condensed phases¹⁷⁸ and notably liquids.¹⁷⁹ This definition is refined by scaling the distance metric by the van der Waals radius of the atom; thus, for instance, assigning more space to oxygen than to hydrogen. For the calculation of distances, we employ a grid based scheme, which makes it robust and allows us to choose an arbitrary resolution for the volume.^{180,181} The resulting Voronoi volume V_V can be compared to the sum of the van der Waals spheres of the atoms in the molecule (counting their intersection only once) V_{vdW} to obtain a volumetric index $V_i = \frac{V_V}{V_{vdW}}$ which

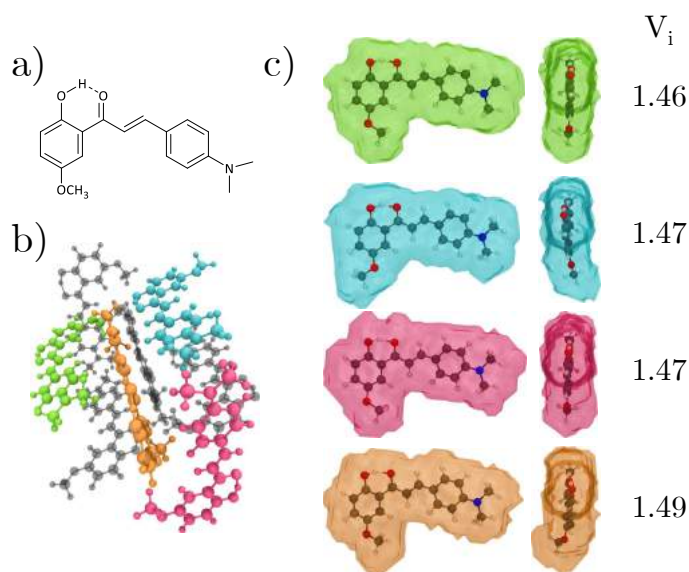


FIGURE 5.4: a) Chemical structure of HC2 b) unit cell of the HC2 crystal with the asymmetric cell highlighted by colour c) Profiles of the Voronoi volumes for the four molecules of the asymmetric cell with their volume index.

gives a normalised indication of the tightness of packing in the crystal for a specific molecule.

The Voronoi volumes introduce a distinction between inequivalent molecules in a crystal, and in fact should average out to V_c if the V_V values are weighted by the amount of occurrences of the molecule per unit cell. They can be used in finite systems such as amorphous clusters and perhaps most importantly, they are visualisable. Seeing the shape of a Voronoi volume can indicate the available space in the aggregate and therefore the areas of the PES least restricted by the environment.

A user may open a cluster geometry file `clust.xyz` with a visualising program and choosing which molecule they wish to calculate the Voronoi volume of with `fromage`. They identify the molecule by marking down the label of any atom belonging to the molecule. Then, upon calling:

```
fro_volumetrics.py clust.xyz -l [atom label]
```

the program will generate the files `voro.cube`, `vdw.cube` and `union.cube` which are the visualisable Voronoi and van der Waals volumes of the molecule, and their union (in the set theory sense). A file called `volumes` contains the integrated volume of each of these.

An illustrative example is a derivative of 2'-hydroxychalcone (HC1), which upon the addition of a methoxy group in *para* position with respect to the hydroxyl group turns off its emissive character in crystal form, bringing the fluorescence quantum yield from 0.32 to less than 0.01.^{134,158} The new molecule is HC2, discussed in Chapter 4. While HC1 exhibits herringbone style packing, HC2 has a complex unit cell structure whose steric constraints on the individual molecules are unclear at first glance. Figure 5.4 shows the difference in tightness of crystal packing between inequivalent monomers of HC2.

Dimeric Arrangement

Excitations in molecular aggregates are not guaranteed to remain confined to one absorbing monomer. Indeed, the electronic wavefunctions of the neighbouring molecules may have enough intermolecular overlap to produce intermolecular electronic interferences in the excited state. This can manifest in all sorts of photochemical processes and is central to exciton governed mechanisms like charge transfer or singlet fission. In crystals, typical packing motifs like herringbone or sheet-like, produce a limited set of archetypal dimer arrangements such as edge-to-face or face-to-face which have generalisable excitonic behaviour for chemically similar molecules.^{107,174} For instance in face-to-face aromatic systems, $\pi\pi$ interactions are a defining feature of the excitonic states.

Regardless of whether a researcher is investigating an amorphous cluster or a crystal structure, it is therefore informative to extract all of the significant dimers in the system and to quantify their geometrical arrangement in order to classify them under the principal dimer archetypes at a glance. In *fromage*, the user can extract the possible dimers whose distance falls below a given threshold. The intermolecular distance can be defined either as the centroid-to-centroid distance or as the nearest intermolecular atom pair distance. The latter can also be complemented by the van der Waals radii of the atoms. If the molecules are in lattice positions, the symmetry of the unit cell will produce groups of dimers identical up to a reflection or rotation. To filter out repeated configurations, all the intermolecular atomic distances are evaluated and sorted which provides a fingerprint for the dimer geometry. Equivalent

dimers are then defined as ones with the same fingerprint up to an RMSD threshold of 10^{-4} Å by default.

The dimeric arrangement can then be characterised quantitatively. An orthonormal set of principal, secondary, and tertiary axes is calculated for each constituent fragment and the angles between same axes of two molecules can be associated to an archetypical dimer, effectively classifying the pair.

The generalised procedure to obtain characteristic vectors for a molecule is shown in Figure 5.5:

1. All atoms of the molecule are projected onto an averaged plane by singular value decomposition.
2. The two longest interatomic distances are identified, forming a quadrilateral $ABCD$ such that $AC > BD$ and AB is the longest side. This imposes an arbitrary but consistent direction for the vectors.
3. The following midpoints are detected: $[AB] \rightarrow H$; $[BC] \rightarrow G$; $[CD] \rightarrow F$; $[DA] \rightarrow E$. The principal and secondary vectors \vec{a} and \vec{b} respectively go from E to G and F to H .
4. The two vectors are normalised and rotated equally until they are perpendicular. The tertiary vector is $\vec{c} = \vec{a} \times \vec{b}$.

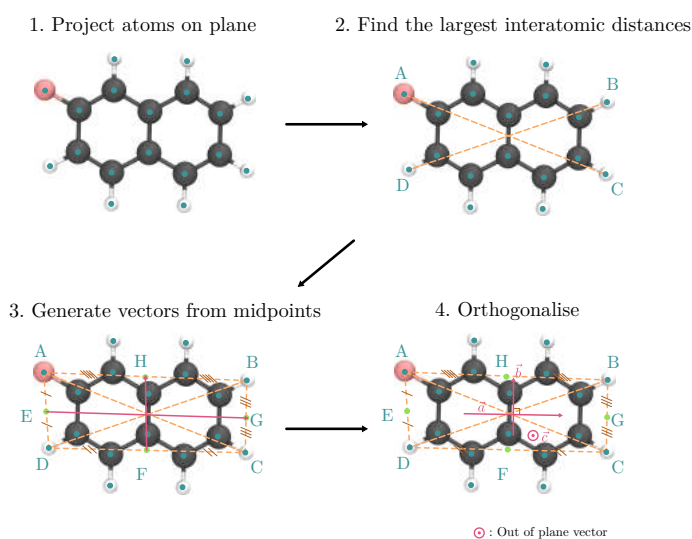


FIGURE 5.5: Algorithm to generate principal axes from a monomer geometry with 2-fluoronaphthalene as an example.

This orthonormal set of axes presumes the significance of a plane which defines the shape of the molecule. Conjugated organic systems, those with optical applications, are often planar due to their aromatic components.⁶ In other cases, the judicious elimination of extraneous atoms from the analysis—a feature present in the code—can reduce the significant coordinates to a plane. To exclude a particular atom type from the calculation, only one atom need be specified, and others with the same identity can be automatically detected using the method outlined in section 5.3.3.

If one same atom is involved in both largest interatomic distances, the procedure remains valid and the points C and D become degenerate. In certain cases, the principal axis should simply be the vector connecting the largest interatomic distance. When this option is selected, the secondary axis becomes the perpendicular axis which lies on the averaged plane. In the general case where one wishes to extract the geometric information from a cluster of dimers `clust.xyz`, one should use the command:

```
fro_dimer_tools.py clust.xyz
```

which will return the file `dimers.dat` containing all of the angles between dimers, the centroid distances and a proposed classification into different archetypical dimers. An additional slip angle is printed, to estimate the amount of face-to-face overlapping area giving rise to $\pi\pi$ interactions. It is defined as the smallest angle between the centroid-to-centroid axis and either tertiary axis of the constituent monomers, as is discussed in Ref. 174.

The above method was used to investigate (2E)-3-(dimethylamino)-1-(2-hydroxy-4-methoxyphenyl)-2-propen-1-one (DMAH). This molecule exhibits lasing behaviour with a fluorescence quantum yield of 0.77 in crystal form compared to 0.19 in PMMA film.⁹⁷ The crystal structure was optimised in Quantum Espresso using PBE-D2 with a plane-wave cutoff of 30 Ry and a 8x6x6 k-point mesh. A cluster of molecules was extracted from its crystal positions and all dimers with centroids falling less than 10 Å from each other were considered. In this case, certain nonessential atoms were ignored in the geometric analysis and the points B and C became degenerate. The results are represented on Figure 5.6. The points at (44°,115°) and

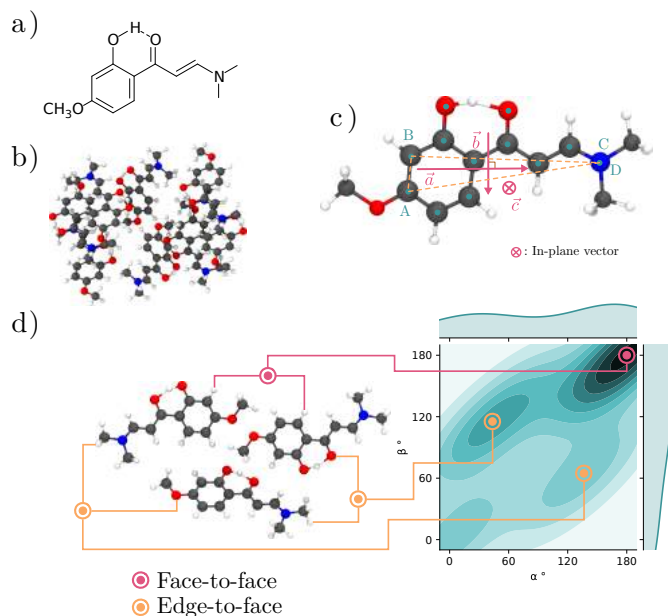


FIGURE 5.6: Case study of DMAH for dimer geometry analysis a) Chemical structure b) Crystalline cluster c) Principal axes determination d) Heatmap of the angles between primary and secondary axes, respectively α and β .

($136^\circ, 65^\circ$) have a large β which is characteristic of the edge-to-face dimers in herringbone packing. However the α values are unusually far from 0° or 180° , showing a packing arrangement specific to this crystal. The point at ($0^\circ, 0^\circ$), is in this case only related to the dimers which monomers form with periodic images of themselves.

5.3.2 Exciton Analysis

Exciton Classification

The exciton model is a framework to characterise different types of many-body electronic excitations in collections of molecules. The excitation can be associated with a localised Frenkel exciton, or a charge-transfer state, which are characterised by the electron density respectively migrating intra- and intermolecularly. A delocalised Frenkel excitation corresponds to an electron density reorganisation throughout the dimer with no net charge transfer. Differentiating between these three behaviours is crucial for the design of organic semiconductors in solar cells, since the process of charge separation and migration is fundamental to their mechanism. One approach is that of analysing the one-electron transition density matrices associated with the

excitation, which contains information about the migration of the charge during the process.¹⁸² This method is implemented in TheoDORÉ.¹⁸³

Refs. 184 and 185 propose a different scheme to qualify the nature of the exciton by comparing the charge density distribution on fragments of a dimer before and after excitation. The original implementation is in a program named CALC DEN which combines Perl and Fortran. The reimplemention in fromage is Python importable, making it an attractive alternative for use and extension depending on the user's preference.

A Mulliken partition scheme can supply integrated orbital-specific density coefficients located on one molecule (A):

$$\rho_A^k = \sum_{\mu \in A, \nu \in A} c_{\mu k} c_{\nu k} S_{\mu\nu} + \sum_{\mu \in A, \nu \in B} c_{\mu k} c_{\nu k} S_{\mu\nu} \quad (5.1)$$

$S_{\mu\nu}$ is the overlap integral between basis functions ϕ_μ and ϕ_ν , and $c_{\mu k}$ is the coefficient of basis function ϕ_μ in molecular orbital k . This density coefficient, ρ_A^k , can be used to produce two indices related to an excitation I :

$$\begin{aligned} \Sigma P_A^I &= \sum_{i \rightarrow j} \sigma_{ij} (C_{i \rightarrow j}^I)^2 (\rho_A^j + \rho_A^i) \\ \Delta P_A^I &= \sum_{i \rightarrow j} \sigma_{ij} (C_{i \rightarrow j}^I)^2 (\rho_A^j - \rho_A^i) \end{aligned} \quad (5.2)$$

Where $C_{i \rightarrow j}^I$ is the TDDFT coefficient corresponding to the excitation from orbital i to j and σ_{ij} is 1 for $i < j$ and -1 for $i > j$, in the case of de-excitation. These indices are in units of e^- and since they are associated to one excited state only, they have bounds $0 \leq \Sigma P_A^I \leq 2$ and $-1 \leq \Delta P_A^I \leq 1$. ΣP_A^I represents the amount of electrons changing populations of atomic orbitals within A for a given excitation. ΔP_A^I , on the other hand, represents the electrons from A populating atomic orbitals of B upon excitation, or *vice versa*.

The combination of the two quantities indicates the behaviour of the electronic density upon excitation; a large reorganisation of density confined to molecule A (labelled LOC(A)) would manifest in an extreme ΣP_A^I , closer to $0 e^-$ for molecule B and closer to $2 e^-$ for molecule A . On the other hand, extreme ΔP_A^I values indicate

a net loss or gain of density by one molecule. Values closer to $-1 e^-$ correspond to a charge transfer from A to B (CT(A \rightarrow B)) and *vice versa* for values close to $1 e^-$. If none of the indices have extreme values, the excitation is delocalised, labelled DELOC. In fromage, an arbitrary threshold is in place by default where an excitation less than $0.5 e^-$ from an extreme value is classified as the corresponding type of exciton.

This method inherits the limitations of Mulliken population analysis, and basis set superposition, but has the advantage of only relying on TDDFT coefficients and overlap matrices, making it much faster than methods which investigate the spatial features of the density difference between excited and ground state.

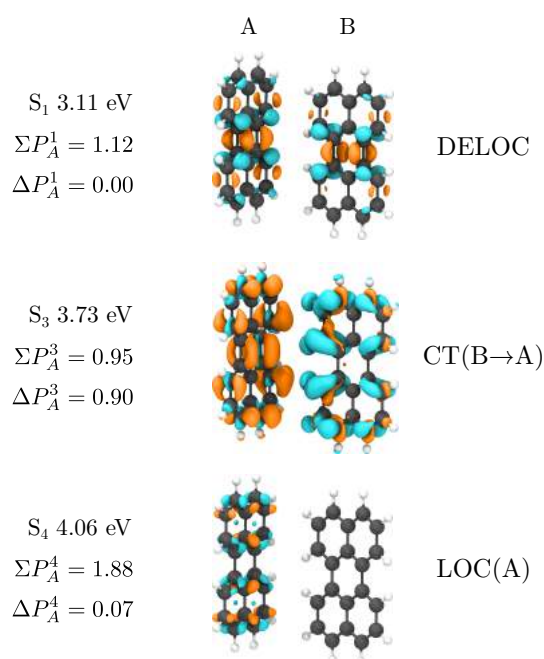


FIGURE 5.7: Transition density of two excited states of a perylene dimer taken from its crystal structure. The excited state energies are reported with respect to S_0 . The electron density migrates from the blue to the yellow areas upon excitation. The exciton classification indices are defined in equation 5.2 and are in units of e^- .

To evaluate the two indices and classify an excitation, the user should first perform a TDDFT or CIS calculation of the dimer. Currently, only Gaussian calculations are supported. They should ensure that the atoms of molecule A appear before those of B in the geometry field and that an `rwf` file is produced by using the option `%rwf=[name].rwf`. Then, given the output files `tddft.log` and `tddft.rwf`, the command line:

fro_exciton_classification.py tddft.log tddft.rwf [number of excitation]

will print out the values of ΣP_A^I and ΔP_A^I along with a suggested classification (DELOC, LOC(A), LOC(B), CT(B→A) and CT(A→B)).

To illustrate the use of this feature, a dimer of perylene was extracted from its experimental crystal structure.¹⁸⁶ Its excited states were calculated in Gaussian16 using TD- ω B97X-D/6-31G(d), and the transition densities analysed using the orbital specific Mulliken partition scheme. The states S_3 and S_4 each had extreme values of the classification indices, with $\Sigma P_A^3 = 0.95 e^-$ and $\Delta P_A^3 = 0.90 e^-$ for the former, indicating a charge transfer from fragment B to a, and $\Sigma P_A^4 = 1.88 e^-$ and $\Delta P_A^4 = 0.07 e^-$ for the latter, corresponding to a transition confined to monomer A. Figure 5.7 corroborates this classification by showing the electronic transition density of a DELOC S_1 and comparing it to the CT(B→A) S_3 and LOC(A) S_4 .

Exciton Coupling Evaluation

The coupling associated with an exciton is a measure of the correlation between the individual isolated excitations within the exciton. Evaluating it can give us a quantitative bearing on the importance of the many-body effects in the excited state process.

Kasha's exciton model is the initial approach, reducing the electronic densities of the fragments to point dipoles and comparing their relative geometry.¹⁰⁷ Under the point dipole approximation (PDA), the exciton coupling between fragments i and j is expressed as:

$$J_{ij}^{\text{Kasha}} = \frac{\boldsymbol{\mu}_i \boldsymbol{\mu}_j}{R^3} - \frac{3(\boldsymbol{\mu}_i \cdot \mathbf{R}_{ij})(\mathbf{R}_{ij} \cdot \boldsymbol{\mu}_j)}{R^5} \quad (5.3)$$

Where $\boldsymbol{\mu}_n$ is the electronic transition dipole moment for monomer n and \mathbf{R}_{ij} is the vector connecting the centroids of monomers i and j .

For a better spatial resolution of the electrostatic interactions, one can instead calculate the interaction between atomic transition charges (ATC)¹⁶⁶:

$$J_{ij}^{\text{ATC}} = \sum_a^{N_i} \sum_b^{M_j} \frac{q_a q_b}{|\mathbf{R}_a^i - \mathbf{R}_b^j|} \quad (5.4)$$

Where R_c^k is the position of atom c of monomer k with ATC q_c and N_k is the number of atoms of monomer k .

For more resolution, the transition electronic density itself can be used, yielding costly but exact Coulombic exciton coupling. Additional correction can be included, for example by including the dielectric response of the environment as a polarisable continuum model as is done in EXAT.¹⁰⁶

These models are purely Coulombic in nature and do not take into account the short-range interaction affecting the excited state behaviour of adjacent molecules, which for example is dominant in charge-transfer states.

Aragó and Troisi have devised a procedure which evaluates the coupling J_{ij} in a dimer by diabatising the Hamiltonian:¹⁰⁵

1. Select an excited state property. For the sake of argument we will use the transition dipole moment (TDM) μ but anything bearing relation to the excited electronic density would be applicable.
2. Evaluate the property for the two lowest excited states, along with the energies, for the dimer. These are the adiabatic energies (E_1^A and E_2^A) and adiabatic TDMs (μ_1^A and μ_2^A).
3. Evaluate the property for both isolated constituent monomers in the first excited state, retaining their orientation. These TDMs are labelled μ_1^{ISO} and μ_2^{ISO} .
4. Calculate the singular value decomposition $(\mu^A)^* \mu^{ISO} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$.
5. Compute the matrix $\mathbf{C} = (\mathbf{U}\mathbf{V}^*)^*$ which is the best unitary transformation matrix mapping the adiabatic to the diabatic basis.
6. Compute the diabatic Hamiltonian:

$$\begin{pmatrix} E_i^D & J_{ij} \\ J_{ji} & E_j^D \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \begin{pmatrix} E_i^A & 0 \\ 0 & E_j^A \end{pmatrix} \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$$

The off-diagonal elements of the Hamiltonian the exciton coupling values J_{ij} . More details can be found in reference 105. The diabatisation method has already been employed in fromage to investigate the aggregate behaviour of propeller-shaped emitters.¹⁷³

We have further extended it to calculate N -dimensional diabatic Hamiltonians, thus, for example, allowing for the calculation of pairwise exciton coupling within a trimer in the excited state, taking into account the influence of the third monomer.

An additional approximate method is that of exploiting the exciton energy splitting of the molecular excited state upon formation of a dimer. In the dimer, the S_1 and S_2 states are separated by twice the magnitude of the exciton coupling, provided that the individual constituent molecules are in perfect resonance.¹⁸⁷ Even in less symmetric cases, this approximation has been used with reasonable success.^{133,174,188}

`fromage` implements this half-gap method, the PDA Kasha model, the ATC Coulombic interaction model and the diabatisation scheme using either ATCs or TDMs as excited state properties.

The script `fro_coupling.py` manipulates Gaussian log files to compute exciton couplings. Several schemes are implemented. For instance, if a user requires the diabatisation method using the TDM excited state property to find the diabatic Hamiltonian containing the first three couplings of a trimer, the steps would be as follows. Carry out Gaussian calculations of the S_1 states of the three constituent monomers and the first three excited states of the trimer, yielding the files `mon_1.log` `mon_2.log` `mon_3.log` and `trim.log`. Then, use the command line:

```
fro_coupling.py -m DIA -p TDM -mf mon_1.log mon_2.log mon_3.log -nf trim.log -ns 3
```

This will print out the diabatic Hamiltonian where the three lower triangular off-diagonal elements correspond to the three exciton couplings of the trimer.

As an illustrative example, the crystal structure of 1,4-bis-(4-styryl-styryl)-benzene (4PV) was used to extract inequivalent dimers and evaluate their couplings. The structure of 4PV has been experimentally shown to contain six monomers, departing from the high symmetry of an ideal herringbone crystal due to variable slight rotation of the extreme phenyl rings.¹⁸⁹ First, the unit cell was optimised using PBE-D2 with a basis set cut-off of 50 Ry and a Monkhorst-Pack grid of 1x2x1 k-points as implemented in Quantum Espresso.¹⁵² Then, the inequivalent dimers were detected by `fromage` using a centroid-centroid distance threshold of 7 Å. The TDMs of these dimers were calculated using TD- ω B97X-D/6-31G(d) in Gaussian16¹³⁹ and processed in `fromage` in order to evaluate exciton couplings using dimer and trimer

diabatic Hamiltonians. The results are shown in Figure 5.8. In this system, edge-to-face dimer arrangements have larger exciton couplings (97, 103 and 105 meV) than face-to-face ones (79 and 91 meV). The largest difference is of 26 meV which represents 25% of the greatest coupling (105 meV). It is expected that in cofacial dimers with couplings of such magnitude, the short-range interactions should account for most of the coupling, making the use of a coupling scheme which accounts for exchange imperative.⁹¹

Overall, the inclusion of the third molecule in the trimer diabatic Hamiltonian reduces the magnitude of the couplings by about 10 meV which is significant since it is in the order of the difference between certain edge-to-face and face-to-face values.

The irregularities in the herringbone packing produce two different face-to-face dimers with respective centroid-centroid distances of 6.05 Å and 6.15 Å. This relatively small increase in distance reduces the exciton coupling by 12 meV (14% of the average value), indicating that the natural packing of 4PV produces in dimers in an excitonically sensitive geometry.

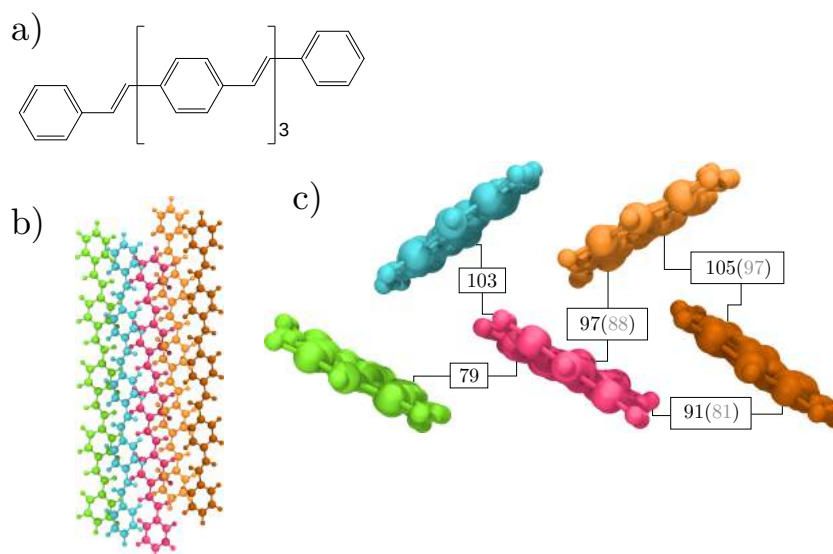


FIGURE 5.8: a) Chemical structure of 4PV b) Top view of a cluster of 4PV molecules taken from their crystal positions c) Side view of the cluster with inequivalent dimers labelled by their exciton coupling value in meV. The values in parentheses are calculated as a trimer.

5.3.3 ONIOM Calculations

Whether it be to calculate an absorption or emission energy, locate nonradiative pathways, or build a potential energy surface, calculating an excited state electronic structure is an often unavoidable step in the study of molecular aggregates. This often proves to be challenging due to the environmental effects in such media.

As discussed in Chapter 2, hybrid method schemes can offer environmental corrections, where an active site calculated with a high accuracy method is embedded in an explicit environment of lower accuracy. Inter-program hybrid method codes are not uncommon. GARLEEK¹⁹⁰ communicates between Gaussian and Tinker for subtractive QM:MM calculations. Chemshe11^{118,171} provides additive QM:MM formulations. The Atomic Simulation Environment (ASE)¹⁶⁹ offers both additive and subtractive QM:MM as a Python library. However to our knowledge no such code yet offers inter-program ONIOM QM:QM' calculations in the excited state. `fro` uses its interfaces with DFTB+, Gaussian, Molcas, and Turbomole to calculate energies and geometry optimisations of different kinds with ONIOM QM:QM'. Namely mechanical embedding, regular electrostatic embedding, Ewald point charge embedding, and self-consistent versions of the last two.

The point charges can originate from Mulliken or RESP calculations. To fit point charges to the Ewald potential, we use the Ewald program.^{63,113,114} It was modified to allow for the use of noninteger charges and redistributed at <https://github.com/Crespo-Otero-group/Ewald> with permission from the original authors, fulfilling the requirements of the Computer Physics Communications license.

The various quantum methods available and tested in `fro` are listed in Table 5.1 along with their corresponding programs.

Setting up hybrid method calculations is typically a technically tedious task, but as many steps as possible are automated in `fro` while retaining the full flexibility of the interfacing programs. The line:

```
fro_prep_run.py
```

accompanied with a few input files such as the unit cell and a configuration file, will prepare the template files for the actual calculations to follow. This may include

TABLE 5.1: Interfaced quantum methods and their availability in electronic structure programs

Method	DFTB+	Gaussian	Molcas	Turbomole
Hartree-Fock	X	✓	X	✓
DFTB	✓	X	X	X
DFT	X	✓	X	✓
TDDFT	X	✓	X	✓
ADC(2)	X	X	X	✓
CC2	X	X	X	✓
CASSCF	X	✓	✓	X
CASPT2	X	X	✓	X

cluster geometry generation, Ewald fitting procedures and self-consistent population analyses.

Then, if the user is satisfied with their newly generated embedded cluster, the line:

```
fro_run.py
```

can perform geometry optimisation or single point calculations.

OEC, OEEC and SC-OEEC calculations all require the distribution of point charges from single molecule population analyses to aggregate geometries or unit cell structures. This is a routine operation for preparing forcefield calculations, where sets of atomic types are associated to a potential or another with predetermined charge values.^{191,192} The atomic type is usually dependent on its neighbouring bonded atoms and the bond types. For QM:QM' calculations, it would be more attractive to use a broader definition of the atomic type which distinguishes between same function atoms at nonequivalent parts of the molecule.

To this end, *froimage* implements a connectivity detection tool which reads population analysis information from a single molecule or unit cell calculation and redistributes it onto any other finite or periodic ensemble of same molecules. The procedure builds a bond order matrix \mathbf{B} where B_{ij} is the shortest path connecting atoms i to j in number of bonds. The construction of the matrix is represented in Figure 5.9 and is performed as follows:

1. Detect the first connections by computing all of the interatomic distances complemented by atomic radii.

2. Check every zero element and detect atom pairs (a, b) which have a connected atom c in common, i.e. $B_{ac} \neq 0 \wedge B_{bc} \neq 0$.
3. Assign $B_{ab} = B_{ac} + B_{bc}$.
4. Repeat from step 2 until convergence of the matrix.

An atom's identity can be completely defined by additionally using the element of the atom corresponding to each row of **B**. A sufficient fingerprint for a given atom is the amount of atoms of a given element which are located a certain amount of bonds away, accompanied by the atom's own element. For example a methyl hydrogen of methanol is sufficiently defined by stating that it is a hydrogen atom with one carbon atom one bond away, one oxygen atom two bonds away, two hydrogen atoms two bonds away and one hydrogen atom three bonds away.

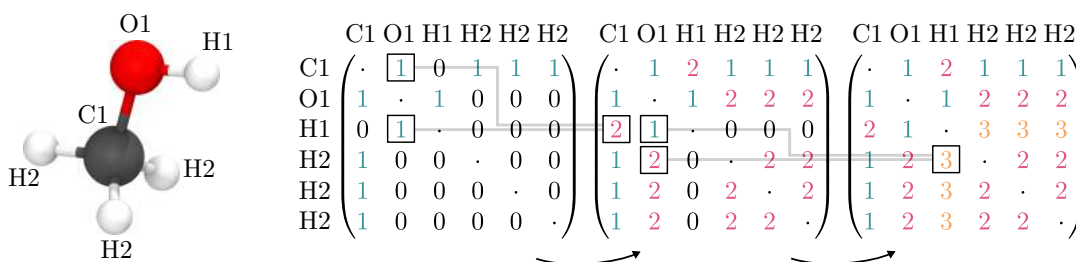


FIGURE 5.9: Generation of the bond order matrix for methanol.

Once all of the atomic identities are sampled from a reference molecule or unit cell, the partial charge values from this reference are distributed onto the desired target. When several reference atoms have the same identity (for instance three hydrogens belonging to the same methyl group), the average charge is retained. In practical terms, given for instance a Mulliken population analysis of methanol calculated in Gaussian with output file `pop.log`, and a Cartesian coordinate file of a cluster of methanol molecules `clust.xyz`, the line:

```
fro_assign_charges.py pop.log clust.xyz
```

will produce a file `out_char` which contains the coordinates of `clust.xyz` with an added column stating the partial charge of each atom.

All geometry optimisation routines use the the BFGS algorithm as implemented in `scipy`.¹⁷⁶ To locate minima on the PES, the ONIOM energy expression (Equation

2.9) is minimised. To locate conical intersections, the penalty function of Levine, Coe, and Martínez is used instead (see Section 4.2).¹⁴³

Example of Use

The optimisation of excited state geometries in detailed condensed phase environments has numerous applications. For instance, the fluorescence of HC1 (see section 5.3.1) can also be switched off by the substitution of a methyl group in *para* position with respect to the hydroxyl group, resulting in another dark compound: (2E)-3-[4-(Dimethylamino)phenyl]-1-(2-hydroxy-5-methylphenyl)-2-propen-1-one (DMAP).^{134,158}

Both HC1 and DMAP can experience excited state intramolecular proton transfer, splitting the excited state into two potential decay pathways. It was recently computationally shown that in HC, both the enol and the keto non-radiative decay pathways were rendered energetically inaccessible by a combination of molecular and crystalline factors.^{131,132}

OEEC was employed to investigate the critical points in the potential energy surface of DMAP. A thorough tutorial on how to perform these calculations is available here: <https://fromage.readthedocs.io/en/latest/tutorial.html>. Geometry optimisations were carried out to locate the ground and excited state minima, as well as its Minimal Energy Conical Intersection Geometries (MECI). The crystal structure was optimised in Quantum Espresso using PBE-D2, a 40 Ry basis cutoff and a 1x2x1 k-point mesh. For the ONIOM calculation, the QM level of theory was TD- ω B97X-D/6-311++G(d,p) while for the QM' level of theory, both HF/STO-3G and DFTB were employed. The energies and gradients were computed in Gaussian16 for TDDFT and HF and in DFTB+ for DFTB. Figure 5.10 shows the relative energies of all the optimised critical points of the PES.

The enol non-radiative decay pathway is significantly above the absorption energy, rendering this channel inaccessible. On the other hand, the keto conical intersection is only about 0.2 eV above the absorption energy, arguably making it accessible *via* thermal fluctuations. The availability of this nonradiative decay channel explains the low fluorescence quantum yield.

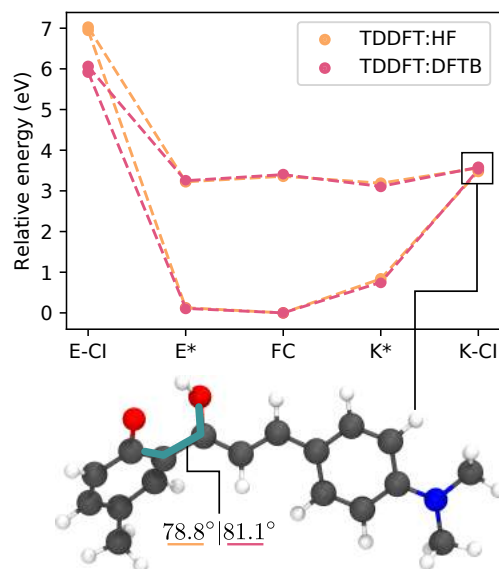


FIGURE 5.10: Relative energies of critical geometries of DMAP evaluated with OEEC at critical points of the PES. The Franck–Condon point (FC), Enol (E^*) and Keto (K^*) excited state minima and Conical Intersection (CI) geometries were located.

The TDDFT:HF and TDDFT:DFTB levels of theory have similar results, both in geometry and energy, with a difference of less than 0.1 eV at each point except for the enol MECI. This outlier can be attributed to the extreme bond stretching occurring between carbons in the back bond in the enol MECI conformation which would render both the Hartree-Fock formalism and the parameterisation of the DFTB calculations inadequate in differing ways. The striking agreement between the two methods is encouraging, given the relatively low computational cost of the semiempirical computation and the prevalence of HF as a ground state theory in QM:QM' protocols.^{13–16}

5.4 Conclusion

We have detailed the principal capabilities of the Python library fromage, which aims to facilitate the computational investigation of the excited states of molecular aggregates. They include geometrical and exciton analysis as well as QM:QM' geometry optimisation tools, which have already been successfully employed in three past publications.^{61,173,174} The features were tested on a diverse array of molecules,

in order to challenge their robustness. They are implemented with enough flexibility that Python literate researchers can employ `fromage` scripts as part of a larger workflow with little added effort. By virtue of being an open source, unit tested and documented piece of software, `fromage` represents an addition to the fast expanding pool of sustainable chemical software libraries. We hope that by enabling researchers to use the framework and manipulate the source code, the field of aggregate photochemistry will continue to mature towards modern reproducible workflows.

Chapter 6

Application

6.1 Introduction

With the new suite of tools implemented in *fromage*, we now focus on applying them to increasingly diverse systems, in the hope to glean some chemical insight. The objective is both to provide robustness to the program, whilst studying the properties of emissive organic molecular crystals .

To fully control the luminescent behaviour of these materials, the excited state radiative and non-radiative decay channels of the molecule need to be understood within a particular condensed phase environment. Despite the abundance in experimental and theoretical investigations of the excited states of molecular crystals, understanding these decay channels is still extremely challenging, and is done on a case-by-case basis for newly discovered compounds. The main obstacle to generalising design rules for these materials is the interconnectedness of their defining properties at the molecular and intermolecular levels.

In this chapter, we analyse the role of different factors affecting the emissive response of thirteen luminescent molecular crystals, which have previously been characterised experimentally (Figure 6.1). We aim to consider a large enough variety of systems that comparisons can be drawn between crystals with different degrees of packing and chemical similarity. Our main focus is to rationalise the Quantum Efficiency of Fluorescence (QEF) of these materials as a competition between radiative and nonradiative decay channels, arising from different inter- and intramolecular factors. We consider a diverse enough range of materials to highlight how differently behaving solute molecules can arrive at the same desirable emissive property

in the crystal phase.

The crystals were gathered into series based on their backbone structures and substituents. *p*-oligophenylenes (nP , $n=3, 4$, and 6) are a family of organic π -conjugated molecules composed of phenyl-rings attached to each other via single bonds in *para*-positions. The DCS series where three phenylene units are connected by vinylene bridges with cyano-group substituents, and additional butoxy and methoxy groups are added to the backbone. We also consider the DSB molecule, which shares the same backbone but has no substituents and 4PV which further extends the phenylene chain by two phenylene units. Additionally, we consider 2-(2'-hydroxyphenyl)benzothiazole (HBT), a molecule exhibiting excited state proton transfer in the solid state. All molecules are represented on Figure 4.1.

Barring certain substitutions of the DCS family and HBT, all systems display amplified spontaneous emission (ASE) in crystal, making them candidates for use as organic single crystal lasers, as detailed in Reference 6. All molecules undergo Solid State Luminescence (SSL),¹³³ some of it induced by crystallisation—Solid State Luminescence Enhancement (SLE).

We first present the computational details of our findings. Then, in order to understand the emissive behaviour of the crystals, we analyse the geometric features of the crystal packing, the excitonic coupling between constituent dimers, and the energetics of the excited states along critical points of their potential energy surfaces. We conclude by comparing the findings between series and assessing the efficacy of our available analysis methods in elucidating competing radiative and nonradiative mechanisms.

6.2 Computational Details

The crystal structure geometries were optimised using PBE-D2 as implemented in Quantum Espresso¹⁵², with a basis set cut-off of 50 Ry and various Monkhorst-Pack Grids chosen in accordance with the unit cell shapes.

The investigation of the multiple molecules was facilitated by the recent development of *fromage*, a Python library dedicated to studying excited state molecular

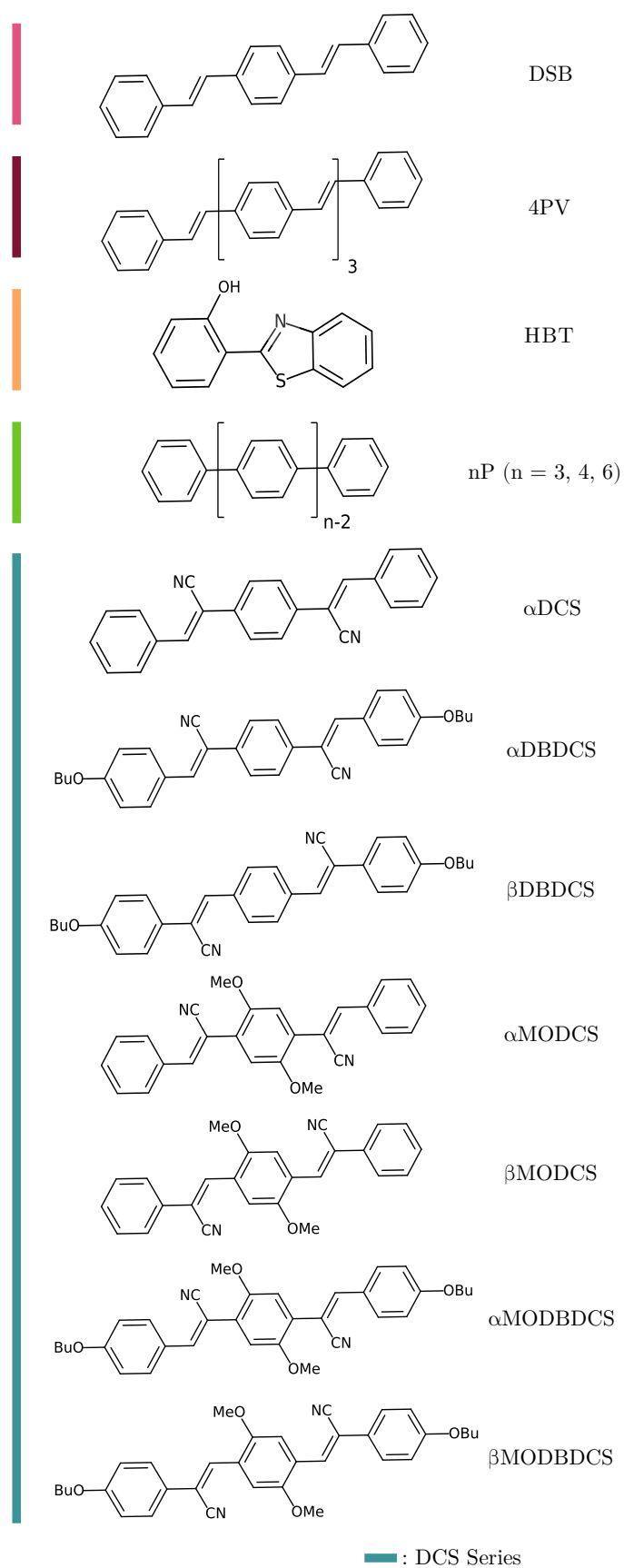


FIGURE 6.1: Molecular structures of the studied systems.

aggregates and crystals. This work showcases the robustness of its features by applying geometry analysis tools, excitonic coupling evaluation, and ONIOM methods to the crystals.

In order to isolate dimers from the lattice, a spherical molecular cluster was extracted from the crystal, and its pairs of molecules with intermolecular contacts smaller than 4 Å were selected. Then, the intermolecular atomic distances of each dimer were evaluated and sorted so as to provide a fingerprint for the dimer configuration. These distances were finally compared between dimers and if their RMSD fell below 10^{-4} Å, the dimers were considered identical and only one was preserved.

To characterise the configurations of the unique dimers, an orthonormal pair of principal and secondary axes was calculated for each constituent fragment, and the angles between same axes of two molecules was evaluated. To obtain the vectors, first, all atoms of the molecule were projected onto an averaged plane via singular value decomposition. The principal axis was defined as the vector tracing the longest interatomic distance and the secondary axis its perpendicular vector on the averaged plane.¹⁹³ This process is represented for a dimer of the 3P crystal in Figure 6.2.

All molecules have rotational symmetry about both of the axes when defined this way apart from HBT which has an inherent orientation. In the case of this molecule, we employed the scheme described in Section 5.3.1 where, by exploiting the two longest interatomic distances on the averaged plane, a set of axes could be defined with consistent orientation.

To evaluate exciton couplings between dimers, the diabatisation scheme by Troisi and Aragó¹⁰⁵, as implemented in *fromage*. The transition dipole moments of the isolated monomers were compared to those in the dimer to construct a diabatic Hamiltonian, whose off-diagonal elements are the exciton couplings. The original algorithm is thoroughly described in the Supporting Information of Reference 105 and in reference 193. The transition dipole moments were calculated in *Gaussian*¹³⁹ using TD- ω B97X-D/6-31G(d).

For QM:QM' calculations, the ONIOM scheme was used, using electrostatic embedding. The excited state level of theory was TDDFT ω B97X-D/6-31G(d), with

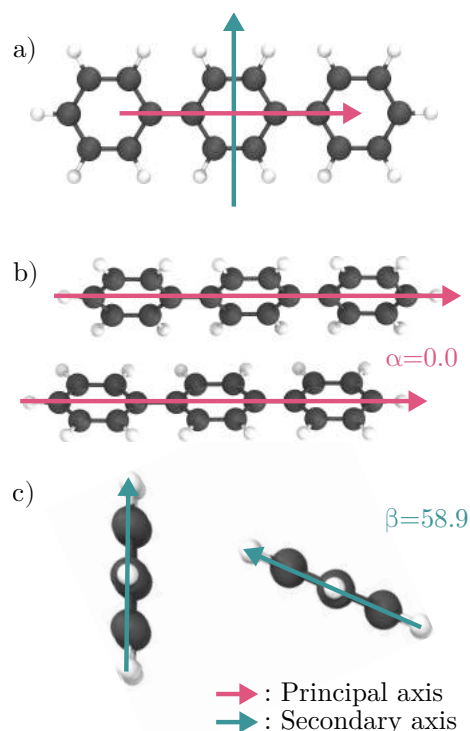


FIGURE 6.2: a) Principal and secondary axes on the 3P monomer b) Top view of a 3P dimer c) Side view of the same dimer.

Gaussian, or ADC(2)/SV(P), with Turbomole¹⁴¹. The high level region was embedded in point charges from RESP calculations of DFT ω B97X-D/6-31G(d) calculated in Gaussian. For the polar molecules HBT and α -DCS, the electrostatic embedding was extended to include long range Coulomb interactions using the ONIOM Ewald Embedded Cluster method (OEEC).⁶¹ The ground state level of theory was DFTB, with DFTB+¹⁴², and the embedding for the central region was done using RESP calculations with PBE/6-31G(d) calculated in Gaussian. Multireference SA-2-CASSCF and MS-2-CASPT2 calculations were performed with Molcas.¹⁴⁰ The active spaces are reported in the Supplementary Information.

To sample the exciton coupling of in the dimeric vibrational space of the crystal, the QM:QM' calculations were carried out on dimers to find their FC point. Then, a normal modes calculation was carried out, from which a Wigner distribution of 200 sample geometries were extracted using Newton-X.^{145,184} The exciton couplings were then evaluated using the diabatisation method.

The Huang-Rhys factors (S_i) for relaxation within the S_1 state, between S_0 and S_1 minima, were evaluated for the members of n P and DCS series in vacuum and

crystal environment using the DUSHIN code.¹⁹⁴ The computations were based on normal modes computed at the optimised S_0 and S_1 geometries in vacuum at the (TD-) ω B97XD/6-31G(d) level and in crystal at the QM:QM' level described above. Reorganisation energies (λ_i) decomposed into normal mode contributions are related to Huang-Rhys factors as follows:

$$\lambda_i = \hbar\omega_i S_i \quad (6.1)$$

The Einstein¹⁸⁴ and Strickler-Berg¹⁹⁵ (SB) relations were employed to evaluate radiative rates and lifetimes of selected fluorophores in solution and crystal. Einstein equation for spontaneous decay relates fluorescence rate (k_r) with emission energy (ΔE) and oscillator strength (f).

$$k_r = \frac{2\Delta E f}{c^3} \quad (6.2)$$

The SB relation takes into account transitions between vibronic wave functions of excited and ground states¹⁹⁵. According to the SB relation, the radiative rate can be evaluated as¹³³

$$k_r = 0.667[cm^2 \times s^{-1}] \frac{\nu_F^3}{\nu_A} n^2 f \quad (6.3)$$

Where ν_F and ν_A are vertical emission and absorption energies (in cm^{-1}), f is the oscillator strength, and n is the refractive index of a solvent.

To evaluate exciton hopping rates, the Marcus scheme was employed:

$$\nu_{ij} = \frac{J_{ij}^2}{\hbar} \sqrt{\frac{\pi}{\lambda k_B T}} \exp \left[-\frac{\lambda}{4k_B T} \right] \quad (6.4)$$

Where J_{ij} is the exciton coupling between excited monomers, \hbar is the reduced Planck's constant, k_B is Planck's constant, λ is the reorganisation energy computed as described above, and T is room temperature of 298 K.^{196,197}



FIGURE 6.3: Illustration of two archetypal packing motifs in molecular crystals.

6.3 Results

6.3.1 Dimeric Arrangement

Due to the non-covalent nature of molecular crystals, photophenomena often occur locally¹⁰⁵ and can be understood on the scale of the nearest neighbour molecules. In particular, the conformational features of the dimer arrangements resulting from the packing of the crystal can elucidate the pairwise interactions affecting the excited states in the crystal.

The packing motifs of these materials are diverse and smoothly varying, making them challenging to classify. However certain patterns have been identified to occur frequently, and we use these as reference points.^{198,199} In particular, herringbone crystals pack in an alternating edge-to-face arrangement, while sheet crystals have all molecules sharing the same orientation arranged in regular layers. These two motifs are represented in Figure 6.3. For the sake of clarity, in this chapter, we employ the terms like *herringbone* and *sheet* to refer to the overall crystal packing, and ones like *edge-to-face* or *face-to-face* to denote specific dimer arrangements.

Table 6.1 summarises the principal data relating to the packing and emission of all systems. We immediately note the prevalence of H dimers arising from the herringbone and sheet packing patterns, which result in significant QEF values, contradicting Kasha's exciton model.¹⁰⁷ The intermolecular processes in these materials therefore must go beyond point-dipole approximations.

The significant dimers of the crystals in consideration were processed in fromage to extract their characteristic angles. The overall crystal packing was found to be split between the series of molecules. The DCS series primarily forms sheets, while the other crystals have a clear bias towards herringbone motifs. The principal axis

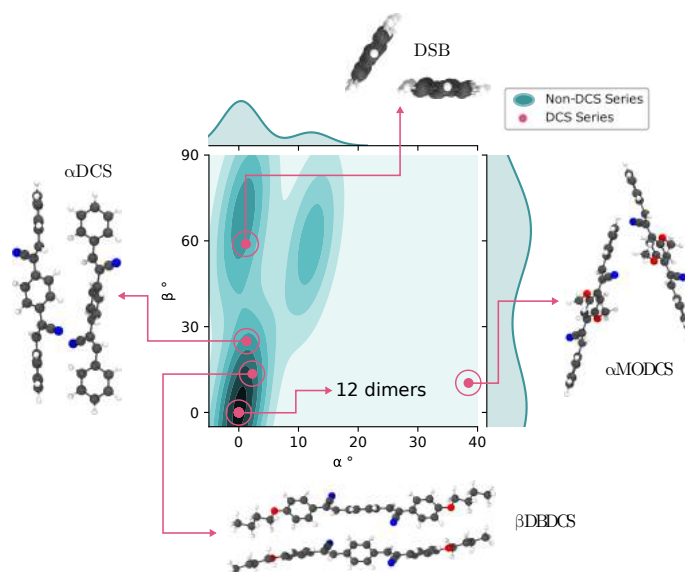


FIGURE 6.4: Distribution of the dimeric arrangements in the DCS series (14 dimers) and the rest of the crystals (14 dimers). The angle between principal axes is plotted against the angle of the secondary axes. Almost all DCS series dimers are perfectly parallel, the exceptions are shown.

angles in face-to-face dimers of sheet crystals are almost always 0° due to the translational symmetry between layers. In contrast, herringbone crystals usually have nearly parallel principal axes and a large array of secondary axis angles. Indeed the tilt between herringbone layers is strongly dependent on the morphology of the constituent fragments.

The angle values for non-DCS series crystals are represented as a density map in Figure 6.4. The herringbone and sheet families of crystals can be spotted by their characteristic densities centred around $(0^\circ, 0^\circ)$ and $(60^\circ, 15^\circ)$ respectively. Herringbone crystals also contribute to the density at $(0^\circ, 0^\circ)$, given that they are made up not only of edge-to-face dimers but also of face-to-face ones. Furthermore, there is more configurational variety in edge-to-face dimers, as evidenced by the wide spread of *beta* angles. The island of dimers of α angles larger than 10° is due only to one 4PV dimer and two HBT dimers. This slight departure from ideal herringbone packing is known in 4PV¹⁸⁹, where the unit cell includes six molecules instead of two. HBT displays slightly misaligned head-to-tail dimers indicating that each layer of the material in the principal axis direction has an alternating orientation.

The DCS series dimers are represented as a points on the same plot. All of the points were found at the $(0^\circ, 0^\circ)$ position, except for four outliers. DSB has a dimer

at $(1.2^\circ, 58.9^\circ)$, which is near the centre of the dense shaded area. This confirms the nature of DSB as a herringbone crystal, which becomes a sheet crystal only following substitutions. α -DCS and β -DBDCS have dimers at $(1.3^\circ, 24.9^\circ)$ and $(2.3^\circ, 13.5^\circ)$ respectively, which break the translational symmetry due to the freedom of rotation of their central aromatic ring. α -MODCS has a point at $(38.5^\circ, 10.3^\circ)$, which is far away from any other considered dimer. The packing of this crystal does not match any of the common packing classifications.

6.3.2 Exciton Coupling

The values of the exciton couplings of the molecules under scrutiny, evaluated at their Frank-Condon region, are a result of structural and chemical features of the dimers formed upon aggregation. We therefore wish to highlight any possible correlations between the packing patterns described above, and the exciton states within the crystal. These states are liable to cause delocalised excitation phenomena, discouraging emission.

We first examine the dependence of the couplings on the distance between constituent fragments of the dimer. Figure 6.5 shows the exciton coupling of each dimer of the crystal structures with respect to the centroid-to-centroid distance of said dimer. We observe a clear monotonic downward trend for dimers belonging to every crystal except for HBT. This trend is in line with the limiting behaviour where fragments become non interacting at infinite distances and the coupling should therefore tend to zero. In the middle to long range, the electrostatic interaction between the two fragments approaches a $1/r$ shape where r is the distance between the centres of mass of each electron cloud. Figure 6.5 does not have the sufficient resolution to suggest an inverse law as opposed to other monotonically decreasing functions. However, we can observe that dimers from different series have similar exciton coupling given similar centroid-to-centroid distance within a range of 50 meV. This is a surprising result because of the inadequacy of centroid distance as a measure of correlation of neighbouring excited states, ignoring the shape of the molecular wave-functions altogether.

HBT constitutes a striking exception, where all of the nearest neighbour dimers

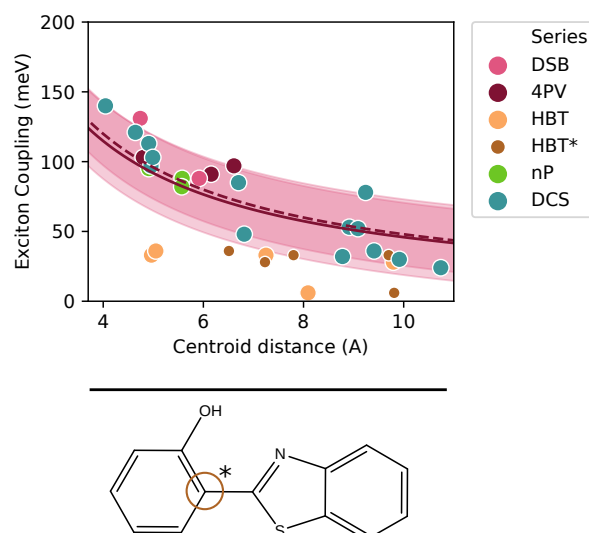


FIGURE 6.5: (top) Exciton coupling (J) as a function of the centroid-to-centroid distance of the constituent monomers of each dimer. The values of every dimer are fitted to an inverse law $f(r) = a/r$ via least squares. The resulting function, with $a = 459$ is plotted in pink and has a standard deviation of 27, represented by the shaded area. A dashed line represents the same fit but using an aromatic carbon (bottom) as a reference for the distance calculation, denoted HBT*. In this case, $a = 480$ and the standard deviation is 23.

have exciton coupling values between 20 and 40 meV. In particular, the two closest dimers, with centroid distances close to 5 Å, are about 60 meV below the fit line. The transition densities of these dimers, compared with the monomer transition density are depicted in Figure 6.6. The excitation of the isolated molecule is mainly localised in the proton transfer moiety, breaking the apparent symmetry of the two constituent rings. Both closest dimer arrangements are aligned in centroid but opposite oriented, effectively distancing the excitation densities. This effect is less pronounced in other molecules because they are all symmetric in orientation of their long axis. By measuring the distance between HBT molecules with an aromatic carbon as a reference point, the exciton coupling values adopt a clearer downward trend, albeit shifted lower than for the other series.

Excluding HBT, the dimers are roughly split into two groups, those below 7 Å in separation, with couplings ranging from 82 to 140 meV and those above 7 Å with couplings from 24 to 64 meV. Those in the former group are overall above the a/r trend line, and those of the latter below. This may be explained by the added proportion of exciton coupling resulting from exchange in the strong coupling regime. Ref.

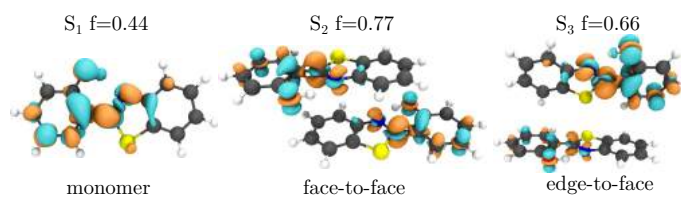


FIGURE 6.6: Transition densities of the bright states of HBT in monomer form and both dimers of least centroid-to-centroid distance. Upon excitation, the density is reorganised from the blue to the orange areas.

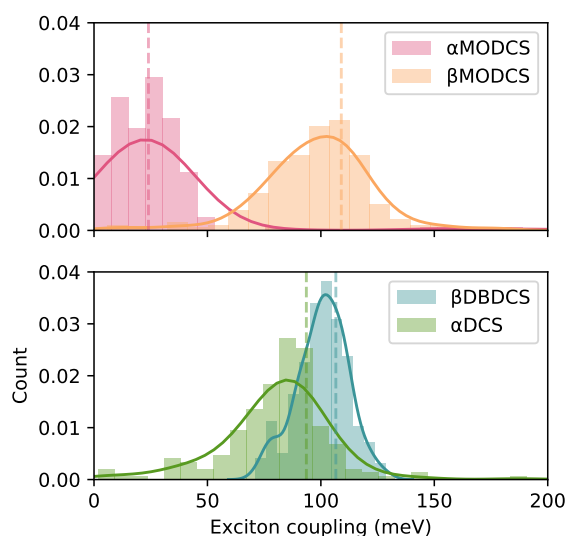


FIGURE 6.7: Exciton coupling values of irregularly packed DSB dimers, sampled from their vibrational phase space.

91 found that when Coulombic coupling exceeds 70 meV in organic semiconductor materials and light-harvesting complexes, the exchange portion of the coupling always shares a sign with its electrostatic counterpart, thus increasing the total coupling. This is consistent with the deviation of the limiting behaviour of the total coupling from a Coulombic inverse law.

Moreover, we observe a clear linear dependence of the exciton coupling on the dimeric S_2 - S_1 gap, depicted in Figure 6.8. This can be understood by definition, in the limit of linear resonant molecules.¹⁶⁶ The S_2 and S_1 states of the dimer are composed of superpositions of equal and symmetric S_1 adiabatic monomer states, and the splitting is only due to the exciton coupling. The remarkable agreement with the fit line implies a strong degree symmetry of the two constituent molecular wavefunctions, characteristic of the herringbone and sheet packing characterised in

Crystal	Series	V_i	J (meV) ^a	Packing ^b	H dimer % ^c	Calculated absorption			Exp. absorption	Φ_f
						Monomer	Dimer S_1	Dimer S_2		
DSB	-	1.41	131	HB	100	3.60	3.47	3.65	3.48 ^{d133}	0.78
4PV	-	1.42	103	HB	100	3.10	2.98	3.19	-	-
HBT	-	1.43	36	HB	83	3.98	3.94	4.00	3.65 ^{e200}	0.77 ^{s201}
3P	nP	1.37	98	HB	100	4.36	4.24	4.43	4.51 ^{d202}	0.67 ²⁰³
4P	nP	1.39	99	HB	100	4.16	4.05	4.23	4.13 ^{d202}	-
6P	nP	2.27	95	HB	100	3.75	3.63	3.82	-	0.30 ^{h204}
α -DCS	DCS	1.41	97	S	100	3.74	3.60	3.80	3.61 ^{f133}	0.90
α -DBDCS	DCS	1.49	53	S	100	3.33	3.19	3.30	3.34 ^{f133}	0.62
β -DBDCS	DCS	1.53	113	S	100	3.30	3.17	3.34	3.19 ^{f133}	0.84
α -MODCS	DCS	1.42	32	-	50	3.66	3.64	3.69	3.45 ^{e133}	0.66
β -MODCS	DCS	1.39	140	S	100	3.06	2.86	3.13	2.95 3.60 ^{f133}	0.73
α -MODBDCS	DCS	1.44	103	S	100	3.27	3.22	3.29	3.43 3.84 ^{f133}	0.42
β -MODBDCS	DCS	1.43	121	S	100	2.80	2.68	2.84	2.87 3.40 ^{f133}	0.46

TABLE 6.1: Photoactive molecular crystals considered. Φ_f : fluorescence quantum yield in crystal, V_i : steric volume index. ^a Largest dimeric exciton coupling in the crystal, ^b Herringbone (HB), Sheet (S) or other (-), ^c fraction of H (not J) dimers in the crystal, ^d tetrahydro-2-methylfuran solvent, ^e cyclohexane solvent, ^f chloroform solvent, ^g powder, ^h film.

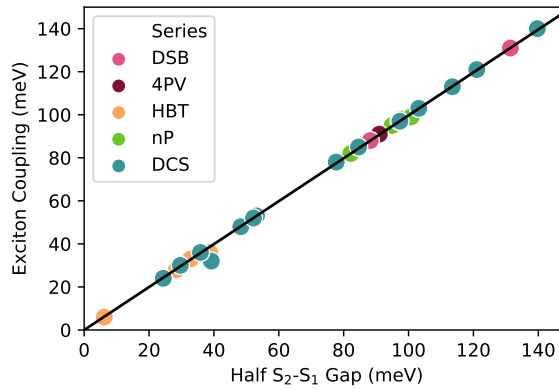


FIGURE 6.8: Exciton coupling as a function of half of the S_2-S_1 energy gap. The linear trend line, obtained *via* least squares, is $f(x) = 0.997x$

the previous section.

We would also like to probe for any link between the geometric dimer arrangements and exciton coupling values. We sample the vibrational phase space of four dimers of the DCS series: two close to sheet-like packing crystals (α -DCS and β -DBDCS), one perfect sheet crystal (β -MODCS) and the unclassified outlier α -MODCS. The results are shown in Figure 6.7.

We can observe a grouping of vibrational ground state exciton coupling values of 94, 107, and 109 meV for the face-to-face dimers (respectively α -DCS, β -DBDCS,

and β -MODCS), well away from the value of 24 meV for α -MODCS. When compared with the corresponding centroid-to-centroid distances—4.95, 4.64, 4.91, and 8.78 Å—we observe the correlation with interatomic distance mentioned previously. β -DBDCS shares a similar interatomic distance than α -DCS and β -MODCS, despite its additional butoxy chains which lengthen the backbone. We can therefore link crystals of this series with similar packing motifs, to a similar interatomic distance, and a similar Franck-Condon geometry exciton coupling.

The broadness of the peaks offers an insight on how the dimers' thermal motions can influence their exciton couplings. From narrowest to broadest, the standard deviations are 11 meV for β -DBDCS, 27 meV for β -MODCS, 35 meV for α -MODCS, and 51 meV for α -DCS. α -MODCS and β -MODCS are the two systems with closest molecular structure, and also the most similar standard deviation in exciton coupling values, despite their radical packing difference. β -DBDCS has a spread of about a third that of β -MODCS, and a fifth that of α -DCS despite them all displaying sheet-like packing. To summarise, in these systems, since exciton coupling is a many-body property, its value at fixed geometry depends on intermolecular structure properties. However the thermal fluctuation in exciton coupling depends on the vibrational phase space of the crystal dimers, which in this case is determined predominantly by molecular structural factors, not the dimer arrangement. Crystal-scale vibrations are not probed by this model, and we cannot comment on the importance of exciton-phonon effects.

6.3.3 Excited-state Relaxation by Series

Whilst geometrical considerations are easily transferable from series to series, their nonradiative decay mechanisms are usually considered to be highly system-specific since they are directly dependent on the molecular conformations available to the molecule in question. In this section, we illustrate this point by investigating the excited-state decay processes of chemically diverse molecules.

HBT

2-(2'-hydroxyphenyl)-benzothiazole (HBT) displays SLE upon its aggregation to herringbone single crystal²⁰⁵ and liquid crystal phases.²⁰⁶ While HBT has low QEF in organic solvents;^{201,207} in the aggregate phase (THF/water solution, 1:9, v/v)

and in powder, the yields are 0.09 and 0.77, respectively.²⁰¹ For this reason, HBT-derivatives are proposed as materials for organic light-emitting diodes (OLEDs) and fluorescent probes.

The underlying excited state relaxation mechanism of HBT-based systems includes excited state intramolecular proton transfer (ESIPT) in vacuum, solution,^{208,209} and crystal^{210,211}. In vacuum and solution, the process is known to be a four-step photophysical cycle enabled by an intramolecular hydrogen-bonding. It consists of a photon absorption, excited-state proton transfer, torsional motion, and the ground-state proton back-transfer^{208,211}. The process is characterised by large reorganisation energies dependent on the nature of the solvent.^{208,211}

Reorganisation energy can be used to deduce the most efficient relaxation pathways within the excited state PES. We optimised the S_0 and S_1 states of keto and enol forms of HBT in solution and in the solid state. Following the excitation to the S_1 state of the enol form, there are two possible pathways: relaxation to the S_1 minimum of the enol form, and ESIPT yielding cis-keto form in the S_1 state. The reorganisation energies in the S_1 state released during these processes are 0.28 eV for the former and 0.38 eV for the latter in cyclohexane and 0.27 eV and 0.39 eV respectively in the solid state. Thus, the ESIPT process is energetically encouraged.

Additionally, in crystal, the enol emission energy has a 0.54 eV difference with the enol absorption energy and 0.27 eV with the keto absorption energy. In contrast, the keto emission has differences of 1.27 eV and 0.56 eV respectively. This indicates reabsorption is much likelier to occur with light emitted from the enol form, thus quenching this radiative decay channel. Indeed fluorescence experiments have observed emission from both the enol and keto forms in solution²¹² and only from the latter in crystal,²¹³ where the packing is closer and the chances for reabsorption greater, only from the latter.

We now attempt to rationalise the SLE mechanism of HBT *via* the lens of the Restricted Access to Conical Intersections (RACI) model outlined by Blancafort *et al.* in References 214 and 135. This framework compares the energy of the MECI energy to the vertical absorption in order to determine the viability of internal conversion through a conical intersection. Our previous work has already proven the efficacy of this method for aromatic ESIPT materials, which indicates that other nonradiative

mechanisms can be put aside for now.^{131,174} We used similar ONIOM Ewald Embedded QM:QM' Cluster methods (OEEC), to optimise critical regions of the solid state potential energy surface. We used the Ewald embedding scheme to account for the long-range electrostatic interactions of the material, which can be important in polar crystals. We chose the ω B97X-D functional since it has reproduced accurate ESIPT optimised geometries in the past, and predicts a QM:QM' absorption energy of 4.20 eV, as compared to the experimental value of 3.65 eV. The results are depicted in Figure 6.9.

The conical intersection which involves only a rotation of the oxygenated aryl bond is the S_1 - S_0 MECI in solution, but becomes very unstable due to the steric hindrance of the nearest neighbour molecules upon crystallisation. The S_1 - S_0 MECI in crystal additionally involves the pyramidalisation of one of the molecular backbone carbons, thus reaching a distorted but spatially compatible geometry within the close packed environment. However due to this distortion, the crystal MECI remains unstable, surpassing the absorption energy by 1.9 eV. If we use MS-2-CASPT2(12,12)/aug-cc-pVDZ as the excited state method instead, using the geometries optimised in TDDFT, the results are similar, with an absorption of 3.88 eV—now only 0.23 eV above the experimental value—and a MECI 2.05 eV above absorption. In this case, the S_1 - S_0 gap at the MECI geometry is 0.36 eV. Further scanning of the CASPT2 PES would help narrow this gap, but would be unlikely to reduce the energy by up to 2.05 eV.

Upon crystallisation, the S_1 - S_0 MECI becomes inaccessible for a molecule excited at FC point. This blocks the principal nonradiative decay channel, and explains the 8 to 9 fold rise in QEF from measurements in organic solvent and in powder samples. In this case, alternative nonradiative decay channels are not important enough in the crystal to prevent the formidable 0.77 efficiency.

nP

p-Hexaphenylene (6P) has been employed as a building block of photonic nanofibers,^{215,216} and as a material for nanolasers,^{217,218} exploiting its amplified spontaneous emission,²¹⁹ thanks to its fluorescence and structural characteristics favourable for growing well-defined molecular architectures.

6P has an experimental fluorescence quantum yield of 0.85 in solution and 0.30

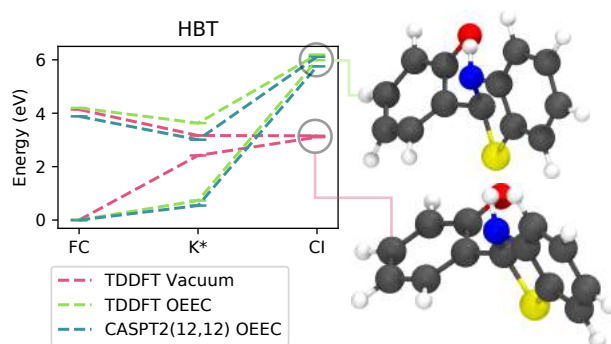


FIGURE 6.9: HBT energy at critical points of its excited state potential energy surface. The vacuum calculation used TD- ω B97X-D/6-31G(d) and the crystal calculation TD- ω B97X-D/6-31G(d):DFTB.

in crystal,²⁰⁴ whereas the much smaller 3P has a yield of 0.82 in solution²²⁰ and 0.67 in crystal.²⁰³ We would like to rationalise this difference, also considering the intermediate case of 4P. In contrast with HBT, these systems are emissive in vacuum, meaning that their excited state process in vacuum is not dominated by nonradiative deactivation.

We optimised the structures of the emitting $\pi\pi^*$ states of 3P, 4P, and 6P, applying TD- ω B97X-D in cyclohexane solvent using PCM and in the crystal phase with the QM/QM' cluster method at the TD- ω B97X-D/DFTB level including one molecule in the QM region. From the computed S_1 energies and oscillator strengths (Table 6.2), several interesting trends can be observed.

As the length of the chain increases from three to four and four to five, the emission energy decreases. The trend is similar in solution and in crystal where in the former, the emission energy decreases by 0.20 eV from 3P to 4P and by 0.17 eV from 4P to 6P, and in the latter the differences are 0.21 eV and 0.28 eV.

We can relate this phenomenon to the degree of delocalisation of the transition density in the different molecular structures. As can be seen in Figure 6.10, in the case of 3P, the S_1 transition density is mostly localised on the central phenyl ring and surrounding C-C bonds, while in the case of 4P and 6P, it is localised on two central phenyl rings and surrounding C-C bonds. This delocalisation destabilises the HOMO and stabilises the LUMO, which contributes to narrowing the S_1 - S_0 energy gap. This could also explain the increases in oscillator strength by 0.54 and 0.90

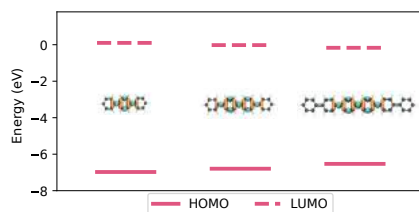


FIGURE 6.10: The energies of HOMO and LUMO orbitals of 3P, 4P, and 6P in the crystal computed at the TD- ω B97X-D/6-31G(d) level. The S_1 transition densities are represented as well.

Molecule	$E(S_1)^{\text{sol}}$	f^{sol}	$k_r^{\text{Ein,sol}}$	$k_r^{\text{SB,sol}}$	Φ_f^{sol}	λ^{vac}	$E(S_1)^{\text{cr}}$	f^{cr}	k_r^{cr}	Φ_f^{cr}	λ^{cr}
3P	3.69	1.37	0.81	1.24	0.82 ²²⁰	0.51	3.83	1.19	0.76	0.80 ^{a203} , 0.67 ^{b203}	0.35
4P	3.49	1.91	1.01	1.54	-	0.54	3.62	1.74	0.99	-	0.37
6P	3.32	2.81	1.35	2.17	-	0.49	3.34	2.86	1.38	0.30 ^d	0.29

TABLE 6.2: Computed S_1 energies in eV ($E(S_1)$) with corresponding oscillator strength (f), radiative rates calculated with the Einstein relation and SB formula in ns^{-1} (k_r^{Ein} , k_r^{SB}), experimental luminescent efficiency (Φ_f), and reorganisation energies in eV (λ) for the n P series. The superscripts indicate the medium, where "sol" is cyclohexane solvent, "vac" is vacuum, and "cr" is crystal. ^aPowder samples. ^b Single crystal.

in solution, and 0.55 and 1.12 in crystal, where a more diffuse transition is correlated with a greater overlap between initial and final wavefunctions and a greater transition dipole moment.

In comparison with the solution, the crystal environment raises the emission energy by 0.1 eV and lowers the oscillator strength by 0.2 for 3P and 4P. These effects are, however, negligible for 6P.

The emission rates, k_r , computed applying Einstein relation in solution and crystal and applying the SB relation in solution, increase with backbone length. This is due to the large increase of oscillator strength relative to the decrease in emission energy. The k_r values in solution and in crystal are very similar throughout the series due to competing effects of crystallisation increasing the emission energy and decreasing the oscillator strength for 3P and 4P.

The rates in solution obtained based on the SB relation are about twice as large as the values obtained from Einstein relation. This is due to the transitions between vibronic wave functions of the excited and ground states, which the Einstein relation neglects.

The decrease in QEF of 6P with respect to 3P is not explained by the behaviour of radiative rates which instead increase with chain length. This raises the question of the importance of nonradiative relaxation pathways in this series. We first examine a rationalisation based on conical intersections, as this was a determining factor for HBT.

The S_1 - S_0 minimum energy crossing points were optimised at the ADC(2)/def-SV(P) level in vacuum and crystal for 3P, 4P, and 6P. TD- ω B97X-D/6-31G(d) was also attempted but electronic convergence problems arose due to the highly distorted conformations involved. Previous studies indicate that ADC(2) can represent accurate S_1 - S_0 crossing topologies in organic chromophores despite being a single reference method.⁷⁶

The optimised vacuum S_1 - S_0 MECI geometries of 3P and 4P, represented in Figure 6.11, correspond to ring puckering conical intersections with puckered phenyl rings on which the S_1 transition densities are localised. The central phenyl ring at the MECI geometry of 3P in vacuum is a prefulvene kind of conical intersection,²²¹ characterised by a half-boat structure with the C_s symmetry. The puckering of the central ring is accompanied by flapping motion of peripheral phenyl rings, resulting in a highly distorted structure with one phenyl ring roughly perpendicular to the puckered ring.

However, in the crystal, ring-puckering and flapping motions are partially hindered due to the tight packing. As a result, the crystal S_1 - S_0 MECI geometry has a different identity, featuring a pronounced puckering of one C atom of the central ring and substantial out-of-plane distortion of H atom attached to it. The rest of the molecule remains in plane. This conical intersection corresponds to another point at the prefulvene CI seam.

Similarly, the S_1 - S_0 MECI structure of 4P in vacuum corresponds to a puckered half-boat structure of one of the central rings, while the other one, on which transition density is also localised at the S_1 minimum, displays slight out-of-plane distortion.

The 4P S_1 - S_0 MECI structure in crystal phase is similar to the one obtained for 3P. The restriction of large out-of-plane motions can be explained by the herringbone packing of these crystals. Figure 6.4 shows how the two prevalent packing

motifs—herringbone and sheet—produce crystalline dimers with roughly parallel principal axes. This type of steric hindrance discourages any motion which would make the backbone of the molecule deviate from the overall principal axis and enter the ground state configurational space of its nearest neighbours.

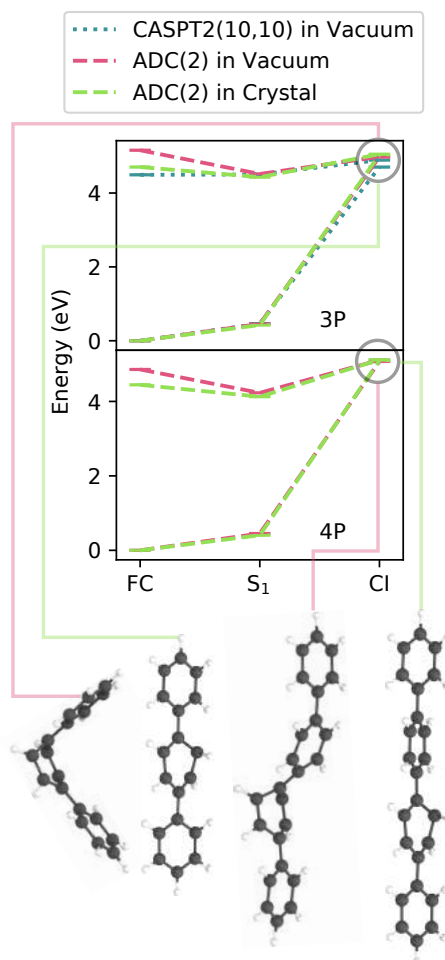


FIGURE 6.11: Energies of S_0 and S_1 states of 3P and 4P at the FC point, S_1 minima, and S_1 - S_0 MECI in vacuum (left) and crystal (right) computed at the RI-ADC(2)/def-SV(P) level. For the comparison, in the case of 3P, the corresponding CASPT2(10,10)/6-31G(d)//CASSCF(10,10)/6-31G(d) and CASPT2(8,8)/6-31G(d)//CASSCF(8,8)/6-31G(d) results are given

For both 3P and 4P, as shown in Figure 6.11, the optimised MECI geometries lie above the S_1 excitations at the FC point, both in vacuum and crystal, implying that this kind of internal conversion is inefficient for these systems. The energy of the 3P vacuum conical intersection, obtained with CASPT2(10,10)/6-31G(d)//CASSCF(10,10)/6-31G(d), lies 0.30 eV above the S_1 excitation at the FC point. ADC(2) successfully describes the region of the conical intersection of 3P

and predicts the MECI energy close to the value obtained with CASPT2(10,10)/6-31G(d)//CASSCF(10,10)/6-31G(d), but it overestimates the vertical excitation at the FC region. In the case of 4P, the vacuum MECI optimised at the ADC(2)/def-SV(P) level is 0.24 eV above the S_1 state at the FC region.

For both systems, the optimised MECI structure in crystal is more energetic compared to the one in vacuum. The MECIs of 3P and 4P lie 0.34 eV and 0.67 eV above the excitation in the FC region, based on the ADC(2)/def-SV(P) optimisations.

The ADC(2)/def-SV(P) optimised MECI of 6P in vacuum corresponds to a puckering conical intersection with a prefulvene-like structure of the central ring. The rest of the chain is highly distorted due to rotation of terminal phenyl rings, as shown in Figure 6.12.

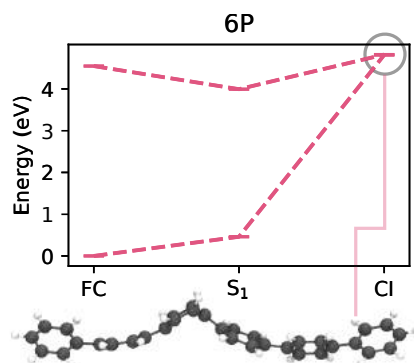


FIGURE 6.12: Energies of S_0 and S_1 states of 6P at the FC point, S_1 minima, and S_1 - S_0 MECI in vacuum computed at the RI-ADC(2)/def-SV(P) level. The MECI structure is represented in the bottom.

The S_1 state at the optimised vacuum S_1 - S_0 MECI is 0.18 eV above the vertical excitation at the FC point, implying that internal conversion is unfavourable for this molecule, as shown in Figure 6.12. The conical intersection optimisation in the crystal could only minimise the S_1 - S_0 gap to 0.3 eV, with an energy inversion between S_1 and S_0 . This shows the inadequacy of single reference methods to characterise conical intersections in the case of 6P, which is also too large for computationally affordable and accurate multireference methods. However, multireference methods confirmed the accuracy of ADC(2) for 3P, so if we assume this to apply to 6P, then we observe a MECI energy several eV above the FC energy, making it inaccessible.

Conical intersections are therefore rightly found to be at least partially inaccessible for 3P and 4P, blocking this nonradiative decay channel. However this barrier is even higher in 6P, which has lower QEF than 3P, indicating the importance of alternative nonradiative decay mechanisms for this crystal.

Another explanation for the increased nonradiative decay rate of 6P due to internal conversion would be a vibrational nonradiative mechanism. This would be consistent with the lower S_1 energy of 6P, thus enabling large vibrational wavefunction overlap. However, the low vibrational reorganisation energies of 3P and 6P are of the same order in solution and in crystal. The Supporting Information shows a reduction of these energies to 10% upon crystallisation for both systems, which does not significantly impact the emissivity of 3P.

Stampfl *et al.* proposed that the decrease of quantum efficiency upon aggregation in 6P is induced by intermolecular excitonic phenomena.²⁰⁴ This conclusion is based on a linearly decreasing luminescence efficiency with temperature, instead of an Arrhenius-type dependence. The former is explained by increased excitonic collision probabilities on higher temperatures, whereas the latter would be associated with intramolecular vibrational radiationless deactivation.

Exciton hopping rates are quadratically dependent on the exciton coupling within a crystal, and inverse exponentially dependent on the reorganisation energy, as shown in Equation 6.4. The exciton coupling in 6P crystals of 95 meV reported in Table 6.1 is of the same order as for 3P, with 98 meV. In contrast, the total reorganisation energy of 6P in crystal is 0.29 eV, compared to 3P's 0.35 eV, as can be reported in Table 6.2. This difference results in a hopping rate 1.95 greater in 6P than in 3P, a ratio approaching that of the different QEFs in crystal. We can postulate that an increased hopping rate leads to more nonradiative decay due to the mobile exciton more readily reaching surface, grain boundary, or bulk defects in the material.

In summary, vibrational and nonadiabatic nonradiative decay channels are mostly blocked in both vacuum and crystal for all members of this series. The drop in QEF of 6P upon crystallisation is due to its particular property of displaying strong exciton coupling, whilst having a low reorganisation energy compared to 3P. We can explain this property by the relative sparsity of 6P, with a steric volume index of 2.27 compared to 3P's 1.37, allowing for less strained reorganisation, but

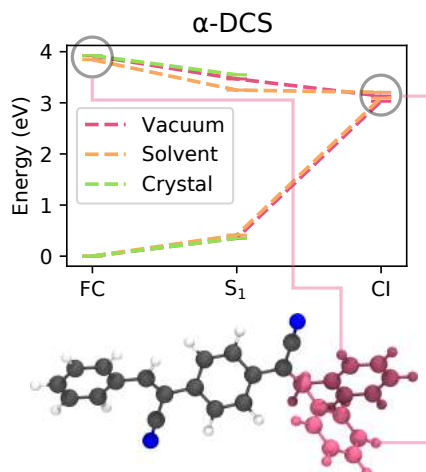


FIGURE 6.13: α -DCS energy at critical points of its excited state potential energy surface. The vacuum and solvent calculation used TD- ω B97X-D/6-31G(d) and the crystal calculation TD- ω B97X-D/6-31G(d):DFTB.

maintaining the high exciton coupling thanks to the length of the 3P backbone maximising π - π interactions.

DCS Series

Finally, we investigate the excited state decay channel of another aromatic molecule with a different structured backbone, based on the DSB molecule. α -DCS is a DSB derivative displaying an impressive rise of QEF from 0.002 to 0.90 from solution to single crystal.¹³³

Its geometry was optimised in chloroform solvent using PCM to find its ground and excited state minima and conical intersection geometry. The results are shown on Figure 6.13. The absorption energy was 3.84 eV, in close agreement with the experimental value of 3.80 eV. The FC minimum was characterised by a tilt of the inner ring with respect to the outer rings of 67.2°, whereas the S_1 optimisation led to a more planar geometry with an angle of 21.3°. This large reorganisation led to an emission energy of 2.84 eV, shifted 1.01 eV away from the absorption energy.

The molecule was reoptimised in its crystal phase using OEEC. Here, the FC geometry had a tilt of 62.9° but the planarisation of S_1 was significantly hindered, only reaching 47.1°. We observe that the crystal packing reduces the flexibility of the molecule, which is calculated to absorb at 3.93 eV and emit at 3.20 eV, producing a Stokes shift of 0.73, a value smaller than in solution by 0.3 eV.

The increased rigidity also has implications for the geometry of the S_1 - S_0 MECI.

The access to the conical intersection geometry in solvent for similar molecules has been characterised by a rotation about a double bond of the backbone, causing one ring to be on a perpendicular plane from the other two, and a pyramidalisation of the carbon connecting the rotated ring to the backbone.²²² We located a similar conical intersection for α -DCS, where the rotation was of 88.8° about the same bond as reported in Reference 222, but involving no pyramidalisation. In either case, the rotation involved in the access to this S_1 - S_0 MECI supposes a large reorganisation which represent a quenched nonradiative decay channel in solution, and a blocked one in crystal. Indeed the crystal packing is too dense to allow for the backbone to draw an arc of nearly a right angle, instead the penalty function MECI optimisation algorithm pursues a double bond stretching CI too distorted to evaluate even with multireference methods.

Other emissive DSB-based molecules have similar molecular backbones and packing, as shown in Figure 6.4, pointing to a similar quenching of the internal conversion decay pathway. They have been investigated in a series of studies for their promising SSL properties.^{133,222,223} They all have cyano-group (CN) substituents on the vinylene units which connect their phenylene rings. DBDCS and MODBDCS have additionally butoxy-groups (OBu) on lateral phenylene rings at their para positions. Apart from CN- and OBU-substituents, the MODBDCS molecules are distinguished by methoxy-groups (OMe) on central phenylene rings in their meta positions. MODCS is characterised by OMe-substitutions on central phenylene rings. α - and β -members of the series differ from each other by positions of CN- groups on vinylene units with respect to the phenylene rings; the former having CN- groups closer to the central ring, and the latter closer to lateral phenylene rings.

All members of the series are emissive in the crystal form, with higher efficiency than in solution. In particular, the α - systems are completely non-emissive in solution. As noted in the previous sections, changes in QEF are understood as a competing change in radiative and non-radiative rates, with the latter being contributed to by differing vibrational wavefunction overlap, intermolecular excitonic processes, and conical intersection accessibilities.

The nonradiative rates only increase upon aggregation for α -DCS and α -MODCS, as is reported in in Table 6.3.¹³³ Therefore, we can expect an important restriction of

Molecule	$k_r^{\text{sol}133}$	$k_{nr}^{\text{sol}133}$	$\Phi_f^{\text{sol}133}$	$\lambda_{\text{low}}^{\text{vac}}$	$k_r^{\text{cr}133}$	$k_{nr}^{\text{cr}133}$	$\Phi_f^{\text{cr}133}$	$\lambda_{\text{low}}^{\text{cr}}$
α -DCS	0.35	175	0.002	0.095	0.43	0.05	0.90	0.046
α -DBDCS	0.5	250	0.002	0.088	0.05	0.02	0.70	0.016
β -DBDCS	0.45	0.39	0.54	0.090	0.14	0.03	0.84	0.035
α -MODCS	0.11	5.4	0.02	0.094	0.19	0.1	0.66	0.054
β -MODCS	0.15	0.60	0.2	0.089	0.04	0.02	0.73	0.031
α -MODBDCS	0.23	77	0.003	0.059	0.09	0.12	0.42	0.009
β -MODBDCS	0.22	0.50	0.31	0.077	0.02	0.02	0.46	0.009

TABLE 6.3: Experimental radiative rates in ns^{-1} (k_r), nonradiative rates in ns^{-1} (k_{nr}), luminescent efficiency (Φ_f), and computed reorganisation energies for vibrational modes of less than 0.031 eV (250 cm^{-1}) in eV (λ_{low}) for the DCS series. The superscripts indicate the medium, where "sol" is chloroform solvent, "vac" is vacuum, and "cr" is single crystal

nonradiative decay mechanisms upon crystallisation for the series.

Regarding vibrational nonradiative decay, it has been proposed to contribute to SLE within the Restriction of Intramolecular Motions (RIM) model described in References 224 and 225. The crystallisation is said to quench low energy vibrational modes, thus impeding the overlap of vibrational wavefunctions between different excited states, and blocking Fermi-Golden-Rule-style nonradiative decay.

As is shown in Table 6.3 these modes are indeed reduced in the crystal phase for the DCS series, however less so than in 3P and 6P, whose nonradiatively decay is thought to principally be through conical intersections and excitonic dissipation. No clear trends emerge linking the quenching of vibrational modes upon crystallisation to the change in QEF of the systems, though they cannot be excluded as a contributing factor to the enhanced emission.

To probe for excitonic dissipation, we can focus on the case of the strongest coupled system, β -MODCS, as seen on Table 6.1, with an exciton coupling of 140 meV. This does not impede a very efficient emission of 0.73, despite its low radiative rate reported in Table 6.3. As shown in Figure 6.4, most crystals share the characteristic face-to-face dimer packing of sheet-like crystals, indicating that the character of their excitonic states should not be radically different. The principal exceptions are α -DCS and α -MODCS, where the former still displays the greatest crystal luminescent efficiency of the series, and the latter a mere 32 meV of exciton coupling. We can conclude that within this series, excitonic states are either present but not of a

dissipative character, or absent.

As for the access to conical intersections, there is reported evidence for its importance in the series. Reference 223 observes a rise in nonradiative decay rates with increasing FC energy in solution. This indicates the comparatively low importance of vibrational decay in the nonradiative rate of this family, due to a lesser overlap between ground and excited state vibrational wavefunctions in the low nonadiabatic coupling regime; leaving the RACI mechanism such as the one previously outlined for α -DCS as the principal cause of SLE.¹⁰⁴

The dominance of conical intersection decay in this series can also be linked to the low but present luminescence of the β - molecules in solution. The position of the CN- substituents upon the rotating section of the double bond which drives the access to the conical intersection, at least in α -DCS, can constitute a hindrance to the motion.

Moreover, the RACI model, depends on the rigidity of the molecules in the excited state, where a smaller conformational freedom of the molecule results in fewer pathways to the conical intersection to restrict. This is consistent with the overall greatest QEF, attributed to the smallest molecule— α -DCS, with 0.90—and the lowest QEF to the ones with the most substituent— α -MODBDCS and β -MODBDCS, with 0.42 and 0.46 respectively.

6.4 Conclusions

In this chapter, we have reviewed the crystalline excited state properties of thirteen organic molecular crystals with luminescent behaviour. We used a geometry analysis tool to characterise the different nearest neighbour dimers present in these crystals and associate them to particular crystal packing motifs. We observed that within one series of similar molecules—DSB derivatives—the packing motif determined the centroid-to-centroid distance of the resulting molecules.

This distance was shown to have direct implications as to the value of the exciton coupling between constituent fragments. Chemically different molecules were found to have similar exciton coupling values within a range of 50 meV. The weakness of this model was highlighted in the HBT crystal, where the centroid of the

molecules is far away from the area with most electronic reorganisation upon excitation.

We also observed that the exciton coupling values accessible to dimers in their vibrational phase space are not particularly dependent on the intermolecular arrangement. Indeed the vibrations are mostly determined by the molecular structure, and dictate the broadness of values of the exciton coupling.

We then investigated the internal conversion decay channels for five of the crystals and how they were affected by their crystal environment. For these luminescent materials, the conical intersection energy was systematically found to be higher than the absorption energy, in crystal, pointing to a quenching due to steric hindrance. Molecules with rotation involved in their vacuum phase conical intersection were likelier to have a higher energy crystal conical intersection. In contrast, ones which had puckered geometries in vacuum found alternate puckering patterns in the crystal with close to equivalent energy profiles. Clear links between the access to the conical intersection and the QEF were drawn for HBT and DCS systems. Vibrational decay was not directly shown to have a significant role in the darkness of any systems in the vacuum, and was further discouraged in the crystal by the quenching of low energy normal modes. Finally, excitonic dissipation was proposed as a determining mechanism explaining the different QEFs of the nP series.

Overall, excited state decay mechanisms remain relatively system specific due to the formidable breadth of chemical space. Fluorescence, internal conversion, and excitonic dissipation are competing mechanisms, interlinked by their relation to crystal structure. The complexity of this relationship is exemplified by the diverse luminescent behaviour in solution of the molecules in this study, despite their consistent efficient luminescence in as crystals. Programs like `fromage` prove themselves to be essential in disentangling the holistic mechanisms behind such phenomena.

Conclusions

Summary

This thesis offered a detailed account of the development, implementation, and application of a new suite of tools for the study of excited states in molecular crystals. Initially, we were interested in applying the self-consistent Ewald embedding scheme of Wilbraham *et. al*¹⁷ to our own set of proton transfer crystals, 2'-hydroxy-chalcones, with a focus on the SLE mechanism. It quickly surfaced that due to the comparative flexibility of our compounds, a simple point charge embedding scheme would need to be complemented by short range interactions, in order to allow for a realistic exploration of the PES.

This led us to the development of the ONIOM Ewald Embedded Cluster models described in Chapter 4. Therein, we used a hierarchy of models to recover the correct emissive behaviour of two crystals, HC1 and HC2, only the former of which displayed Solid State Luminescent Enhancement. Our new and most accurate method combined ONIOM QM:QM' with Ewald point charge embedding, in order to eliminate the electrostatic truncation error that accompanies traditional ONIOM QM:QM' in periodic systems. Not only were we able to correctly describe non-Born-Oppenheimer regions of the PES in order to describe Solid State Luminescent Enhancement in these materials using the Restricted Access to Conical Intersections model, we also found emission energies within 0.2 eV of the experimentally observed ones, a result overestimated both by QM:MM and traditional QM:QM'.

The implementation of a hybrid level of theory method into a program immediately hinted to the development of a programming library, since allowing the user a great degree of flexibility was paramount. This paved the way for the development of `fromage`, described in Chapter 5. Along with these ONIOM methods, many other tools were developed by our group since 2016 to aid in the study of molecular crystal

excited states, and they were all implemented in `fromage`. Working with molecules, finite cluster models, and unit cells all at once started off as a very time intensive task from a practical point of view, because of the niche geometry manipulation operations which were ubiquitous as part of the workflow, and the decentralisation of post-processing analysis tools. `fromage` has streamlined these processes significantly by translating these geometrical objects into Python. This allowed for the development of tools to analyse characteristic dimer geometries within aggregates and manipulate different classes of clusters. It was also used to investigate excited state electronic structures, allowing for the classification of excitonic states, and the evaluation their exciton coupling.

With this arsenal of tools at hand, we broadened our scope, to investigate larger families of compounds. Chapter 6 offered a study of thirteen emissive molecular crystals, with different luminescent properties in solution. The competition between radiative and nonradiative decay pathways was explored, highlighting how both vibrational decay and internal conversion *via* conical intersection needed to be quenched in the crystal, without introducing additional excitonic dissipation mechanisms. In our systems, low frequency normal modes were significantly quenched by crystallisation, disfavoured vibrational decay, regardless of its importance in vacuum. The conical intersection was systematically found to be level with or higher in energy than the absorption energy in crystal, thus blocking this channel by the means of steric hindrance. Conical intersections tended to involve puckering in the crystal, which was found to destabilise them more than rotational conical intersections. Energy dissipation through excitonic delocalisation was controlled for by comparing both exciton coupling values and reorganisation energies between different dimers in the crystal.

Of the objectives outlined in the introduction, the first three (extending system-environment interactions, identifying an optimal point charge scheme, probing the response of the environment) are addressed by Chapter 4. The fourth objective, of producing a tailored implementation of these methods to molecular crystals is addressed by Chapter 5. The role of Chapter 6 is to demonstrate the use and robustness of the outcomes of the previous two chapters.

This doctoral project has therefore produced all of the key steps to enable the

modelling of photochemical processes in molecular crystals within a level of accuracy which would previously have been impossible for certain systems. The `fromage` package is already in use on at least three continents, and will unlock new areas of study for future researchers. Readers can find the source code and documentation here:

<https://github.com/Crespo-Otero-group/fromage>

<https://fromage.readthedocs.io>

Outlook

Point Charge Description

The point charge approximation for the potential emanating from an atom, upon which hinges the ONIOM QM:QM' formalisms described herein, has shown to reach its limits when combined with too diffuse densities. This limits the amount of systems for which these methods could be applicable, as well as the size of the basis set.

We can fathom a whole family of substitutes for the point charges of the molecules, each with its own advantages. A first solution would be to introduce a local repulsive term to the Coulomb interaction of each point charge, accounting for the Pauli repulsion of the atom. This would require parameterisation, but keep the complexity of the embedding scheme to a minimum. Alternatively, smoothing out the point charge potential as Gaussian functions, with a finite cusp, would limit unphysical solutions due to infinite potential wells, but would not account for atomic repulsion.

Embedding the excited state Hamiltonian higher orders of the multipole expansion would further help recover mutual polarisation more naturally than in the self-consistent scheme, whilst keeping the modelling cost to a minimum. Here, the implementation aspect becomes arduous due to the scarcity of electronic structure codes accepting this sort of embedding. Additionally, determining the value of the multipoles representing the environment is not trivial. This polarisation could also be achieved by combining the point charges with a polarisable continuum model,²²⁶ which would involve parameterising the dielectric constant of the material.

Finally, the interaction between system and environment could be treated in a Rayleigh-Shrödinger perturbative way, accounting for accurate Coulomb interactions, polarisation, dispersion, and exchange provided a symmetry-adapted formulation.^{227,228} The implementation of a Symmetry-Adapted Perturbation Theory to excited states with ONIOM has yet to be fully developed.

Geometry Optimisation

Currently, the implementation of ONIOM schemes only accounts for the optimisation of region **1** of the crystal. Optimising region **2** would allow for the modelling of many-body reorganisation within the crystal, which could be critical in some systems.

Allowing for the optimisation of region **2** would require a careful consideration of the conservation of energy in the ONIOM equation, and importantly an Ewald embedding for this region.

Additionally, the current method for optimising conical intersections is the penalty function method. This has the advantage of forgoing the need for derivative and nonadiabatic coupling vectors. However those quantities are available in several packages already interfaced with `fromage`, meaning that they could be exploited to locate conical intersections faster and more precisely.

Ewald Potential

Currently, the Madelung sum of the crystal is accounted for by varying point charge values to reproduce an Ewald potential. This is all handled by the program `Ewald`,^{113,114} which was initially devised for ionic systems with a quantum region much smaller than the clusters which are relevant to molecular crystals.

A new implementation of this Ewald scheme would benefit the program hugely, by being tailored to these cluster geometries, and removing implementation problems which arise at large length scales. Other schemes also exist to mimic the result of an Ewald sum which may require fewer point charges or computational power to match,²²⁹ such as the Wolf sum.²³⁰

Bibliography

- [1] Y. Li, L. Xu and B. Su, *Chem. Commun.*, 2012, **48**, 4109–4111.
- [2] R. T. K. Kwok, C. W. T. Leung, W. Y. Lam and T. B. Z, *Chem. Soc. Rev.*, 2015, **44**, 4228–4238.
- [3] J. Huang, N. Sun, Y. Dong, R. Tang, P. Lu, P. Cai, Q. Li, D. Ma, J. Qin and Z. Li, *Adv. Funct. Mater.*, 2013, **23**, 2329–2337.
- [4] Z. Zhao, Z. Li, J. W. Y. Lam, J.-L. Maldonado, G. Ramos-Ortiz, Y. Liu, W. Yuan, J. Xu, Q. Miao and B. Z. Tang, *Chem. Commun.*, 2011, **47**, 6924.
- [5] M. P. Aldred, G.-F. Zhang, C. Li, G. Chen, T. Chen and M.-Q. Zhu, *J. Mater. Chem. C*, 2013, **1**, 6709.
- [6] J. Gierschner, S. Varghese and S. Y. Park, *Adv. Opt. Mater.*, 2016, **4**, 348–364.
- [7] A. W. Hains, Z. Liang, M. A. Woodhouse and B. A. Gregg, *Chemical Reviews*, 2010, **110**, 6689–6735.
- [8] S. B. Jo, H. H. Kim, H. Lee, B. Kang, S. Lee, M. Sim, M. Kim, W. H. Lee and K. Cho, *ACS Nano*, 2015, **9**, 8206–8219.
- [9] S. Humbel, S. Sieber and K. Morokuma, *J. Chem. Phys.*, 1996, **105**, 1959–1967.
- [10] M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber and K. Morokuma, *J. Phys. Chem.*, 1996, **100**, 19357–19363.
- [11] M. A. Kochman, A. Bil and C. A. Morrison, *Phys. Chem. Chem. Phys.*, 2013, **15**, 10803–10806.
- [12] M. A. Kochman and C. A. Morrison, *J. Chem. Theory Comput.*, 2013, **9**, 1182–1192.

- [13] D. Presti, F. Labat, A. Pedone, M. J. Frisch, H. P. Hratchian, I. Ciofini, M. C. Menziani and C. Adamo, *J. Chem. Theory Comput.*, 2014, **10**, 5577–5585.
- [14] D. Presti, F. Labat, A. Pedone, M. J. Frisch, H. P. Hratchian, I. Ciofini, M. Cristina Menziani and C. Adamo, *J. Comput. Chem.*, 2016, **37**, 861–870.
- [15] D. Presti, A. Pedone, I. Ciofini, F. Labat, M. C. Menziani and C. Adamo, *Theor. Chem. Acc.*, 2016, **135**, 86.
- [16] D. Presti, L. Wilbraham, C. Targa, F. Labat, A. Pedone, M. C. Menziani, I. Ciofini and C. Adamo, *J. Phys. Chem. C*, 2017, **121**, 5747–5752.
- [17] L. Wilbraham, C. Adamo, F. Labat and I. Ciofini, *J. Chem. Theory Comput.*, 2016, **12**, 3316–3324.
- [18] D. R. Bowler and T. Miyazaki, *Rep. Prog. Phys.*, 2012, **75**, 036503.
- [19] A. Szabo and N. S. Ostlund, *Modern quantum chemistry: introduction to advanced electronic structure theory*, Courier Corporation, 2012.
- [20] C. Ullrich, *Time-Dependent Density-Functional Theory: Concepts and Applications (Oxford Graduate Texts)*, Oxford University Press, 2012.
- [21] L. Kantorovich, *Quantum Theory of the Solid State: An Introduction*, Springer Netherlands, Dordrecht, , 2004.
- [22] E. Schrödinger, *Ann. Phys.*, 1926, **384**, 361–376.
- [23] M. Born and R. Oppenheimer, *Ann. Phys.*, 1927, **389**, 457–484.
- [24] J. C. Slater, *Phys. Rev.*, 1929, **34**, 1293–1322.
- [25] D. R. Hartree and W. Hartree, *Proc. Royal Soc. Lond. A*, 1935, **150**, 9–33.
- [26] V. Fock, *Z. Phys.*, 1930, **61**, 126–148.
- [27] V. Fock, *Z. Phys.*, 1930, **62**, 795–805.
- [28] C. Møller and M. S. Plesset, *Phys. Rev.*, 1934, **46**, 618–622.
- [29] J. Olsen, P. Jørgensen, T. Helgaker and O. Christiansen, *J. Chem. Phys.*, 2000, **112**, 9736–9748.

- [30] J. Čížek, *J. Chem. Phys.*, 1966, **45**, 4256–4266.
- [31] D. Feller and D. A. Dixon, *J. Chem. Phys.*, 2001, **115**, 3484–3496.
- [32] J. Almlöf, A. V. Nemukhin and A. Heiberg, *Int. J. Quantum Chem.*, 1981, **20**, 655–667.
- [33] B. O. Roos, P. R. Taylor and P. E. Sigbahn, *Chem. Phys.*, 1980, **48**, 157–173.
- [34] P. E. Siegbahn, J. Almlöf, A. Heiberg and B. O. Roos, *J. Chem. Phys.*, 1981, **74**, 2384–2396.
- [35] B. O. Roos, in *AB Initio Methods in Quantum Chemistry - II*, Wiley, 2007, vol. 69, pp. 399–445.
- [36] K. Andersson, P. Å. Malmqvist, B. O. Roos, A. J. Sadlej and K. Wolinski, *J. Phys. Chem.*, 1990, **94**, 5483–5488.
- [37] K. Andersson, P. Å. Malmqvist and B. O. Roos, *J. Chem. Phys.*, 1992, **96**, 1218–1226.
- [38] J. J. McDouall, K. Peasley and M. A. Robb, *Chem. Phys. Lett.*, 1988, **148**, 183–189.
- [39] P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864–B871.
- [40] W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133–A1138.
- [41] B. Tanatar and D. M. Ceperley, *Phys. Rev. B*, 1989, **39**, 5005–5016.
- [42] G.-X. Zhang, A. M. Reilly, A. Tkatchenko and M. Scheffler, *New J. Phys.*, 2018, **20**, 063020.
- [43] J. P. Perdew, K. Burke, M. Ernzerhof, D. of Physics and N. O. L. . J. Quantum Theory Group Tulane University, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.
- [44] D. Nazarian, P. Ganesh and D. S. Sholl, *J. Mat. Chem. A*, 2015, **3**, 22432–22440.
- [45] A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100.
- [46] Lee, Yang and Parr, *Phys. Rev. B*, 1988, **37**, 785–789.
- [47] C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158–6170.

- [48] J. Heyd, G. E. Scuseria and M. Ernzerhof, *J. Chem. Phys.*, 2003, **118**, 8207–8215.
- [49] J. Heyd, G. E. Scuseria and M. Ernzerhof, *J. Chem. Phys.*, 2006, **124**, 219906.
- [50] P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623–11627.
- [51] I. Y. Zhang, J. Wu and X. Xu, *ChemComm*, 2010, **46**, 3057–3070.
- [52] J.-D. Chai and M. Head-Gordon, *J. Chem. Phys.*, 2008, **128**, 084106.
- [53] A. D. Becke, *J. Chem. Phys.*, 1997, **107**, 8554–8560.
- [54] Y. Shao, Y. Mei, D. Sundholm and V. R. I. Kaila, *J. Chem. Theory Comput.*, 2020, **16**, 587–600.
- [55] S. Grimme, *J. Comput. Chem.*, 2006, **27**, 1787–1799.
- [56] W. M. C. Foulkes and R. Haydock, *Phys. Rev. B*, 1989, **39**, 12520–12536.
- [57] G. Seifert and J. O. Joswig, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2012, **2**, 456–465.
- [58] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai and G. Seifert, *Phys. Rev. B*, 1998, **58**, 7260–7268.
- [59] D. Selli, G. Fazio and C. Di Valentin, *J. Chem. Phys.*, 2017, **147**, 164701.
- [60] H. J. Monkhorst and J. D. Pack, *Phys. Rev. B*, 1976, **13**, 5188–5192.
- [61] M. Rivera, M. Dommett and R. Crespo-Otero, *J. Chem. Theory Comput.*, 2019, **15**, 2504–2516.
- [62] C. Kittel, *Introduction to Solid State Physics*, John Wiley & Sons, Inc., New York, 2005, vol. 8th ed.
- [63] J. Weber and J. Schmedt auf der Günne, *Phys. Chem. Chem. Phys.*, 2010, **12**, 583–603.
- [64] J. Torras, S. Bromley, O. Bertran and F. Illas, *Chem. Phys. Lett.*, 2008, **457**, 154–158.

- [65] K. P. McKenna, P. V. Sushko and A. L. Shluger, *J. Am. Chem. Soc.*, 2007, **129**, 8600–8608.
- [66] A. Y. Toukmaji and J. A. Board, *Comput. Phys. Commun.*, 1996, **95**, 73–92.
- [67] J. Sauer, *Chem. Rev.*, 1989, **89**, 199–255.
- [68] M. E. Casida, C. Jamorski, K. C. Casida and D. R. Salahub, *J. Chem. Phys.*, 1998, **108**, 4439–4449.
- [69] J. Schirmer and A. B. Trofimov, *J. Chem. Phys.*, 2004, **120**, 11449–11464.
- [70] A. B. Trofimov and J. Schirmer, *J. Phys. B*, 1995, **28**, 2299–2324.
- [71] A. Trofimov and J. Schirmer, *Chem. Phys.*, 1997, **214**, 153–170.
- [72] K. Sneskov and O. Christiansen, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2012, **2**, 566–584.
- [73] O. Christiansen, H. Koch and P. Jørgensen, *Chem. Phys. Lett.*, 1995, **243**, 409–418.
- [74] J. Finley, P. Å. Malmqvist, B. O. Roos and L. Serrano-Andrés, *Chem. Phys. Lett.*, 1998, **288**, 299–306.
- [75] M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer and W. Thiel, *J. Chem. Phys.*, 2008, **128**, 134110.
- [76] D. Tuna, D. Lefrancois, Ł. Wolański, S. Gozem, I. Schapiro, T. Andruniów, A. Dreuw and M. Olivucci, *J. Chem. Theory Comput.*, 2015, **11**, 5758–5781.
- [77] B. G. Levine and T. J. Martínez, *Annu. Rev. Phys. Chem.*, 2007, **58**, 613–634.
- [78] M. R. Silva-Junior, M. Schreiber, S. P. A. Sauer and W. Thiel, *J. Chem. Phys.*, 2008, **129**, 104103.
- [79] M. R. Silva-Junior, M. Schreiber, S. P. A. Sauer and W. Thiel, *J. Chem. Phys.*, 2010, **133**, 174318.
- [80] B. Helmich-Paris, *J. Chem. Theory Comput.*, 2019, **15**, 4170–4179.
- [81] B. Helmich-Paris, *J. Chem. Phys.*, 2019, **150**, 174121.

- [82] T. Shiozaki, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2018, **8**, 1–7.
- [83] J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999–3094.
- [84] B. Mennucci, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2012, **2**, 386–404.
- [85] S. Miertuš, E. Scrocco and J. Tomasi, *Chem. Phys.*, 1981, **55**, 117–129.
- [86] T. Kato, *Commun. Pure Appl. Math.*, 1957, **10**, 151–177.
- [87] N. H. March, *Phys. Rev. B*, 1986, **33**, 88–89.
- [88] A. Biancardi, J. Barnes and M. Caricato, *J. Chem. Phys.*, 2016, **145**, 224109.
- [89] A. Warshel and M. Levitt, *J. Mol. Biol.*, 1976, **103**, 227–249.
- [90] F. Maseras and K. Morokuma, *J. Comput. Chem.*, 1995, **16**, 1170–1179.
- [91] R. P. Fornari, P. Rowe, D. Padula and A. Troisi, *J. Chem. Theory Comput.*, 2017, **13**, 3754–3763.
- [92] H. P. Hratchian, P. V. Parandekar, K. Raghavachari, M. J. Frisch and T. Vreven, *J. Chem. Phys.*, 2008, **128**, 034107.
- [93] R. S. Mulliken, *J. Chem. Phys.*, 1955, **23**, 1833–1840.
- [94] F. Martin and H. Zipse, *J. Comput. Chem.*, 2005, **26**, 97–105.
- [95] M. A. Hagräs and W. J. Glover, *J. Chem. Theory Comput.*, 2018, **14**, 2137–2144.
- [96] M. Persico and G. Granucci, *Photochemistry*, Springer International Publishing, Cham, 2018.
- [97] B. Tang, H. Liu, F. Li, Y. Wang and H. Zhang, *ChemComm*, 2016, **52**, 6577–6580.
- [98] R. C. Hilborn, *Am. J. Phys.*, 1982, **50**, 982–986.
- [99] P. Dirac, *Proc. Royal Soc. Lond. A*, 1927, **114**, 243–265.
- [100] Y. Niu, Q. Peng, C. Deng, X. Gao and Z. Shuai, *J. Phys. Chem. A*, 2010, **114**, 7817–7831.
- [101] Q. Peng and Z. Shuai, in *ACS Symposium Series*, ACS, 2016, vol. 1226, pp. 35–59.

- [102] C. A. Mead and D. G. Truhlar, *J. Chem. Phys.*, 1982, **77**, 6090–6098.
- [103] D. R. Yarkony, *Acc. Chem. Res.*, 1998, **31**, 511–518.
- [104] D. Escudero, in *Photodeactivation Channels of Transition Metal Complexes: A Computational Chemistry Perspective*, Springer International Publishing, 2019, vol. 29, pp. 259–287.
- [105] J. Aragó and A. Troisi, *Phys. Rev. Lett.*, 2015, **114**, 026402.
- [106] S. Jurinovich, L. Cupellini, C. A. Guido and B. Mennucci, *J. Comput. Chem.*, 2018, **39**, 279–286.
- [107] M. Kasha, H. R. Rawls and M. Ashraf El-Bayoumi, *Pure Appl. Chem.*, 1965, **11**, 371–392.
- [108] N. J. Hestand and F. C. Spano, *Chem. Rev.*, 2018, **118**, 7069–7163.
- [109] A. Severo Pereira Gomes and C. R. Jacob, *Annu. Rep. Prog. Chem.*, 2012, **108**, 222–277.
- [110] F. Libisch, C. Huang and E. A. Carter, *Acc. Chem. Res.*, 2014, **47**, 2768–2775.
- [111] J. Cheng, K. Yu, F. Libisch, J. M. Dieterich and E. A. Carter, *J. Chem. Theory Comput.*, 2017, **13**, 1081–1093.
- [112] J. Cheng, F. Libisch, K. Yu, M. Chen, J. M. Dieterich and E. A. Carter, *J. Chem. Theory Comput.*, 2017, **13**, 1067–1080.
- [113] M. Klintenber, S. Derenzo and M. Weber, *Comput. Phys. Commun.*, 2000, **131**, 120–128.
- [114] S. E. Derenzo, M. K. Klintenber and M. J. Weber, *J. Chem. Phys.*, 2000, **112**, 2074–2081.
- [115] K. Jug and T. Bredow, *J. Comput. Chem.*, 2004, **25**, 1551–1567.
- [116] D. Stueber, F. N. Guenneau and D. M. Grant, *J. Chem. Phys.*, 2001, **114**, 9236–9243.
- [117] A. A. Sokol, S. T. Bromley, S. A. French, C. R. A. Catlow and P. Sherwood, *Int. J. Quantum Chem.*, 2004, **99**, 695–712.

- [118] S. Metz, J. Kästner, A. A. Sokol, T. W. Keal and P. Sherwood, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2014, **4**, 101–110.
- [119] I. V. Abarenkov, *Phys. Rev. B*, 2007, **76**, 165127.
- [120] P. V. Sushko and I. V. Abarenkov, *J. Chem. Theory Comput.*, 2010, **6**, 1323–1333.
- [121] O. Kobayashi and S. Nanbu, *Chem. Phys.*, 2015, **461**, 47–57.
- [122] K. Nam, J. Gao and D. M. York, *J. Chem. Theory Comput.*, 2005, **1**, 2–13.
- [123] Y. Wang, K. Lopata, S. A. Chambers, N. Govind and P. V. Sushko, *J. Phys. Chem. C*, 2013, **117**, 25504–25512.
- [124] P. V. Sushko, L. Qiao, M. Bowden, T. Varga, G. J. Exarhos, F. K. Urban, D. Barton and S. A. Chambers, *Phys. Rev. Lett.*, 2013, **110**, 077401.
- [125] C. Sousa, S. Tosoni and F. Illas, *Chem. Rev.*, 2013, **113**, 4456–4495.
- [126] G. Cui and W. Yang, *J. Chem. Phys.*, 2011, **134**, 204115.
- [127] L. Wilbraham, M. Louis, D. Alberga, A. Brosseau, R. Guillot, F. Ito, F. Labat, R. Métivier, C. Allain and I. Ciofini, *Adv. Mater.*, 2018, **30**, 1800817.
- [128] P. Cieplak, F.-Y. Dupradeau, Y. Duan and J. Wang, *J. Phys. Condens. Matter*, 2009, **21**, 333102.
- [129] A. Warshel, M. Kato and A. V. Pisliakov, *J. Chem. Theory Comput.*, 2007, **3**, 2034–2045.
- [130] W. Xie and J. Gao, *J. Chem. Theory Comput.*, 2007, **3**, 1890–1900.
- [131] M. Dommett and R. Crespo-Otero, *Phys. Chem. Chem. Phys.*, 2017, **19**, 2409–2416.
- [132] M. Dommett, M. Rivera and R. Crespo-Otero, *J. Phys. Chem. Lett.*, 2017, **8**, 6148–6153.
- [133] J. Shi, L. E. Aguilar Suarez, S. J. Yoon, S. Varghese, C. Serpa, S. Y. Park, L. Lüer, D. Roca-Sanjuán, B. Milián-Medina and J. Gierschner, *J. Phys. Chem. C*, 2017, **121**, 23166–23183.

- [134] X. Cheng, K. Wang, S. Huang, H. Zhang, H. Zhang and Y. Wang, *Angew. Chem. Int. Ed.*, 2015, **54**, 8369–8373.
- [135] R. Crespo-Otero, Q. Li and L. Blancafort, *Chem.: Asian J.*, 2019, **14**, 700–714.
- [136] Q. Li and L. Blancafort, *ChemComm*, 2013, **49**, 5966–5968.
- [137] M. Rivera, M. Dommett and R. Crespo-Otero, *fromage: A FRamework for Molecular AGgregate Excitations*, 2018, <https://github.com/Crespo-Otero-group/fromage>.
- [138] M. Rivera, M. Dommett and R. Crespo-Otero, *Documentation for fromage*, 2018, <https://fromage.readthedocs.io/>.
- [139] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Menucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian 16 Revision B.01*, 2016, Gaussian Inc. Wallingford CT.
- [140] F. Aquilante, J. Autschbach, R. K. Carlson, L. F. Chibotaru, M. G. Delcey, L. De Vico, I. Fdez Galván, N. Ferré, L. M. Frutos, L. Gagliardi, M. Garavelli, A. Giussani, C. E. Hoyer, G. Li Manni, H. Lischka, D. Ma, P. Å. Malmqvist, T. Müller, A. Nenov, M. Olivucci, T. B. Pedersen, D. Peng, F. Plasser, B. Pritchard, M. Reiher, I. Rivalta, I. Schapiro, J. Segarra-Martí, M. Stenrup, D. G. Truhlar, L. Ungur, A. Valentini, S. Vancoillie, V. Veryazov,

- V. P. Vysotskiy, O. Weingart, F. Zapata and R. Lindh, *J. Comput. Chem.*, 2016, **37**, 506–541.
- [141] *TURBOMOLE V6.2 2010, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>.*
- [142] B. Aradi, B. Hourahine and T. Frauenheim, *J. Phys. Chem. B A*, 2007, **111**, 5678–5684.
- [143] B. G. Levine, J. D. Coe and T. J. Martínez, *J. Phys. Chem. B*, 2008, **112**, 405–413.
- [144] S. Ruiz-Barragan, K. Morokuma and L. Blancafort, *J. Chem. Theory Comput.*, 2015, **11**, 1585–1594.
- [145] M. Barbatti, M. Ruckebauer, F. Plasser, J. Pittner, G. Granucci, M. Persico and H. Lischka, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2014, **4**, 26–33.
- [146] S. Gozem, F. Melaccio, A. Valentini, M. Filatov, M. Huix-Rotllant, N. Ferré, L. M. Frutos, C. Angeli, A. I. Krylov, A. A. Granovsky, R. Lindh and M. Olivucci, *J. Chem. Theory Comput.*, 2014, **10**, 3074–3084.
- [147] R. Crespo-Otero, A. Mardykov, E. Sanchez-Garcia, W. Sander and M. Barbatti, *Phys. Chem. Chem. Phys.*, 2014, **16**, 18877–18887.
- [148] M. Barbatti and R. Crespo-Otero, *Top. Curr. Chem.*, 2016, **368**, 415–444.
- [149] L. W. Chung, W. M. C. Sameera, R. Ramozzi, A. J. Page, M. Hatanaka, G. P. Petrova, T. V. Harris, X. Li, Z. Ke, F. Liu, H.-B. Li, L. Ding and K. Morokuma, *Chem. Rev.*, 2015, **115**, 5678–5796.
- [150] M. Kepenekian, V. Robert, B. Le Guennic and C. De Graaf, *J. Comput. Chem.*, 2009, **30**, 2327–2333.
- [151] B. Le Guennic, S. Borshch and V. Robert, *Inorg. Chem.*, 2007, **46**, 11106–11111.
- [152] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj,

- M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauszero, A. P. Seitsonen, A. Smogunov, P. Umari and R. M. Wentzcovitch, *J. Phys. Condens. Matter*, 2009, **21**, 395502.
- [153] J. Hutter, M. Iannuzzi, F. Schiffmann and J. VandeVondele, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2014, **4**, 15–25.
- [154] G. Henkelman, A. Arnaldsson and H. Jónsson, *Comput. Mater. Sci.*, 2006, **36**, 354–360.
- [155] E. Sanville, S. D. Kenny, R. Smith and G. Henkelman, *J. Comput. Chem.*, 2007, **28**, 899–908.
- [156] M. Yu and D. R. Trinkle, *J. Chem. Phys.*, 2011, **134**, 064111.
- [157] W. Tang, E. Sanville and G. Henkelman, *J. Phys. Condens. Matter*, 2009, **21**, 084204.
- [158] N. I. Zahid, M. S. Mahmood, B. Subramanian, S. Mohd Said and O. K. Abou-Zied, *J. Phys. Chem. Lett.*, 2017, **8**, 5603–5608.
- [159] R. Crespo-Otero and M. Barbatti, *Theor. Chem. Acc.*, 2012, **131**, 1–14.
- [160] M. Barbatti, G. Granucci, M. Ruckebauer, F. Plasser, R. Crespo-Otero, J. Pittner, M. Persico and Lischka, *NEWTON-X: A Package for Newtonian Dynamics Close to the Crossing Seam. Version 2*, 2016, www.newtonx.org.
- [161] L. Wilbraham, M. Savarese, N. Rega, C. Adamo and I. Ciofini, *J. Phys. Chem. B*, 2015, **119**, 2459–2466.
- [162] J. M. Skelton, E. Lora da Silva, R. Crespo-Otero, L. E. Hatcher, P. R. Raithby, S. C. Parker and A. Walsh, *Faraday Discuss.*, 2015, **177**, 181–202.
- [163] X.-L. Peng, S. Ruiz-Barragan, Z.-S. Li, Q.-S. Li and L. Blancafort, *J. Mat. Chem. C*, 2016, **4**, 2802–2810.
- [164] R. Crespo-Otero and M. Barbatti, *Chem. Rev.*, 2018, **118**, 7026–7068.

- [165] M. Barbatti, Z. Lan, R. Crespo-Otero, J. J. Szymczak, H. Lischka and W. Thiel, *J. Chem. Phys.*, 2012, **137**, 22A503.
- [166] F. C. Spano and D. Beljonne, *Photophysical Properties of Molecular Aggregates "101"*, World Scientific, Hackensack, 2016, pp. 93–130.
- [167] R. Dovesi, A. Erba, R. Orlando, C. M. Zicovich-Wilson, B. Civalleri, L. Maschio, M. Rérat, S. Casassa, J. Baima, S. Salustro and B. Kirtman, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2018, **8**, e1360.
- [168] C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, *Acta Crystallogr. B*, 2016, **72**, 171–179.
- [169] A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng and K. W. Jacobsen, *J. Phys. Condens. Matter*, 2017, **29**, 273002.
- [170] *RDKit: Open-source cheminformatics*, <http://www.rdkit.org>.
- [171] Y. Lu, M. R. Farrow, P. Fayon, A. J. Logsdail, A. A. Sokol, C. R. A. Catlow, P. Sherwood and T. W. Keal, *J. Chem. Theory Comput.*, 2019, **15**, 1317–1328.
- [172] A. V. Akimov, *J. Comput. Chem.*, 2016, **37**, 1626–1649.
- [173] L. Stojanović and R. Crespo-Otero, *ChemPhotoChem*, 2019, **3**, 907–915.
- [174] M. Dommett, M. Rivera, M. T. H. Smith and R. Crespo-Otero, *J. Mater. Chem. C*, 2020, **8**, 2558–2568.
- [175] T. Oliphant, *NumPy: A guide to NumPy*, USA: Trelgol Publishing, 2006–, <http://www.numpy.org/>.
- [176] P. P. Eric Jones, Travis Oliphant, *SciPy: Open source scientific tools for Python*, 2001–, <http://www.scipy.org/>.

- [177] G. Voronoi, *J. Reine Angew. Math.*, 1907, **133**, 97–178.
- [178] P. Procacci and R. Scateni, *Int. J. Quantum Chem.*, 1992, **42**, 1515–1528.
- [179] J. S. Hunjan and B. C. Eu, *J. Chem. Phys.*, 2010, **132**, 134510.
- [180] J. L. Abascal, M. A. Gonzalez, J. L. Aragones and C. Valeriani, *J. Chem. Phys.*, 2013, **138**, 084508.
- [181] G. Menzl, M. A. Gonzalez, P. Geiger, F. Caupin, J. L. Abascal, C. Valeriani and C. Dellago, *Proc. Natl. Acad. Sci. U.S.A.*, 2016, **113**, 13582–13587.
- [182] F. Plasser and H. Lischka, *J. Chem. Theory Comput.*, 2012, **8**, 2777–2789.
- [183] F. Plasser, *TheoDORE: a package for theoretical density, orbital relaxation, and exciton analysis*, 2017.
- [184] R. Crespo-Otero and M. Barbatti, *Theor. Chem. Acc.*, 2012, **131**, 1237.
- [185] K. Sen, R. Crespo-Otero, O. Weingart, W. Thiel and M. Barbatti, *J. Chem. Theory Comput.*, 2013, **9**, 533–542.
- [186] M. Botoshansky, F. H. Herbstein and M. Kapon, *Helv. Chim. Acta*, 2003, **86**, 1113–1128.
- [187] C.-P. Hsu, *Acc. Chem. Res.*, 2009, **42**, 509–518.
- [188] J. Gierschner, D. Oelkrug and H.-j. Egelhaaf, *J. Phys. Chem. Lett.*, 2013, **4**, 2686–2697.
- [189] P. F. Van Hutten, J. Wildeman, A. Meetsma and G. Hadziioannou, *J. Am. Chem. Soc.*, 1999, **121**, 5910–5918.
- [190] J. R.-G. Pedregal, I. Funes-Ardoiz, G. Sciortino, J.-E. Sánchez-Aparicio, G. Ujaque, A. Lledós, J.-D. Maréchal and F. Maseras, *J. Comput. Chem.*, 2018, 1–6.
- [191] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.
- [192] J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graph.*, 2006, **25**, 247–260.

- [193] M. Rivera, M. Dommett, A. Sidat, W. Rahim and R. Crespo-Otero, *J. Comput. Chem.*, 2020, **41**, 1045–1058.
- [194] J. R. Reimers, *J. Chem. Phys.*, 2001, **115**, 9103–9109.
- [195] S. J. Strickler and R. A. Berg, *J. Chem. Phys.*, 1962, **37**, 814–822.
- [196] V. Stehr, R. F. Fink, B. Engels, J. Pflaum and C. Deibel, *J. Chem. Theory Comput.*, 2014, **10**, 1242–1255.
- [197] A. Kimura, T. Kakitani and T. Yamato, *J. Phys. Chem. B*, 2000, **104**, 9276–9287.
- [198] G. R. Desiraju and A. Gavezzotti, *Acta Crystallogr. B*, 1989, **45**, 473–482.
- [199] J. E. Campbell, J. Yang and G. M. Day, *J. Mater. Chem. C*, 2017, **5**, 7574–7584.
- [200] J. Lee, C. H. Kim and T. Joo, *J. Phys. Chem. A*, 2013, **117**, 1400–1405.
- [201] M. Du, B. Huo, M. Li, A. Shen, X. Bai, Y. Lai, J. Liu and Y. Yang, *RSC Adv.*, 2018, **8**, 32497–32505.
- [202] T. G. Pavlopoulos and P. R. Hammond, *J. Am. Chem. Soc.*, 1974, **96**, 6568–6579.
- [203] R. Katoh, K. Suzuki, A. Furube, M. Kotani and K. Tokumaru, *J. Phys. Chem. C*, 2009, **113**, 2961–2965.
- [204] J. Stampfl, S. Tasch, G. Leising and U. Scherf, *Synthetic Metals*, 1995, **71**, 2125 – 2128.
- [205] Q. Chen, C. Jia, Y. Zhang, W. Du, Y. Wang, Y. Huang, Q. Yang and Q. Zhang, *J. Mater. Chem. B*, 2017, **5**, 7736–7742.
- [206] W. Zhang, T. Sakurai, M. Aotani, G. Watanabe, H. Yoshida, V. S. Padalkar, Y. Tsutsui, D. Sakamaki, M. Ozaki and S. Seki, *Adv. Opt. Mater.*, 2019, **7**, 1801349.
- [207] P. Majumdar and J. Zhao, *J. Phys. Chem. B*, 2015, **119**, 2384–2394.
- [208] M. Barbatti, A. J. A. Aquino, H. Lischka, C. Schrieffer, S. Lochbrunner and E. Riedle, *Phys. Chem. Chem. Phys.*, 2009, **11**, 1406–1415.

- [209] S. Pijeu, D. Foster and E. G. Hohenstein, *J. Phys. Chem. A*, 2017, **121**, 4595–4605.
- [210] W. E. Brewer, M. L. Martinez and P. T. Chou, *J. Phys. Chem.*, 1990, **94**, 1915–1918.
- [211] R. Hu, S. Li, Y. Zeng, J. Chen, S. Wang, Y. Li and G. Yang, *Phys. Chem. Chem. Phys.*, 2011, **13**, 2044–2051.
- [212] J. Lee, C. H. Kim and T. Joo, *J. Phys. Chem. A*, 2013, **117**, 1400–1405.
- [213] W. Zhang, Y. Yan, J. Gu, J. Yao and Y. S. Zhao, *Angew. Chem. Int. Ed.*, 2015, **54**, 7125–7129.
- [214] Q. Li and L. Blancafort, *Chem. Commun.*, 2013, **49**, 5966–5968.
- [215] A. Andreev, G. Matt, C. J. Brabec, H. Sitter, D. Badt, H. Seyringer and N. S. Sariciftci, *Adv. Mater.*, 2000, **12**, 629–633.
- [216] Y. Duan, C. Ju, G. Yang, E. Fron, E. Coutino-Gonzalez, S. Semin, C. Fan, R. S. Balok, J. Cremers, P. Tinnemans, Y. Feng, Y. Li, J. Hofkens, A. E. Rowan, T. Rasing and J. Xu, *Adv. Funct. Mater.*, 2016, **26**, 8968–8977.
- [217] F. Quochi, F. Cordella, A. Mura, G. Bongiovanni, F. Balzer and H.-G. Rubahn, *Appl. Phys. Lett.*, 2006, **88**, 041106.
- [218] F. Quochi, A. Andreev, F. Cordella, R. Orrú, A. Mura, G. Bongiovanni, H. Hoppe, H. Sitter and N. Sariciftci, *J. Lumin.*, 2005, **112**, 321 – 324.
- [219] F. Quochi, M. Saba, F. Cordella, A. Gocalinska, R. Corpino, M. Marceddu, A. Anedda, A. Andreev, H. Sitter, N. S. Sariciftci, A. Mura and G. Bongiovanni, *Adv. Mater.*, 2008, **20**, 3017–3021.
- [220] V. Mukundam, K. Dhanunjarao, R. Mamidala and K. Venkatasubbaiah, *J. Mater. Chem. C*, 2016, **4**, 3523–3530.
- [221] I. Schapiro, F. Melaccio, E. N. Laricheva and M. Olivucci, *Photochem. Photobiol. Sci.*, 2011, **10**, 867–886.
- [222] M. A. Izquierdo, J. Shi, S. Oh, S. Y. Park, B. Milián-Medina, J. Gierschner and D. Roca-Sanjuán, *Phys. Chem. Chem. Phys.*, 2019, **21**, 22429–22439.

- [223] J. Shi, M. A. Izquierdo, S. Oh, S. Y. Park, B. Milián-Medina, D. Roca-Sanjuán and J. Gierschner, *Org. Chem. Front.*, 2019, **6**, 1948–1954.
- [224] Z. Shuai and Q. Peng, *Phys. Rep.*, 2014, **537**, 123–156.
- [225] Q. Peng, Y. Yi, Z. Shuai and J. Shao, *J. Am. Chem. Soc.*, 2007, **129**, 9333–9339.
- [226] D. Loco, E. Polack, S. Caprasecca, L. Lagardère, F. Lipparini, J.-P. Piquemal and B. Mennucci, *J. Chem. Theory Comput.*, 2016, **12**, 3654–3661.
- [227] B. Jeziorski, R. Moszynski and K. Szalewicz, *Chem. Rev.*, 1994, **94**, 1887–1930.
- [228] G. Jansen, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2014, **4**, 127–144.
- [229] I. Fukuda and H. Nakamura, *Biophys. Rev.*, 2012, **4**, 161–170.
- [230] D. Wolf, P. Keblinski, S. R. Phillpot and J. Eggebrecht, *J. Chem. Phys.*, 1999, **110**, 8254–8282.

Appendix A

Supporting Information for the Ewald Embedded Cluster

A.1 Crystal Structures

The experimental crystal structures for HC1 and HC2 are accessible on the Cambridge Structural Database with codes 941991 and 1061608 respectively.¹³⁴

A.2 Absorption and Emission Spectra

To consider the effect of vibrations on the position of the absorption and emission maxima, we simulate both spectra with the nuclear ensemble approach as implemented in Newton-X.^{160,184} They are shown in Figure A.1. For each scheme, the monomer was first optimised to the FC and K* geometries for absorption and emission. The vibrational modes were calculated using ω B97X-D/6-311++G(d,p) in Gaussian¹³⁹ in the presence of the point charges from the embedding model.

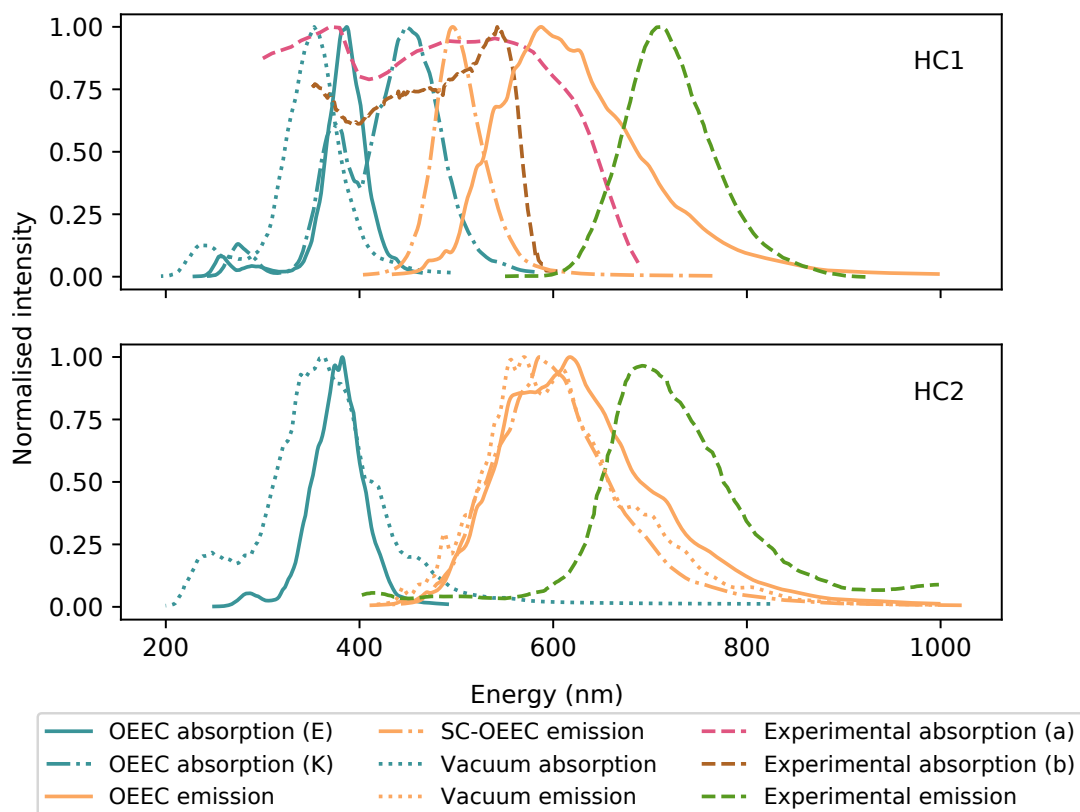


FIGURE A.1: Absorption and emission spectra for both molecules with experimental data¹³⁴ for comparison. The experimental emission for HC2 is smoothed with a Savitzky-Golay filter.

A.3 Population Analysis for Aggregates

RESP charges respond directly to the electrostatic environment of the atom. As such they are a measure of the quality of the embedding method. In Figure A.2 we compare the RESP charges of notable atoms on monomers embedded in different charge backgrounds to those calculated in tetramers embedded using OEEC. In this case the central molecule is in the FC geometry. The results for \mathbf{K}^* are in Figure A.3.

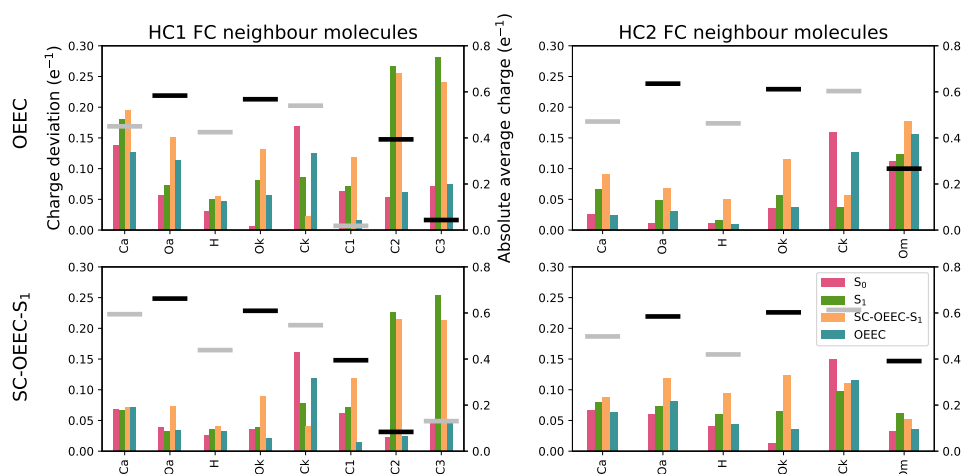


FIGURE A.2: Deviation of charge on notable atoms from the explicitly calculated tetramer with the relaxing molecule in **FC** geometry. In the top row, the molecule is optimised with **OEEC** and in the bottom with **SC-OEEC-S₁**. We compare the charge obtained from the monomer population analysis of in vacuum S_0 , S_1 and in **EEC** and **SC-OEEC-S₁**. The absolute average charge in the tetramer is indicated by horizontal ticks, positive values are in grey and negative ones in black. The atom labels are the ones found in Figure 1 of the main text.

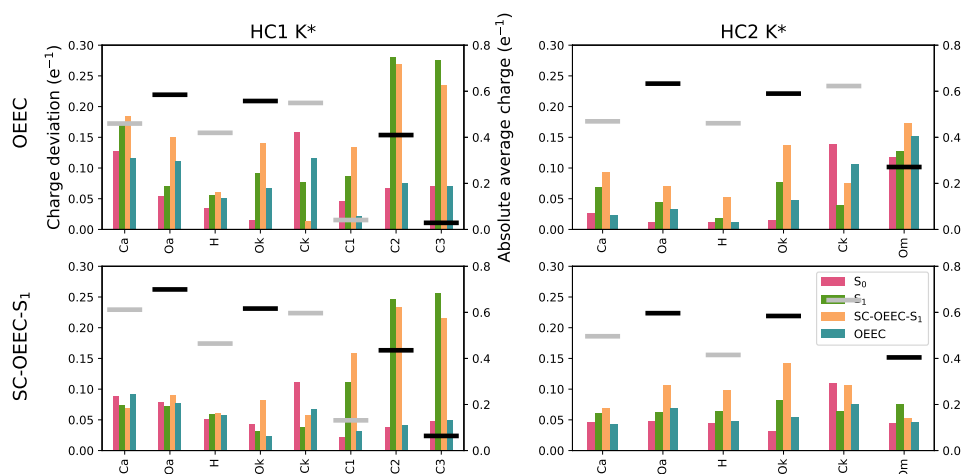


FIGURE A.3: Deviation of charge on notable atoms from the explicitly calculated tetramer with the relaxing molecule in **K*** geometry. In the top row, the molecule is optimised with **OEEC** and in the bottom with **SC-OEEC-S₁**. We compare the charge obtained from the monomer population analysis of in vacuum S_0 , S_1 and in **EEC** and **SC-OEEC-S₁**. The absolute average charge in the tetramer is indicated by horizontal ticks, positive values are in grey and negative ones in black. The atom labels are the ones found in Figure 1 of the main text.

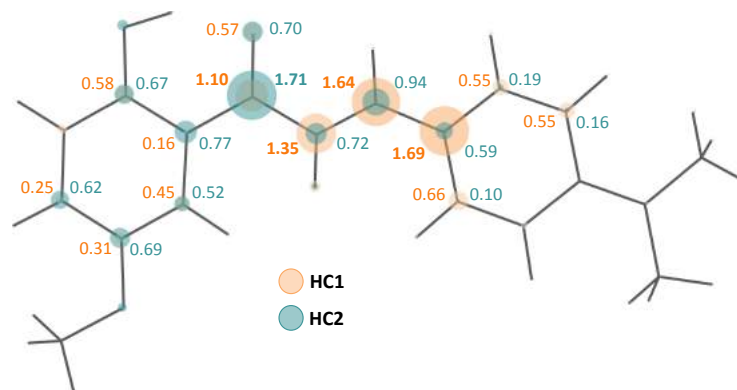


FIGURE A.4: Largest differences in RESP charges between S_1 and S_0 for HC1 and HC2 in vacuum in units of $10^{-1} e^{-1}$. We only include atoms where the difference was more than $0.01 e^{-1}$ in either molecule. The diameter of the circles is proportional to the charge difference.

A.4 Vertical Excitation for Aggregate

The bright state of **HC2** has a different identity depending on the charge background. We illustrate this in Figs A.5 and A.6.

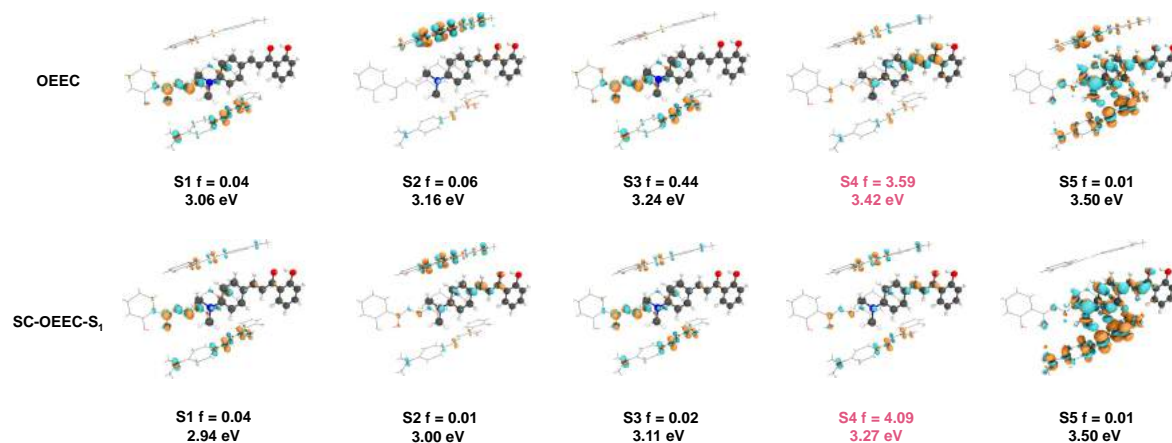


FIGURE A.5: Density plots for the absorption of HC1 under OEEC and SC-OEEC. The for each excited state n , the difference in electronic density $S_n - S_0$ is plotted where positive values are shown in orange and negative values in blue. The bright state is highlighted in pink.

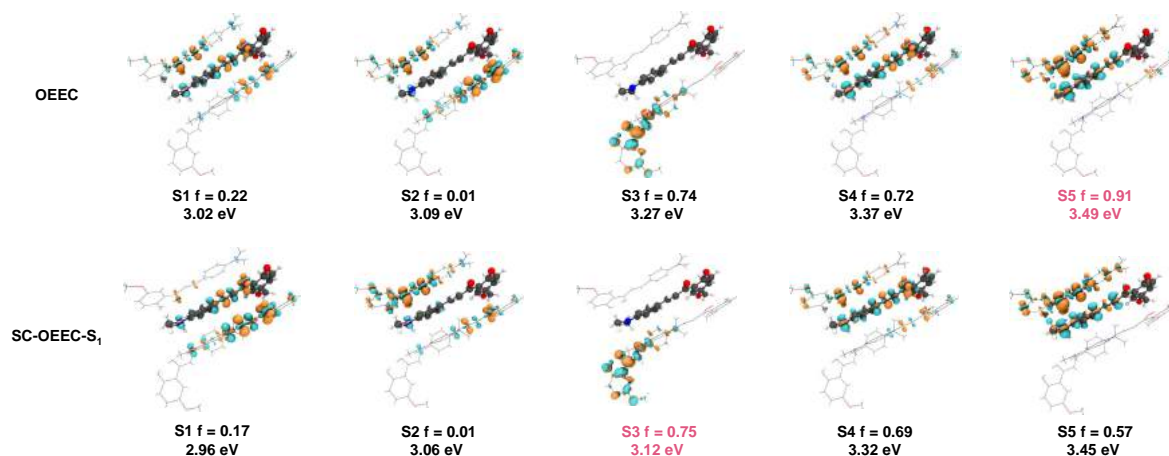


FIGURE A.6: Density plots for the absorption of HC2 under OEEC and SC-OEEC. For each excited state n , the difference in electronic density $S_n - S_0$ is plotted where positive values are shown in orange and negative values in blue. The bright state is highlighted in pink.

A.5 Exciton Coupling in Dimers

Dimeric exciton couplings were computed using the diabatisation scheme devised by Arag3 and Troisi.¹⁰⁵ In this scheme, a diabatisation excited state property is selected. (1) It is first evaluated for the dimer in S_1 and S_2 , along with the excited state energy. (2) It is also calculated for the constituent monomers of the dimer. (3) Then, a singular-value decomposition is employed to determine the matrix which best transforms the property from the adiabatic to the diabatic basis. (4) Finally, the newly obtained diabatisation matrix is used to transform the adiabatic Hamiltonian into the diabatic one. The latter's off-diagonal elements are the desired J exciton coupling values.

More details are offered in section 5.3.2. The results for HC1 and HC2 are shown in Table A.1. Also shown are the couplings obtained using half of the energy gap between S_1 and S_2 in the dimer.

TABLE A.1: Exciton couplings (eV) for the different fragments involved in the tetramers computed with the diabatisation scheme. In parenthesis is half the energy splitting between S_1 and S_2 . The dimers A, B and C are shown in the Figure 6 of the main text

Dimer	HC1	HC2
A	0.012 (0.012)	0.116 (0.117)
B	0.108 (0.109)	0.148 (0.148)
C	0.061 (0.063)	0.001 (0.003)

A.6 Results with CC2 and CASPT2

TABLE A.2: Table of absorption and emission energies with RI-CC2 and multireference methods for both model systems after geometry optimisation. The energies are in eV

Cluster model	Method	HC1		HC2	
		FC	K*	FC	K*
OEEC	RI-CC2/TZVP	2.99	1.08	3.11	0.97
	RI-CC2/TZVP//TD- ω B97X-D/6-311++G(d,p)	3.08	1.48	3.20	1.62
	CASPT2/6-31G(d)//CASSCF/6-31G(d)	3.53	1.41	-	-
	CASPT2/6-31G(d)//TD- ω B97X-D/6-311++G(d,p)	3.52	1.74	-	-
SC-OEEC S_1	RI-CC2/TZVP	2.61	1.41	2.99	1.03
	RI-CC2/TZVP//TD- ω B97X-D/6-311++G(d,p)	2.68	2.19	3.07	1.73
	CASPT2/6-31G(d)//CASSCF/6-31G(d)	3.06	1.50	-	-
	CASPT2/6-31G(d)//TD- ω B97X-D/6-311++G(d,p)	3.01	2.27	-	-

Appendix B

Supporting Information for the Applications

B.1 Geometries

The optimised geometries of the molecules discussed in the main paper are deposited at: https://github.com/Crespo-Otero-group/paper_data

B.2 Crystal Structures

All crystal structures were obtained from the Cambridge Crystallographic Database. The ones available freely are listed in Table B.1. As of writing this paper, α MODCS, β MODCS, α MODBDCS, and β MODBDCS do not have deposited structures.²²³

Crystal	CCDC ID
DSB	921998
4PV	129139
HBT	727487
3P	1269382
4P	1245769
6P	1319661
α DCS	1247728
α DBDCS	778284
β DBDCS	969314

TABLE B.1: Cambridge Crystallographic Data Centre (CCDC) IDs for the corresponding crystal structures.

System	Reorganisation energy (cm ⁻¹)					
	Vacuum			Crystal		
	$\omega \leq 250$	$\omega > 250$	Total	$\omega \leq 250$	$\omega > 250$	Total
3P	912.7	3221.3	4134	64.6	2791.4	2856
4P	1423.7	2948.3	4372	383.6	2560.4	2944
6P	912.6	3075.4	3988	104.3	2230.7	2335
α DCS	764.9	2550.1	3315	370.7	2335.3	2706
α DBDCS	705.9	2228.1	2934	129.6	1916.4	2046
β DBDCS	724.6	2593.4	3318	286.2	2672.8	2959
α MODCS	756.2	2675.8	3432	433.7	2013.3	2447
β MODCS	721.6	1777.4	2499	252.3	1990.7	2243
α MODBDCS	473.3	1971.7	2445	71.7	1878.3	1950
β MODBDCS	617	2106	2723	75.6	1792.4	1868

TABLE B.2: Reorganisation energies for normal modes of the n P and DCS series, in vacuum and crystal, split between low and high energy normal modes.

B.3 Active Space

CASSCF and CASPT2 were carried out for 3P and HBT. The active spaces are reported in Figure B.1.

B.4 Full Normal Mode Analysis

The reorganisation energies in vacuum and in crystal discussed in the main paper are only the ones below 250 cm⁻¹. Here, we include the rest of the energies:

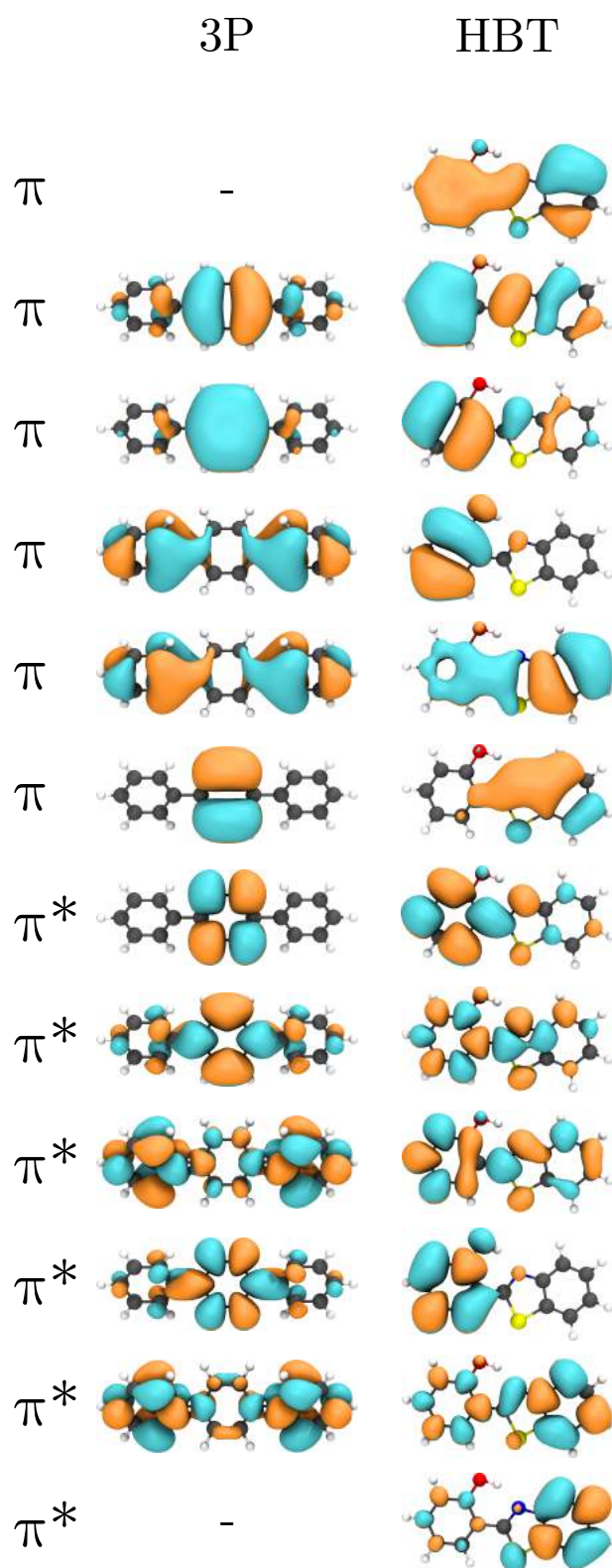


FIGURE B.1: Active spaces of 3P and HBT for multireference calculations.