Quantum Chemical Calculations of the Excited States of Porphyrins

Christian J. Suess

A thesis presented for the degree of

Doctor of Philosophy



School of Chemistry

University of Nottingham

September 2017

Quantum Chemical Calculations of the Excited States of Porphyrins

Christian Suess

Abstract

The development of optical multidimensional spectroscopic techniques has opened up new possibilities for the study of biological processes. Ultrafast two-dimensional ultraviolet spectroscopy experiments have determined the rates of tryptophan $\rightarrow$ heme electron transfer and excitation energy transfer for the two tryptophan residues in myoglobin. Here we show that accurate prediction of these rates can be achieved using Marcus theory in conjunction with time dependent density functional theory (TDDFT). Key intermediate residues between the donor and acceptor are identified, and in particular the residues Val68 and Ile75 play a critical role in calculations of the electron coupling matrix elements. Our calculations demonstrate how small changes in structure can have a large effect on the rates, and show that the different rates of electron transfer are dictated by the distance between the heme and tryptophan residues, while for excitation energy transfer the orientation of the tryptophan residues relative to the heme is important.

The absorption and fluorescence spectroscopy of a series of porphyrin based systems have been studied. The range of systems has been selected in order to investigate the influence of both the electronic and geometric configuration on the photophysical properties. The origin of the bathochromic shift in the absorption and fluorescence spectra of substituted porphyrins is attributed to both steric distortions of the ring and electronic effects of the substituents. Three DFT based approaches have been used to model and calculate these properties. The approach using the maximum overlap method (MOM) predicted the largest discrepancy from

experimental results whilst TDDFT calculated shifts within 0.05 eV of experimental values. Finally a third method labelled as a 'hybrid' approach has been used, where the MOM is employed to optimise excited state geometries and single point TDDFT calculations are used to evaluate the vertical excitation energies. This approach improves on the excitation energies predicted by the MOM but does not improve on the values that a full TDDFT calculation produced. However, this 'hybrid' approach is computationally less demanding. There is distinct trade-off between accuracy and feasibility of calculations, where this 'hybrid' method of MOM and TDDFT becomes beneficial and useful.

High resolution spectra of both free base and metallocentred complexes of porphyrin are calculated for the Q band region. Calculations on the vibronic structure of porphyrin are performed using DFT and TDDFT. Both Franck–Condon (FC) and Herzberg–Teller (HT) approaches have been used to predict the frequency and intensity of vibronic bands in the simulated absorption spectra with respect to the $S_0 \rightarrow S_1$ electronic transition as the summation of contributions from both schemes characterise the electronic transitions and provide a high resolution description of the Q-band. Chapter five shows that the first electronic transition into the singlet excited state is vibronically active in the Q_x region of the absorption spectrum, in good agreement with experimental data. The HT scheme appears to have more significant contributions and provides more insight into resolving the vibrationally active area of the absorption spectra. HT contributions to the electronic transition dipole moments are essential to assign the weak vibrational transitions and reproduce the experimental spectral profile. In order to provide a detailed account of the vibronic structure it is necessary to assign the vibrational transitions using both HT and FC schemes. The importance of theoretical calculations are highlighted here and can help the general understanding of the photophysical properties of porphyrins.

For Frank

Acknowledgements

I would like to thank Jonathan Hirst and Nick Besley for all their help, support and guidance throughout my research. I know I must have frustrated them on many occasions but their patience throughout has allowed me to thoroughly enjoy the challenge of my PhD. Far and away the best prize that life has to offer is the chance to work hard at work worth doing. I am grateful for the opportunity to do work worth doing with people worth doing it for.

I would also like to thank Christof Jäger for giving me the chance to continue to develop myself as a scientist. The support as both a colleague and as a friend have been invaluable.

Under certain circumstances PhD life can be very lonely. I am lucky and grateful to have met some truly fantastic people who have influenced me along the way. So I would like to thank everyone in the extended Chemistry community at Nottingham University. I have really enjoyed spending time with you all!

Finally, to Emma, who supported me on this journey. Words can not justify the extent of her contributions towards this achievement.

Contents

1 Introduction **17**

1.1 Porphyrin Background 17

1.2 Porphyrins in nature 19

1.3 Uses and applications of porphyrins 20

1.4 Spectroscopic characteristics 21

1.4.1 Gouterman's four orbital model 23

1.4.2 Molecular orbitals 25

1.5 Porphyrin Varieties 26

1.5.1 Larger Units 29

1.5.2 Scope of Thesis 30

2 Overview of Theoretical Methods **32**

2.1 Theoretical and Computational Methods 32

2.2 The Quantum Mechanical Approach 32

2.2.1 The Hamiltonian 33

2.2.2 The Born-Oppenheimer Approximation 34

2.3 Hartree-Fock Theory 35

2.3.1 Linear Combination of Atomic Orbitals 41

2.3.2 Fock Matrix 43

2.3.3 Restricted and Unrestricted Hartree Fock Theory 45

2.4 Basis Sets 46

2.4.1 Slater and Gaussian Type Functions 46

2.4.2 Polarized and Diffuse Functions 49
2.4.3 Dunning Basis Sets 49
2.5 Density Functional Theory 51
2.5.1 Hohenberg-Kohn Theorems 52
2.5.2 Kohn-Sham DFT 52
2.5.3 Exchange and Correlation Functionals 56
2.5.4 Long-range corrected functionals 58
2.5.5 Dispersion Corrections within DFT 60
2.6 Excited State Methods 62
2.6.1 Time Dependent Density Functional Theory 63
2.6.2 Tamm-Dancoff Approximation 64
2.6.3 Maximum Overlap Method 65
2.6.4 Single Excitation Configuration Interaction 68
2.6.5 Configuration Interaction 68
2.6.6 Complete Active Space 69
2.7 Polarizable Continuum Models 69
2.8 *Ab Initio* Molecular Dynamics 70
2.9 General Theory of Molecular Spectroscopy 71
2.9.1 Franck-Condon Principle 73
2.10 Methods for predicting electron transfer 78
2.10.1 Generalised Mulliken Hush 79
2.10.2 Fragment Charge Difference 79

3 Quantum Chemical Calculations of Tryptophan→Heme Electron and Excitation Energy Transfer Rates in Myoglobin 81

3.1 Introduction 82
3.2 Methodology 88
3.3 Results and Discussion 92
3.4 Conclusions 99

4 Absorption and Fluorescence Spectroscopy of Porphyrins 100

4.1 Introduction 101

4.2 Methodology 108

4.3 Results and Discussion 112

4.4 Conclusion 123

5 Vibrationally resolved porphyrin spectra using the Franck-Condon and Herzberg-Teller approach 125

5.1 Introduction 126

5.2 Methodology 130

5.3 Results and Discussion 130

5.4 Conclusion 138

6 Conclusions 139

6.1 Radical approaches towards protein engineering 140

6.2 An outlook to the future of simulations 141

List of Figures

1.1 Free Base and Metal Centred Porphyrin 17

1.2 IUPAC numbering of carbons in porphyrin with the *meso*, *alpha* and *beta* positions labelled. 18

1.3 Electron delocalisation route of porphyrin shown in blue. 19

1.4 Skeletal structures of chlorophyll and heme shown together to highlight similarities. 20

1.5 Typical UV-Vis spectra for free base porphyrin. A magnified view of the Q-band is shown inset.[1] 22

1.6 Typical UV-Vis spectra for a metal centred porphyrin.[2] 23

1.7 Molecular orbitals for free base porphyrin (left) and metalloporphyrin (right). 24

1.8 Molecular orbital diagram for free base porphyrin. Values are obtained from a time dependent density functional theory calculation (6-311G*/CAM-B3LYP) which is explained in greater detail later. Not to scale 25

1.9 Free base porphyrin with substiuents highlighted and labelled. 27

1.10 An Azo linked zinc porphyrin dimer. 28

1.11 UV/vis spectra of porphyrin monomer and dimer. (Unpublished results from an experimental collaborator). 29

1.12 An example of just one of the many varieties of porphyrin oligomers (6-Porphyrin Nanoring)[3]. 30

1.13 A sapphyrin molecule with five pyrole units and larger centre. 31

2.1 Behaviour of e^x where $x = r$ (solid line, STO) and $x = r^2$ (dashed line, GTO) 47

2.2 The radial behaviour of various basis functions in atom-centered coordinates. The bold solid line at top is the STO ($\zeta = 1$) for the hydrogen 1s function; for the one-electron H system, it is also the exact solution of the Schrödinger equation. Nearest to it is the contracted STO-3G 1s function that has been optimised to match the STO. It is the sum of the many single Gaussian functions. Note that from 0.5 to 4.0 a.u., the STO-3G orbital matches the correct orbital closely. However, near the origin there is a notable difference and, were the plot to extend to very large r, it would be apparent that the decay of the STO-3G orbital is more rapid than the correct orbital. . 48

2.3 Franck–Condon principle energy diagram. Since electronic transitions are very fast compared with nuclear motions, vibrational levels are favored when they correspond to a minimal change in the nuclear coordinates. 74

3.1 The work presented in this chapter is published in the Journal of Computational Chemistry[4] and features on the front cover of the Journal for June 2017. 82

3.2 The rates of tryptophan $\rightarrow$ heme electron transfer and excitation energy transfer in myoglobin are simulated based upon Marcus theory in conjunction with quantum chemical calculations and reveal the structural factors affecting the rates of these processes. 83

3.3 Marcus parabolas give a visual representation of the values ΔG and λ. ΔG is the free energy difference between the excited state of the donor acceptor. λ is difference in energy for the excited state and the optimised excited state. 85

3.4 Schematic of the electronic configurations for ET and EET processes from the S_1 state of a tryptophan residue to the heme. 90

3.5 Atoms highlighted in red are held fixed in position during the optimisation of the structure of the heme, intervening residue and tryptophan. 91

3.6 Dominant electron transfer pathways for Trp14 and Trp7 to heme calculated by the pathway tunnelling model. For details of the residues see the main text. 93

3.7 The reduced myoglobin system consisting of heme, Trp14, Trp7 and the E-Helix. Different colours represent each amino acid. 95

3.8 Modified heme and tryptophan system. Original tryptophan orientations shown in green and modified orientations shown in red. 97

4.1 Modified Jablonski diagram showing in a schematic way the energies of the various states of the porphyrin molecule. 101

4.2 The different configurations of porphyrins. Left: Saddled, Middle: Planar, Right: Domed. 102

4.3 Absorption spectra of metalloporphyrin. The 'B' band occurs around 400 nm and is attributed to an $S_0 \rightarrow S_2$ transition. The 'Q' band occurs around 500 nm and is a result of an $S_0 \rightarrow S_1$ transition. Due to the relative weaker intensity it is highlighted in the figure.[1] 103

4.4 The Stokes shift which is the difference between the emission (blue) and absorption (red) maximum. 103

4.5 Core porphyrin unit with labelled atomic indices. 106

4.6 The metal versions of the porphyrin structures with the base unit highlighted. (Top Left: Zinc Tetraphenylporphyrin (ZnTPP) Top Right: Zinc-tetra-tert-Butyl Porphyrin (ZnT(t-Bu)P) Bottom Left: Zinc-octaethyl-tetrephenylporphyrin (ZnOETPP) Bottom Right: Zinc-dodecaphenylporphyrins (ZnDPP). 107

4.7 A comparison of the excited state geometries obtained using TDDFT (Red) and MOM (Blue). The RMSD between the two optimised systems is 0.05 Å. 109

4.8 A simplified diagram of an electronic interactions at both absorption and fluorescence. 109

4.9 A comparison of basis sets for predicting excitation energies for free base porphyrin. 113

4.10 A comparison of absorption spectra for average of the results from the AIMD simulations (Green), Single point gas phase (Blue) and Experiment (Red)[1]. 115

4.11 A comparison of stoke shifts (in eV) calculated using different methods against experiment. MOM $R^2 = 0.33$, 'Hybrid' $R^2 = 0.73$ and TDDFT $R^2 = 0.74$. 117

4.12 Potential Energy Surface diagram for a zinc centred porphyrin dimer where the dihedral angle between the two porphyrin units is rotated through 360°. Energy units are hartrees. 117

4.13 A side view showing a porphyrin dimer with dihedral angle between the two monomers being rotated through 360° 118

4.14 The HOMO and LUMO of a zinc porphyrin dimer at the lowest energy (27°) configuration and the highest energy (90°) configuration. 119

4.15 The UV spectra of a porphyrin dimer where the two monomers are at different dihedral angles. The graph is 'noisy' but shows how the dimer can be manipulated to absorb light over a large range in frequency. 119

4.16 Orbital diagrams of the HOMO and LUMO of each of the zinc (left) and free base (right) substituted porphyrin complexes. 120

4.17 Optimised structural parameters of H_2TPP (left) and H_2T(t-Bu)P (right) of the S_0 and S_1 states for both the MOM and TDDFT methods. For atomic indices see figure 4.5. 121

4.18 Optimised structural parameters of ZnTPP (left) and ZnT(t-Bu)P (right) of the S_0 and S_1 states for both the MOM and TDDFT methods. For atomic indices see figure 4.5. 121

5.1 IR spectra for the first four excited state geometries responsible for the four peaks in the Q band for free base porphyrin. Top Left is excited state 1, Top Right is excited state 2, Bottom Left is excited state 3 and Bottom Right is excited state 4. 132
5.2 Computed stick spectra (298K) of H_2P using only FC contributions (top), HT mechanism (middle) and a combined model including both FC and HT contributions, FCHT (bottom). 133
5.3 Q band of free base porphyrin with scaled HT and FC factors. 134
5.4 Q band of free base porphyrin with scaled HT and FC factors. This figure focuses on the 620-630 nm region. 135
5.5 IR spectra for the first two excited state geometries responsible for the two peaks in the Q band for zinc centred porphyrin 135
5.6 Comparisons of the FCHT absoprtion stick spectra calculated at 0.5K (left) and 298K (right). 137

6.1 QueE protein with the radical SAM enzyme nestled in the active site. 140
6.2 Radical SAM enzyme (CPH4) complexed with an Mg ion, arginine and surrounding waters. 141
6.3 An overview of how RSE's are computed. Absolute energies of both substrate and its radical are calculated on MD snapshots. In the above figure the conformation of the substrates changes to a more stable configuration after around 60ns. 142

List of Tables

1.1 λ_{max} values free base porphyin (monomer) and the Azo linked dimer. 28

2.1 A comparison of the orbitals present for the first and second row atoms using popular correlation consistent basis sets. 50

3.1 Computed excitation energies in eV with CIS and TDDFT with different exchange-correlation functionals. [a]Experimental data from references[1,5,6] 89

3.2 Atoms frozen in the geometry optimisations. 91

3.3 Computed pathways model ($|V_{DA}^{Beratan}|$) and TDDFT ($|V_{DA}^{TDDFT}|$) coupling values in eV for Trp14 $\rightarrow$ heme electron transfer for different structural models. 3AA and 1AA indicate three and one intervening amino acid residue included in the calculation, see text for details. The average values are evaluated using seven different crystal structures with the Trp14 + Heme + 3AA structural model. 94

3.4 Computed pathways model ($|V_{DA}^{Beratan}|$) and TDDFT ($|V_{DA}^{TDDFT}|$) coupling values in eV for Trp7 $\rightarrow$ heme electron transfer for different structural models. 3AA and 1AA indicate three and one intervening amino acid residue included in the calculation, see text for details. The average values are evaluated using seven different crystal structures with the Trp7 + Heme + 3AA structural model. 94

3.5 Computed TDDFT (CAM-B3LYP/6-31G*) coupling values in eV for ET and EET for the Trp + Heme with three intervening amino acid residues model. 96

3.6 Calculated relaxation times in ps. For τ_{ES} the $|V_{DA}|$ is computed using TDDFT for the S_1 state, for τ_{GS} $|V_{DA}|$ is computed using TDDFT for the S_0 state and for τ_{hybrid} $|V_{DA}|$ is evaluated using the pathways tunnelling model. 98

4.1 Descriptors for M, R and R' for atomic labels in figure 4.5. 106

4.2 Experimentally observed excitation energies of free base porphyrin in eV. 107

4.3 Calculated excitation energies of free base porphyrin using various levels of theory. Results presented are gas phase and in eV. 107

4.4 λ_{max} values in eV of the Q and B band for free base porphyrin calculated using a variety of functionals. [a]Experimental data from Ref.[1] . 112

4.5 Excitation energies (eV) of the Q band absorption max and the emission maximum 116

4.6 Calculated Stokes Shifts in eV. [a]Experimental data from Ref.[7]. 116

5.1 Relevant normal modes (n) of the two electronic states for H_2P. δ is the dimensionless displacement. The three normal modes m'$_1$, m'$_2$ and m'$_3$ of S_0 (identified by a prime) on which the S_1 mode n is mainly projected, together with their weights $J(m'_x, n)^2$, where J is the Duschinsky matrix as demonstrated by Santoro et al[8]. Values = 0.000 refer to values <0.001. The frequency difference between n and the normal mode m'$_1$ ($\Delta\omega = \omega - \omega_{m'1}$). 134

5.2 Relevant normal modes of the two electronic states for ZnP. δ is the dimensionless displacement. The three normal modes m'$_1$, m'$_2$ and m'$_3$ of S_0 (identified by a prime) on which the S_1 mode n is mainly projected, together with their weights $J(m'_x, n)^2$, where J is the Duschinsky matrix as demonstrated by Santoro et al[8]. Values = 0.000 refer to values <0.001. The frequency difference between n and the normal mode m'$_1$ ($\Delta\omega = \omega - \omega_{m'1}$). 136

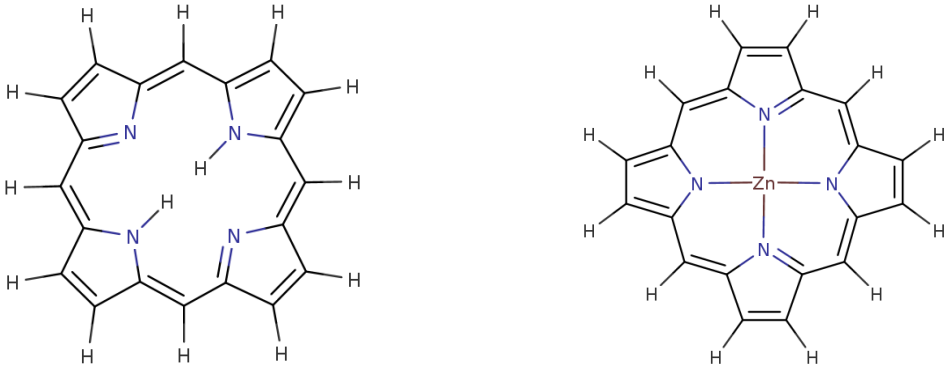
6.1 Price/Performance for MD using GROMACS and AMBER for BCL-XL. Costs are in £/ns (smaller is better); figures in brackets are ns/day (larger is better) 143

Chapter 1

Introduction

1.1 Porphyrin Background

Nature is an architect and a remarkably good one. Millions of years of evolution have allowed failed designs to be discarded, augment those that worked and design the molecules of the future. A molecule that has been subjected to her ruthless testing is porphyrin, a molecule so well engineered that it has been referred to as the 'pigment of life' and arguably responsible for life as we know it.



(a) Free Base Porphyrin (b) Metalloporphyrin (zinc)

Figure 1.1: Free Base and Metal Centred Porphyrin

The name porphyrin is derived from the ancient Greek word for the colour purple, which is fitting as the core porphyrin system produces intense red and purple colours. A porphyrin unit is a large aromatic macrocycle which is an amalgamation of four pyrrole units connected by methine bridges as shown in figure 1.1. They

have the ability to form many different bonds from a variety of locations known as: *alpha*, *beta* and *meso*, shown in figure 1.2. A porphyrin that contains nothing in the centre is referred to as free base porphyrin[1] with its stable configuration having D_{2h} symmetry.[9]

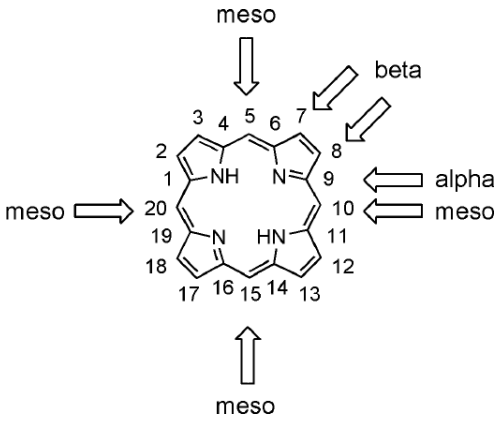


Figure 1.2: IUPAC numbering of carbons in porphyrin with the *meso*, *alpha* and *beta* positions labelled.

The centre of a porphyrin molecule is large enough to accommodate a metal atom facilitated by the the presence of delocalised electrons in the system to achieve a metal-to-porphyrin bond. The energies of the orbitals of both the metal and the porphyrin are similar which allows not only for porphyrin-to-metal charge transfer but metal-to-porphyrin backbonding as well.[10] A porphyrin containing a metal in the centre is referred to as a metalloporphyrin and is widely regarded having its stable configuration in the D_{4h} symmetry group.[9] The reason for the change in symmetry is as a result of removing two central hydrogens and replacing them with one metal which is connected to four nitrogens, which increases the symmetry.[1,11]

Aromaticity

Porphyrins are highly-conjugated cyclic molecules which contain 26 π-electrons. As they share similar properties they can be directly compared to the quintessential aromatic benzene. Like benzene porphyrin obeys Hückel's rule of aromaticity (4n+2 delocalised π-electrons) which states that *In order to be aromatic, a molecule must*

have a certain number of π electrons (electrons with π bonds, or lone pairs within p orbitals) within a closed loop of parallel, adjacent p orbitals.[12] Although they have 26 π electrons, only 22 electrons are included in a single route of delocalisation as shown in figure 1.3.[13]

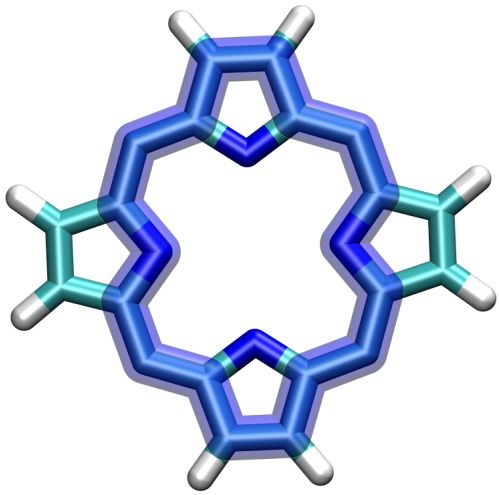


Figure 1.3: Electron delocalisation route of porphyrin shown in blue.

1.2 Porphyrins in nature

There are a multitude of porphyrin configurations available due to the great variety of substituents and binding locations on the subunit. This has led to porphyrin being incredibly useful in a wide range of biological systems and the obvious choice as Nature's catalyst. The sheer amount of opportunities has allowed Nature to engineer porphyrins to implement many of her grandest designs.

A few commonly found porphyrinoid complexes ubiquitous in nature are: The heme group in hemo- and myoglobin used for oxygen transport in blood,[14,15] Vitamin B_{12} (cobalamin) which is a water-soluble vitamin that has a key role in the normal functioning of the brain and nervous system[16] and chlorophyll that is vital towards energy harvesting from light[17] (shown in figure 1.4). These are all of great importance for many biological processes to function correctly. Even though all these

compounds share the same core porphyrin backbone they all behave differently and and display unique and vibrant colours. Electronic excitations are responsible for mediating these colour changes and why these complexes have a range of different colours in the visible region of the electromagnetic spectrum.[18] Seemingly insignificant alterations to the structure have dramatic impact on both the colour and the function of the system. This phenomenon is investigated in Chapter 4.

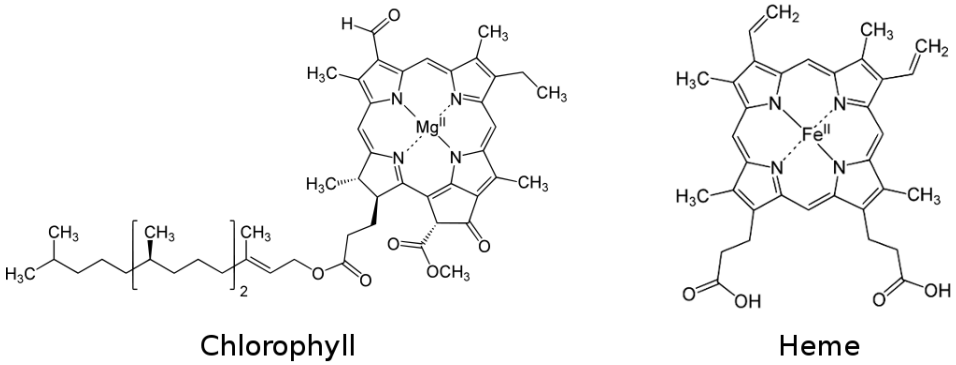


Figure 1.4: Skeletal structures of chlorophyll and heme shown together to highlight similarities.

The heme group is responsible for the intense red colour of blood. Deoxygenated hemoglobin is a darker red than oxygenated hemoglobin as the amount of oxygen bound to the iron has an impact on the conjugated system.[15] Chlorophyll is responsible for the green appearance of plants (the name itself means green leaf, chloros is Greek for green and phyllon means leaf).[19] A close look at the heme and chlorophyll molecules shows that the main difference between the two is that chlorophyll has a reduction in conjugation and it is this seemingly insignificant factor that is responsible for the change in colour.

1.3 Uses and applications of porphyrins

The uses for porphyrins are diverse and numerous. Traditionally, owing to the fact porphyrins display such vibrant colours, they were used simply as dyes. In more

recent years, they still find purpose as dyes only now with a more modern focus being utilised in dye sensitised solar cells (DSSC's) serving to maximise visible light absorption.[20,21] The optical limiting properties of porphyrins are exploited to make effective limiters of light which is useful in making photo-chromic glass.[22] The uses of porphyrins are extended to medicinal applications which range from being used to combat drug resistant[14] pathogens to being used to sterilise equipment. The main focus for porphyrin based medical research is as a photo sensitiser in photo dynamic therapy (PDT) which is used as a treatment for cancer.[23–25]

PDT works by delivering the porphyrin to the target tissue 'exciting' it to a singlet state which then crosses to a stable triplet state which destroys the harmful cells. Since porphyrins can be excited with visible light, their use in PDT has gathered significant attention as use of more harmful x and γ rays are no longer required.[26] This is advantageous over traditional chemotherapy methods as it is site specific meaning only cancerous cells are targeted.

1.4 Spectroscopic characteristics

Spectroscopy provides the foundation for many important analytical techniques in science and technology designed to determine various properties of matter in the solid, liquid and gas phases. Electronic spectroscopy, where electron transitions are studied, is one of these techniques which probes the absorption and emission of light by matter providing information on its structure, composition, and dynamics. The range of systems that can be studied using electronic spectroscopy extends from inorganic compounds to biologically relevant molecules, nanostructures, the Earth's atmosphere and distant galaxies. Spectroscopic experiments are complemented by the theoretical spectroscopy that exploits the predictive power of physical models and facilitates interpretation of experimental results. Of particular importance is the ability to accurately predict spectra theoretically since this provides a route to

computer-aided design of materials with desirable properties, it could also help to significantly decrease the number of laboratory experiments to be carried out.

Porphyrins undergo photo-induced electronic excitations. Visible light will cause an excitation from an occupied orbital to an unoccupied orbital in the electronic ground state to form an electronic excited state.[27] A common approach for studying the excited states of porphyrin systems is ultraviolet/visible (UV/Vis) spectroscopy which is a quick, effective and relatively easy technique. Free base porphyrin and metalloporphyrin display similar spectra. However, the Q band region of the spectra acts as a fingerprint and shows the distinguishable characteristics.

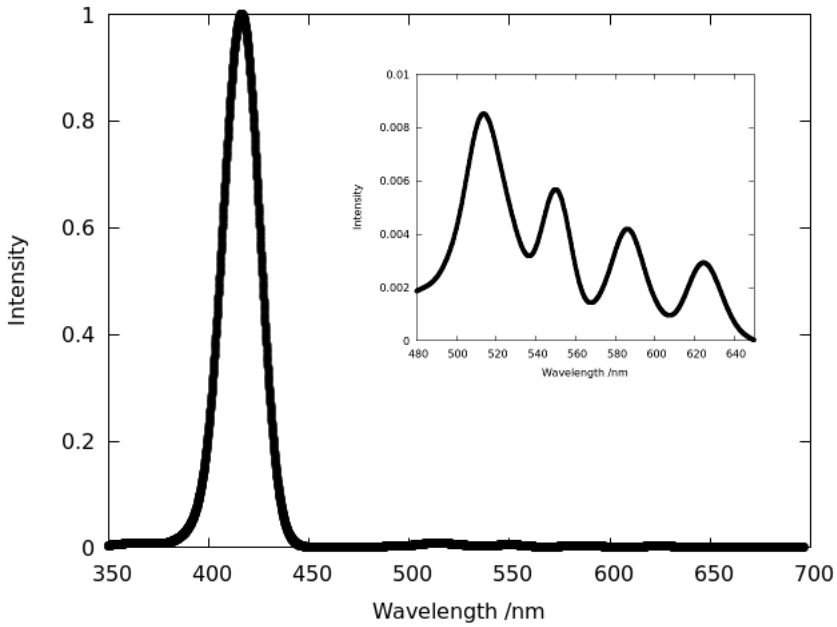


Figure 1.5: Typical UV-Vis spectra for free base porphyrin. A magnified view of the Q-band is shown inset.[1]

Figures 1.5 and 1.6 show the typical UV-Vis spectra for porphyrin complexes which lie between 350-700 nm (1.5eV – 3.5eV) and have two distinct regions: a strong peak at around 400-450 nm, arising from an excitation of S_0 to S_2 transition, known as the B band (Soret) and a weak broader peak Q band around 550-600 nm which arises from a S_0 to S_1.

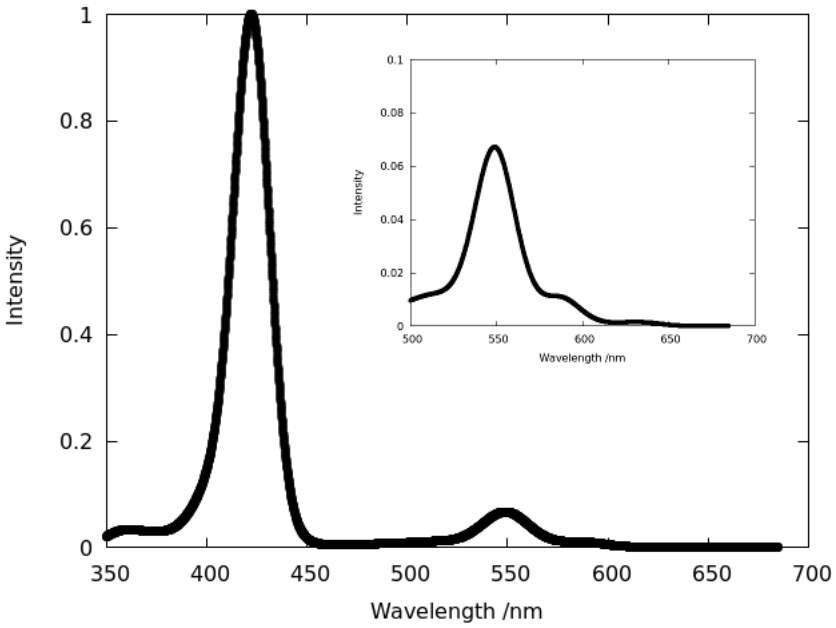


Figure 1.6: Typical UV-Vis spectra for a metal centred porphyrin.[2]

The internal conversion from S_2 to S_1 is very rapid and thus fluorescence is only detectable from the S_1 state. The B and Q-bands arise from π to π^* transitions. The Gouterman four orbital model suggests the presence of frontier molecular orbitals, two HOMO (π, highest occupied molecular orbitals) and two LUMO (π^*, lowest unoccupied molecular orbitals, figure 1.7). The two π orbitals are of almost the same energy, resulting in two almost equal absorption bands, namely the a_{1u} to e_g and a_{2u} to e_g (figure 1.7). The electron-electron configuration interaction that mixes the two transitions splits them giving rise to the two bands, the B and Q band. The former being intense and of short wavelength and the latter being weak and of long wavelength.

1.4.1 Gouterman's four orbital model

The key difference between the spectra for free base and metalloporphyrin are the peaks in the Q band region. Free base porphyrin has four peaks while metalloporphyrin only has two which is caused by the differences in symmetry between the two systems. Owing to the higher symmetry in metalloporphyrin the two lowest

unoccupied molecular orbitals (LUMO) are degenerate, seen in figure 1.7. This is explained by Gouterman's four orbital model[6]. The B and Q bands are a result of π to π^* transitions. Gouterman's four orbital model postulates that the two highest occupied molecular orbitals (HOMO, π) and the two LUMOs (π^*) are responsible for and can explain the spectroscopic characteristics seen in the UV/Vis spectra.[6]

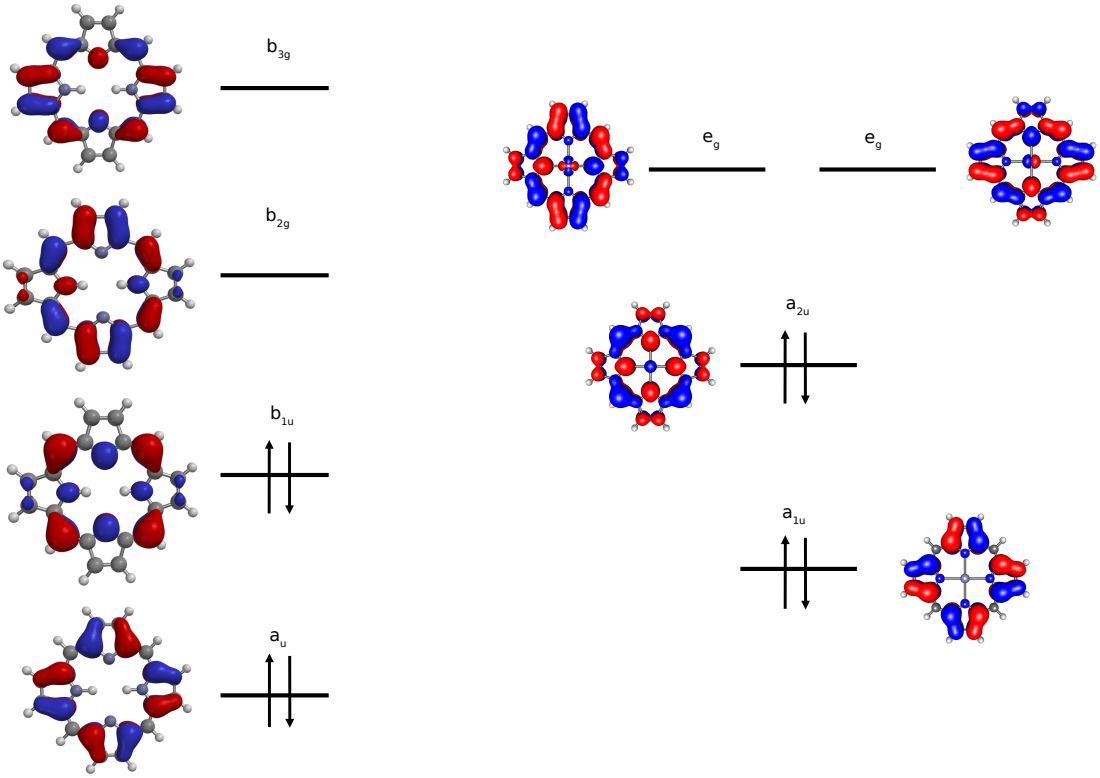


Figure 1.7: Molecular orbitals for free base porphyrin (left) and metalloporphyrin (right).

Free base porphyrin has a lower symmetry (D_{2h}) than metalloporphyrin which means the two LUMOs are no longer degenerate and split into x and y symmetric orbitals. The four bands are now observed as four different non degenerate transitions are possible. This also accounts for the lowering of the intensity and broadening of the B band. The degeneracy of the LUMOs for metalloporphyrin means that only two non-degenerate transitions occur.

1.4.2 Molecular orbitals

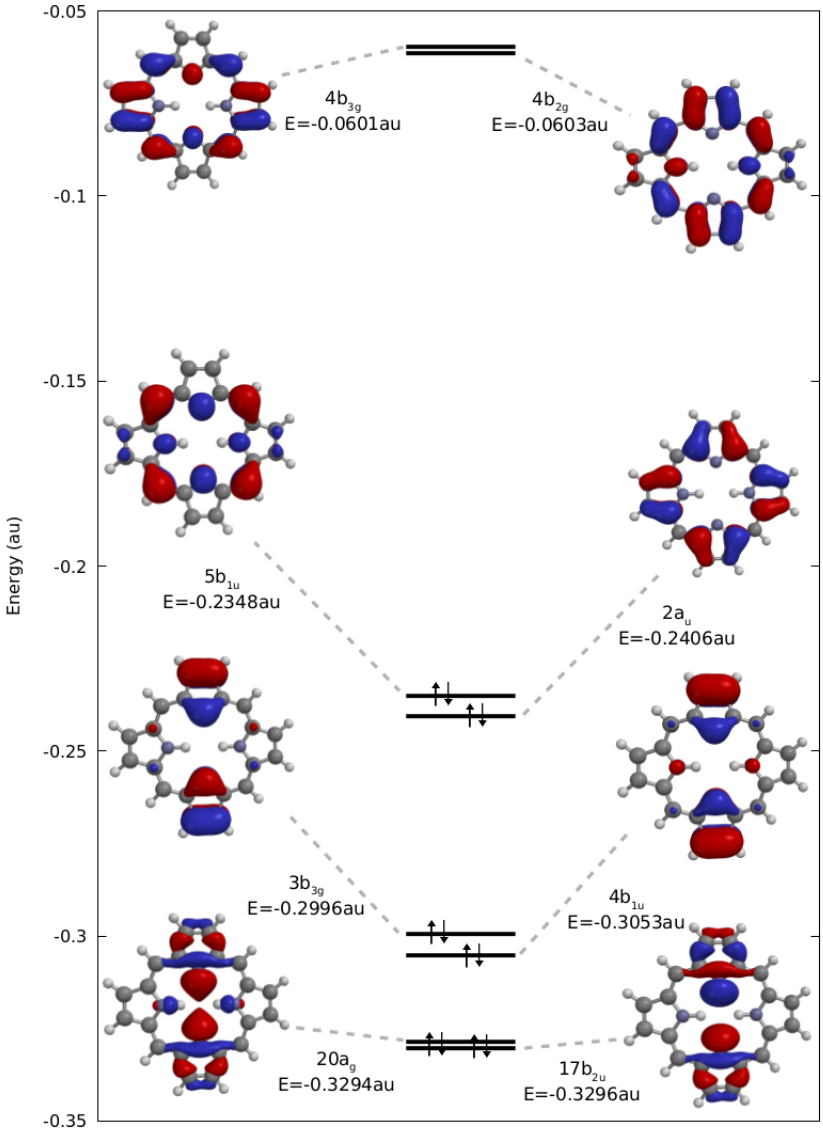


Figure 1.8: Molecular orbital diagram for free base porphyrin. Values are obtained from a time dependent density functional theory calculation (6-311G*/CAM-B3LYP) which is explained in greater detail later.

Figure 1.8 shows the two LUMOs and the six HOMOs for free base porphyrin with calculated energies included. The two LUMOs are very close in terms of energy but importantly are not degenerate. This means that there will be four peaks in the Q band as predicted by Gouterman's four orbital model. The LUMOs have a similar form, except they are split into symmetric x and y polarised orbitals. As a consequence of this, it is possible to trace the transitions and predict which orbitals give rise to the individual peaks in the visible spectrum. This is why six HOMOs

have been included in the MO diagram, since even though they do not give rise to significant peaks they still do have transitions which should not be ignored. It can help further the understanding of porphyrin complexes knowing which orbitals give rise to which peaks but it can also be of benefit to understand which orbitals do not give rise to significant peaks and why. Centrosymmetric (molecules with an inversion centre) molecular transitions must satisfy Laporte's rule. This only allows gerade (g) to ungerade (u) or ungerade to gerade transitions and all other transitions such as g to g or u to u are forbidden. However, there are exceptions to this rule. Porphyrins are large aromatic molecules and sometimes the centre of inversion can become distorted and be subject to vibronic effects permitting these transitions to be weakly allowed. As a result of this, formally forbidden transitions can be observed with very low oscillator strength, which is the focus of Chapter Five.

1.5 Porphyrin Varieties

As described earlier, two possible reaction positions are available for porphyrins. Substitution at both the *meso* and *beta* positions (shown in figure 1.2) will both perturb the electron distribution for the porphyrin molecule.[28,29] Out of the two reaction sites, the *meso* substituents display a larger perturbation than the *beta* due to them being more directly involved in the conjugation of the system, which has a larger effect on the frontier orbitals.[28] The idea that porphyrins can be *'tuned'* drives the study of substituent addition to porphyrins and the effect they have on both the electronic structure and absorbance.[30]

There are a variety of different possible configurations due to the numerous binding sites present in porphyrins. To make discussion easier, individual phenyl substituents have been labelled: *A, B, C* and *D* (see figure 1.9). For free base porphyrin it is possible make more unique structures than metalloporphyrin due to the increase in symmetry. Analysis of each different configuration was performed in

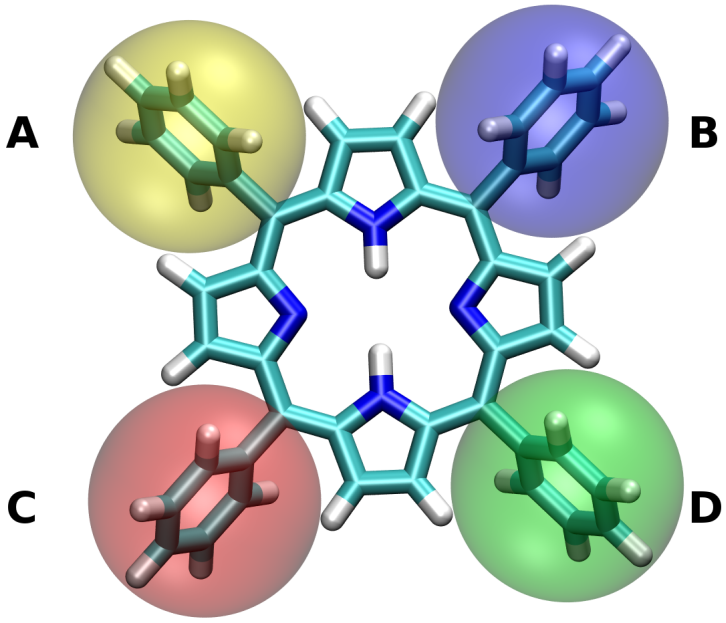


Figure 1.9: Free base porphyrin with substiuents highlighted and labelled.

order to see how addition of substituents affected the spectroscopic properties. It was seen that for both free base porphyrin and metalloporphyrin there is a direct relationship between substituents and the absorbance spectrum. Additional phenyl groups on the core porphyrin unit results in a greater bathochromic (red) shift.

It is also possible to bridge two porphyrin monomer units together to create a porphyrin dimer (figure 1.10). These have attracted significant attention as they share many properties with the monomer, but importantly have the ability to be tuned to a greater degree than the monomer.[22] In relation to porphyrins use in photo dynamic therapy, having an absorbance that is greater than 650 nm is desirable since it means the target porphyrin is excited and not another biologically active porphyrin that is native to the body such as the heme group found in the globin proteins. To reach these longer wavelengths by addition of substituents is possible with the monomer but becomes difficult and expensive. Hence without altering the porphyrin structure, connecting two monomers to create a dimer is a particularly promising solution to achieving these desired wavelengths.[23,31]

The spectrum of the porphyrin dimer (figure 1.11) shows it has a split B band

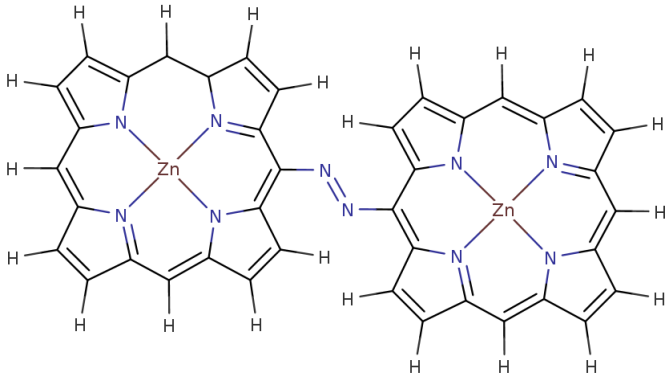


Figure 1.10: An Azo linked zinc porphyrin dimer.

with two peaks B_1 and B_2 at around 400 nm and 460 nm. The intensity of these two peaks are around half that of the monomer but are much broader. The Q band is more intense in the dimer in comparison with the monomer and again like the B band is much broader, it is also shifted into the 600 - 700nm region. An overall bathochromic shift (red shift) is observed.[22]

The splitting of the B band can be explained by the exciton model proposed by Kasha for molecular spectroscopy.[32] The Q band of the monomer has two orbitals which account for the different peaks, an orbital polarised along the x axis and an orbital polarised along the y axis which is shown in figure 1.7 by the two degenerate e_g orbitals of metalloporphyrin. This is now effectively seen in the B band of the dimer for which there are two bands B_x and B_y. The B_2 band peak has a bathochromic shift and occurs at around 450 nm. This is a result of both transitions being parallel along the meso-meso bond. The transitions in the B_1 band are 90° to each other.[33]

System	λ_{abs}/nm **(Relative Intensity)**
Monomer	415 (0.9), 542 (0.04)
Dimer	420 (0.44), 456 (0.42), 559 (0.1)

Table 1.1: λ_{max} values free base porphyin (monomer) and the Azo linked dimer.

When connecting multiple monomers together a consideration towards how the units are bridged is necessary. A variety of bridges for the two monomers exist

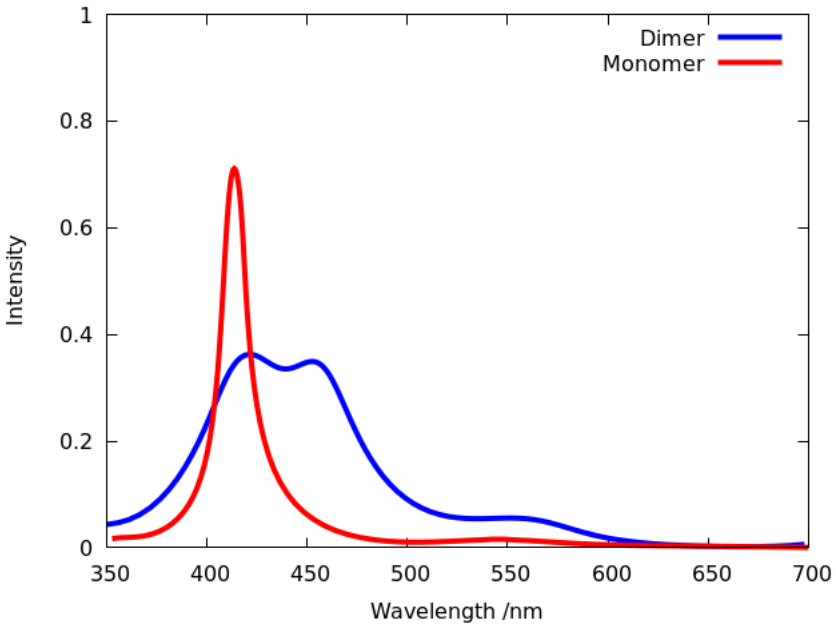


Figure 1.11: UV/vis spectra of porphyrin monomer and dimer. (Unpublished results from an experimental collaborator).

for which the azo linked porphyrin has been suggested as the most promising inter porphyrin communicator.[22,34] The orbitals of the nitrogens (sp^2) are sterically less demanding than those of alternatives such as carbon (sp^3) which allow for smaller dihedral angles between the two porphyrin monomers and also allows for good conjugation across the bridge. This feature becomes increasingly important with the application in DSSC's where a dynamic link between the two monomers that can alter depending on the intensity of light becomes beneficial.

1.5.1 Larger Units

Structures larger than the dimer are also of interest. Oligomers have recently been used to create effective molecular wires.[3,35] Charge transfer through porphyrins takes place by a phase coherent tunnelling mechanism which has a small distance dependance.[36,37] Large cyclic porphyrin polymers have been bound to carbon nanotubes in an effort to improve photovoltaic devices.[38] An alternative idea is to expand the monomer unit by addition of more pyrole units. A five membered pyrole unit known as sapphyrin[39] is generating interest as a porphyrin alternative due to its likeness

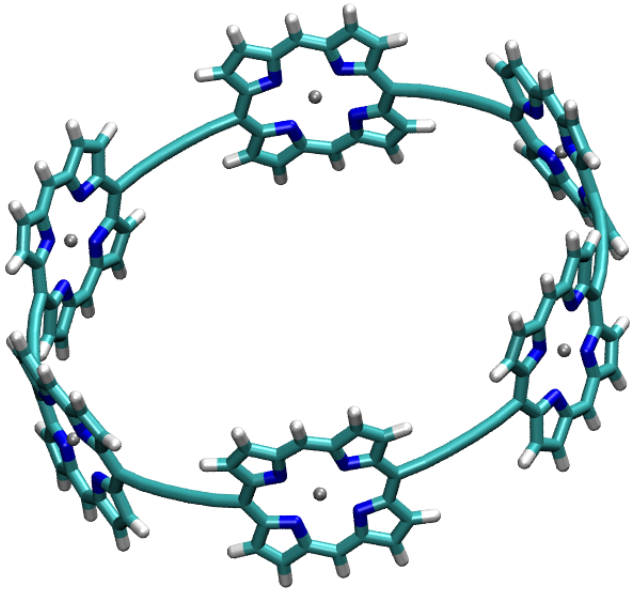


Figure 1.12: An example of just one of the many varieties of porphyrin oligomers (6-Porphyrin Nanoring)[3].

to porphyrins but with increased possibilities. Even larger expanded porphyrin systems consisting of six, seven and eight pyrole units have been recently synthesised.[40]

1.5.2 Scope of Thesis

The main scope of this thesis is to further the understanding of the electronic structure of porphyrins. A cloud of uncertainty remains around electron transfer in proteins. Porphyrin is ubiquitous throughout nature and facilatates most of the electron transfer reactions that occur within proteins. If this mechanism is fully understood and can be modelled accurately we can mutate proteins and design new systems solely on computers. This is very advantageous in the world of protein engineering.

A method of measuring charge transfer in proteins is by use of fluorescence spectroscopy. Heme proteins do not fluoresce making it impossible to distinguish the modes of charge transfer. By use of computational methods looking at the

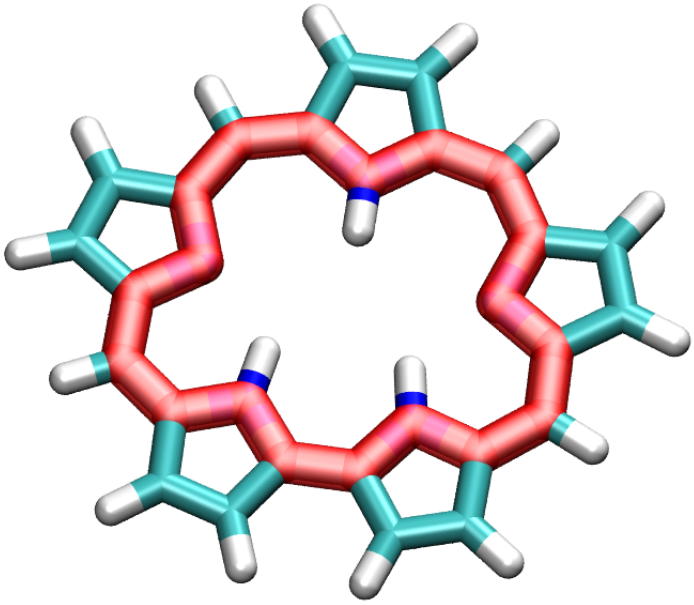


Figure 1.13: A sapphyrin molecule with five pyrole units and larger centre.

electronic configuration of the targeted systems it is possible to determine whether electrons decay by electron or excitation energy transfer. This requires high level theoretical calculations on large systems which is enabled by use of sophisticated excited state methods which are computationally feasible on large systems. The individual electronic excitations and vibronic bands are studyed in order to resolve the individual peaks of absorption spectra at high resolution.

Chapter 2

Overview of Theoretical Methods

2.1 Theoretical and Computational Methods

This chapter aims to introduce and describe all methods used during this work, both computationally and theoretically.

2.2 The Quantum Mechanical Approach

The fundamental equation in quantum mechanics is the time dependent Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\Psi = \hat{H}\Psi \tag{2.1}$$

here: $\hbar$ is the reduced Planck constant ($\hbar = \frac{h}{2\pi}$), t is time, Ψ is the wavefunction, $\hat{H}$ is the Hamiltonian operator and i describes the imaginary unit.

This is applicable for time dependent systems. For time independent systems, the stationary states can be described using the time independent Schrödinger equation.

$$\hat{H}\Psi = E\Psi \tag{2.2}$$

where E is the total energy of the system.

The solutions of the time independent Schrödinger equation for a system give the stationary state wavefunctions, $\Psi_n(\mathbf{r})$, and energies E_n which correspond to the different energy levels of a system. The lowest energy solution is known as the ground state whilst higher energy solutions are referred to as excited states. In order to solve equation 2.2 the Hamiltonian operator, $\hat{H}$, is required.

2.2.1 The Hamiltonian

The Hamiltonian consists of kinetic ($\hat{T}$) and potential ($\hat{V}$) energy operators.

$$\hat{H} = \hat{T} + \hat{V} \tag{2.3}$$

For a molecular system, neglecting relativistic effects, the kinetic and potential energy terms can be written (in atomic units):

$$\hat{T} = -\frac{1}{2}\sum_{A}\frac{1}{m_A}\bigtriangledown_A^2 - \frac{1}{2}\sum_{i}\bigtriangledown_i^2 \tag{2.4}$$

$$\hat{V} = \sum_{A>B}\frac{Z_A Z_B}{r_{AB}} - \sum_{A}\sum_{i}\frac{Z_A}{r_{Ai}} + \sum_{i>j}\frac{1}{r_{ij}} \tag{2.5}$$

The operator $\hat{T}$ (equation 2.4) describes the kinetic energy where A denotes the nuclei in the system, i to an electron in the system and m_A to the mass of the nucleus A. The electron mass is equal to one as atomic units are used.

The operator $\hat{V}$ (equation 2.5) denotes the potential energy. Here A and B denote nuclei in the system, i and j describe electrons and Z_A is the charge of the nucleus A.

A combination of equations 2.4 and 2.5 describes the complete molecular hamiltonian shown in equation 2.6.

$$\hat{H} = -\frac{1}{2}\sum_{A}\frac{1}{m_A}\bigtriangledown_A^2 - \frac{1}{2}\sum_{i}\bigtriangledown_i^2 + \sum_{A>B}\frac{Z_A Z_B}{r_{AB}} - \sum_{A}\sum_{i}\frac{Z_A}{r_{Ai}} + \sum_{i>j}\frac{1}{r_{ij}} \tag{2.6}$$

2.2.2 The Born-Oppenheimer Approximation

Already discussed is that relativistic effects are ignored as the inclusion in results for the systems studied in this thesis are often negligible. Similarly, since the nuclei are significantly heavier than the electrons it is possible to use the Born-Oppenheimer Approximation[41] in order to describe the system. This assumes that the motion of the nuclei is so slow relative to the electrons that they may essentially be treated as fixed in position with respect to the electrons. This allows the total wavefunction to be split into a nuclear and electronic wavefunction.

$$\Psi_{total}(\mathbf{R};\mathbf{r}) = \Psi_{electronic}(\mathbf{R};\mathbf{r})\Psi_{nuclear}(\mathbf{R}) \tag{2.7}$$

where $\mathbf{R}$ and $\mathbf{r}$ correspond to the spatial coordinates of the nuclei and electrons respectively.

The electronic $\hat{H}$ can now be described where the kinetic energy of the nuclei is zero and the nuclear-nuclear repulsion is a constant ($\hat{V}_{nn}$). A separate Hamiltonian term $\hat{T}_n$ is now used to describe the kinetic energy of the nuclei.

$$\hat{H}_{total} = \hat{H}_e + \hat{T}_n + \hat{V}_{nn} \tag{2.8}$$

where:

$$\hat{H}_e = -\frac{1}{2}\sum_{i}\bigtriangledown_i^2 - \sum_{Ai}\frac{Z_A}{r_{Ai}} + \sum_{i>j}\frac{1}{r_{ij}} \tag{2.9}$$

$$\hat{T}_n = -\frac{1}{2}\sum_{A}\frac{1}{m_A}\bigtriangledown_A^2 \tag{2.10}$$

This now allows the electronic Shrödinger equation to be described as:

$$\hat{H}_e \Psi_e(\mathbf{R};\mathbf{r}) = E_e(\mathbf{R})\Psi_e(\mathbf{R};\mathbf{r}) \tag{2.11}$$

and its nuclear counterpart,

$$(\hat{T}_n + E_e(\mathbf{R}))\Psi_n(\mathbf{R}) = E_{tot}\Psi_n(\mathbf{R}) \tag{2.12}$$

The electronic Schrödinger equation (2.11) can be used to find and determine the potential energy surface (PES) of a molecular system. Often the focus is on energy minima, transition states and low energy pathways. The energy minima correspond to the nuclear positions that result in stable structures of a system. These low energy minima can be used to study chemical reactions.

The electronic Schrödinger equation theoretically will yield exact results for the electronic energy of a molecular system with considerations to the approximations. The equation can be solved exactly for one electron systems. However for any system with more than one electron an exact solution to the electronic Schrödinger equation (2.11) becomes impossible. This is due to the electron-electron repulsion term which explicitly couples the motion of the electrons. However, it is possible to obtain approximate solutions. The field of quantum chemistry is concerned with finding approximate solutions to the electronic Schrödinger equation. There are many different approaches that yield solutions of varying levels of accuracy, and the methods relevant to the work presented in this thesis are described in the following sections.

2.3 Hartree-Fock Theory

Hartree-Fock (HF) theory is an *ab initio* (from first principles) method for solving the electronic Schrödinger equation (equation 2.11) for a many electron system.[42,43]

In Hartree-Fock theory the wavefunction is approximated by a single Slater determinant.[44]. The total wavefunction obeys the Pauli principle that states that *the wavefunction must be antisymmetric with respect to the interchange of any two electrons*. This means the wavefunction must change sign with exchange of electrons. As an example in a two electron system one can assume that combining the product of single particle wavefunctions with their respective spatial coordinates will give the correct electronic wavefunction (equation 2.13 which is known as the Hartree product. However, expressing the wavefunction in this form is incorrect since the electronic wavefunction is not antisymmetric with respect to electron exchange. In order to satisfy the Pauli principle it is necessary to construct the overall electronic wavefunction as a linear combination of the original Hartree product and the exchanged electron Hartree product (equation 2.16).

$$\Psi_H(\mathbf{1},\mathbf{2}) = \psi_1(\mathbf{1})\psi_2(\mathbf{2}) \tag{2.13}$$

which does not satisfy the Pauli principle since:

$$\Psi_H(\mathbf{1},\mathbf{2}) = \Psi_H(\mathbf{2},\mathbf{1}) \tag{2.14}$$

However the HF wavefunction does satisfy the condition

$$\Psi(\mathbf{1},\mathbf{2}) = -\Psi(\mathbf{2},\mathbf{1}) \tag{2.15}$$

$$\Psi_e(\mathbf{1},\mathbf{2}) = \frac{1}{\sqrt{2!}}\left\{\psi_1(\mathbf{1})\psi_2(\mathbf{2}) - \psi_1(\mathbf{2})\psi_2(\mathbf{1})\right\} \tag{2.16}$$

This wavefunction can also be expressed in terms of a 2x2 determinant

$$\Psi_e(\mathbf{1},\mathbf{2}) = \frac{1}{\sqrt{2!}}\begin{vmatrix}\psi_1(\mathbf{1}) & \psi_2(\mathbf{1})\\ \psi_1(\mathbf{2}) & \psi_2(\mathbf{2})\end{vmatrix} \tag{2.17}$$

This can be generalised to N electron systems leading to the determinant presented in Equation 2.17 (two electron system) which can be expanded to an N

electron system as shown in Equation 2.18 where N is the number of electrons, ψ_1 through to ψ_N are one electron molecular orbitals and (**1**) to (**N**) are the spatial coordinates of the electrons.

$$\Psi_{HF} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{1}) & \psi_2(\mathbf{1}) & ... & \psi_N(\mathbf{1}) \\ \psi_1(\mathbf{2}) & ... & ... & ... \\ ... & ... & ... & ... \\ \psi_1(\mathbf{N}) & \psi_2(\mathbf{N}) & ... & \psi_N(\mathbf{N}) \end{vmatrix} \tag{2.18}$$

The one electron molecular orbitals in the Slater determinant are known as the *spin* orbitals which are the products of a spatial orbital and a spin function α or β with a spin coordinate ω. For the i-th molecular orbital depending on the coordinates of the k-th electron the form of the molecular orbital is shown equation in 2.19.

$$\psi_i(\mathbf{x}_k) = \psi_i^{\alpha}(\mathbf{r}_k)\alpha(\omega_k) \tag{2.19}$$

where ($\mathbf{r_k}$) describes the spatial coordinates of electron k in spatial orbital ψ_i^{α} with the spin coordinate ω_k.

The atomic or molecular energy of a system can be expressed in terms of the total wavefunction, Ψ. It is calculated by integrating the product of the complex conjugate of the wavefunction Ψ^* with respect to the three spatial coordinates and one spin coordinate for each electron. This leads to the time independent Schrödinger equation (equation 2.2) to take the form seen in equations 2.20 and 2.21. Here $d\tau$ represents the integration over all space and the spin coordinate.

$$E \int \Psi^* \Psi d\tau = \int \Psi^* \hat{H} \Psi d\tau \tag{2.20}$$

$$E = \frac{\int \Psi^* \hat{H} \Psi d\tau}{\int \Psi^* \Psi d\tau} \tag{2.21}$$

If the wavefunctions are normalised $\int \Psi^* \Psi d\tau = 1$ then equation 2.21 becomes

$$E = \int \Psi^* \hat{H} \Psi d\tau \tag{2.22}$$

or expressed in Dirac's notation,

$$E = \left\langle \Psi | \hat{H} | \Psi \right\rangle \tag{2.23}$$

Now the electronic Hamiltonian (equation 2.9) can be separated into one and two electron terms giving one and two electron Hamiltonians, $h(i)$ and $v(i,j)$ where:

$$\hat{h}(i) = -\frac{1}{2} \bigtriangledown_i^2 - \sum_A \frac{Z_A}{r_{Ai}} \tag{2.24}$$

$$\hat{v}(i,j) = \frac{1}{r_{ij}} \tag{2.25}$$

The nuclear-nuclear term (V_{nn}, which is a constant due to the Born-Oppenheimer approximation) can be added to form the total Hamiltonian for an N-electron system (equation 2.26). It is now possible to treat the one-electron, two-electron and nuclear repulsion terms separately.

For the one electron term for an electron i at position $\mathbf{r}$ the energy E_i is simply evaluated as a sum of one electron integrals and the total one electron energy is evaluated as a sum of one-electron integrals given in equation 2.27.

$$\hat{H}_e = \sum_i^N h(i) + \sum_{i>j}^N v(i,j) + V_{nn} \tag{2.26}$$

$$E_i = \langle \psi_i | h(i) | \psi_i \rangle = \int \psi_i^*(\mathbf{r}) \left(-\frac{1}{2} \bigtriangledown_i^2 - \sum_A \frac{Z_A}{r_{Ai}} \right) \psi_i(\mathbf{r}) d\mathbf{r} \tag{2.27}$$

For the two electron terms with electrons i and j with positions $\mathbf{r_1}$ and $\mathbf{r_2}$ the energy becomes more complicated to calculate. It is the sum of the Coulomb energy and the exchange energy, where the former is evaluated as a sum of Coulomb integrals and the latter is a sum of exchange integrals. The Coulomb integral takes

the form

$$(ii|jj) = \int\int \frac{\psi_i^*(\mathbf{r_1})\psi_i(\mathbf{r_1})\psi_j^*(\mathbf{r_2})\psi_j(\mathbf{r_2})}{r_{12}} d\mathbf{r_1} d\mathbf{r_2} \tag{2.28}$$

This represents a positive contribution of energy (destabilisation) of a system due to electron-electron Coulombic repulsions. Equation 2.28 represents an electron pair ij where an electron i at position $\mathbf{r}_1$ interacting with electron j at position $\mathbf{r}_2$. The exchange integrals are more complicated and take the form

$$(ij|ji) = \int\int \frac{\psi_i^*(\mathbf{r_1})\psi_j(\mathbf{r_1})\psi_j^*(\mathbf{r_2})\psi_i(\mathbf{r_2})}{r_{12}} d\mathbf{r_1} d\mathbf{r_2} \tag{2.29}$$

Unlike the Coulomb integrals, the exchange integrals do not have a classical analogy and instead corresponds to the energy gained by making the wavefunction antisymmetric. It is now possible to write the HF energy as equation 2.30 for both one and two-electron terms.

$$E = \sum_i^N E_i + \sum_{i>j}^N \{(ii|jj) - (ij|ij)\} + V_{nn} \tag{2.30}$$

The Coulomb and exchange integrals can be rewritten in terms of their respective operators $\hat{J}$ and $\hat{K}$, as described in equations 2.31 and 2.32.

$$\hat{J}\psi_i(\mathbf{r_1}) = \sum_j \int \frac{\psi_j^*(\mathbf{r_2})\psi_j(\mathbf{r_2})}{r_{12}} d\mathbf{r_2}\psi_i(\mathbf{r_1}) \tag{2.31}$$

$$\hat{K}\psi_i(\mathbf{r_1}) = \sum_j \int \frac{\psi_j^*(\mathbf{r_2})\psi_i(\mathbf{r_2})}{r_{12}} d\mathbf{r_2}\psi_j(\mathbf{r_1}) \tag{2.32}$$

Now a more convenient form of the HF energy from equation 2.30 can be utilised

$$E = \sum_i^N E_i + \frac{1}{2}\sum_i^N \left\langle \psi_i | \hat{J} - \hat{K} | \psi_i \right\rangle \tag{2.33}$$

The previous equations show that the HF energy has been determined through the one and two-electron terms. It would be advantageous to solve these equations

in a manner that is computationally efficient.

First it is worth noting that HF theory follows the variational principle, which states that *energy calculated from the form seen in (2.23) must be either greater than or equal to the lowest energy solution (true ground-state energy) of the molecule.* This means that the lower the energy calculated by the HF method the closer it is to its true value. The HF method aims to find the set of orbitals that minimise the energy.

The HF equations can be written as a set of one electron equations where:

$$\hat{F}\psi_i(r) = \varepsilon_i\psi_i(\mathbf{r}) \tag{2.34}$$

here ε_i are the orbital energies and $\hat{F}$ is the Fock operator. This is an effective one-electron operator that takes the form of the one electron operators encountered in equation 2.30. This gives

$$\hat{F}(1) = h(1) + \hat{J} - \hat{K} \tag{2.35}$$

The HF equations presented in equation 2.34 are known as pseudoeigenvalue equations. This means that they take the form of an eigenvalue equation but are known as a pseudoeigenvalue equation because the Fock operator is dependent on the wavefunction on which it acts. In this case the Fock operator acts on a wavefunction to generate an energy value multiplied by the wavefunction which is analogous to the time-independent Schrödinger equation (2.2).

The HF equations allows both the molecular orbitals and the energy levels to be calculated. The HF method begins with an initial estimation of the orbitals allowing the Fock operator to be constructed. The initial Fock operator can then operate on the estimated molecular orbitals to give an approximate solution for the

energy levels (ε_i) and hence new functions for improved molecular orbitals can be introduced. This process occurs over a series of iterations in an effort to minimise the energy until the energies and molecular orbitals converge to satisfy a certain degree of accuracy. The final pair of solutions is known as self consistent and allows the minimum energy to be obtained from an initial guess of molecular orbitals.

2.3.1 Linear Combination of Atomic Orbitals

Although it seems viable to calculate the properties of an atomic or molecular system it is not possible to directly calculate the HF equations as the final converged orbitals are based on the estimated orbitals. It is usual to represent the molecular orbitals as a linear combination of functions known as basis functions. This was first proposed by Roothaan and Hall[45,46] where the basis functions have been based on atomic orbitals. This method is known as the linear combination of atomic orbitals (LCAO) and the takes the following form

$$\psi_i = \sum_{\alpha}^{N_b} c_{\alpha i}\chi_\alpha \tag{2.36}$$

Where ψ_i describes the i-th molecular orbital in a system, $c_{\alpha i}$ is the coefficient corresponding to the α-th basis function of the molecular orbital and χ_α represents the α-th basis function. N_b is the total number of basis functions. These depend on the collection of specific basis functions being used known as a basis set which are described further in section 2.4.

It is possible to combine the HF equations with the LCAO approach to give the following

$$\sum_{\alpha}^{N_b} c_{\alpha i}\hat{F}\chi_\alpha = \varepsilon_i \sum_{\alpha}^{N_b} c_{\alpha i}\chi_\alpha \tag{2.37}$$

As the values $c_{\alpha i}$ are coefficients, $\hat{F}$ only operates on the basis functions χ_α. It

is possible to multiply equation 2.37 by the basis function χ_β

$$\sum_{\alpha=1}^{N_b} c_{\alpha i}\chi_\beta \hat{F}\chi_\alpha = \varepsilon_i \sum_{\alpha=1}^{N_b} c_{\alpha i}\chi_\beta\chi_\alpha \tag{2.38}$$

and then integrate,

$$\sum_{\alpha=1}^{N_b} c_{\alpha i}\left\langle \chi_\beta|\hat{F}|\chi_\alpha\right\rangle = \varepsilon_i \sum_{\alpha=1}^{N_b} c_{\alpha i}\left\langle \chi_\beta|\chi_\alpha\right\rangle \tag{2.39}$$

which can be rewritten as

$$\sum_{\alpha=1}^{N_b} c_{\alpha i}(F_{\beta\alpha} - \varepsilon_i S_{\beta\alpha}) = 0 \tag{2.40}$$

where

$$F_{\beta\alpha} = \left\langle \chi_\beta|\hat{F}|\chi_\alpha\right\rangle \tag{2.41}$$

and

$$S_{\beta\alpha} = \left\langle \chi_\beta|\chi_\alpha\right\rangle \tag{2.42}$$

here, $S_{\beta\alpha}$ describes the overlap between basis functions known as the overlap integrals. Equation 2.40 can be written in a matrix notation to give equation 2.43.

$$\mathbf{FC} = \mathbf{SC}\varepsilon \tag{2.43}$$

where $\mathbf{F}$ is the Fock matrix, $\mathbf{C}$ is the coefficient matrix containing the basis function coefficients, $\mathbf{S}$ is the basis function overlap matrix and ε is a diagonal matrix of the orbital energies of the system.

2.3.2 Fock Matrix

Equation 2.43 is solved using a self-consistent field approach. This approach has the following steps:

1. Obtain an initial guess of the orbitals using the coefficients $c_{\alpha i}$.
2. Calculate the necessary integrals for the system and form a Fock matrix and overlap matrix from these integrals.
3. Solve the Roothaan-HF equations (2.43 and 2.40) to give a new set of molecular orbitals.
4. Test for convergence by inspecting the density matrix. If the change is below a threshold then convergence has been met. If not procedure is repeated until the convergence is met.
5. If convergence has not been met, take the new set of molecular orbitals and return to step 2. Keep repeating until the convergence criteria are met.

Each iteration in a HF calculation solves the Roothaan-Hall equations (2.40). These form spatial molecular orbitals from N_b basis functions. This number of basis functions can vary from a minimum, which is equal to the number of doubly occupied molecular orbitals in a system, to effectively an infinite amount. The number of basis functions used can greatly affect the computational cost. Hence, it is necessary to use an appropiate basis set to satisfy the calculation.

The most computationally expensive part of the SCF procedure outlined above is the evaluation of the two-electron Coulomb and exchange integrals. This results in N_b^4 integrals in total which means Hartree-Fock calculations scale with the number of basis functions to the degree N_b^4. This results in the overall cost of the calculation scaling by 16 for either doubling the size of the system or for doubling the amount of basis functions.

Methods to speed up HF calculations exist. One successful method of interest looks at screening integrals so only the relevant ones are evaluated in order to improve computational efficiency. However, increasing the system size will always lead to a rise in computational cost. This scaling results in calculations on large systems becoming very computationally demanding meaning most HF calculations are restricted to relatively small molecular systems. Another computational limit that becomes apparent on larger systems calculations is the need for large amounts of computer memory. This work looks at pushing these boundaries and details many different methods for electronic structure calculations on large systems.

A criticism of HF theory is its inability to model electron-electron repulsion with sufficient accuracy. The repulsion between electrons is included by having each electron move in an average electronic field with respect to the other electrons in the system. This means the probability of an a electron being in one area in the system is only dependent on the average positions of other electrons. This allows electrons in HF theory to be in close proximity of each other which yields an unphysical representation. In a real system electron repulsion of one electron would cause the average electron cloud of the other electrons to change so that electrons would not occupy space so close to another. As a result of this HF theory is known to overestimate energies as it models electrons too close to each other, resulting in a large electron-electron repulsion value with respect to a real system. This error is known as the correlation energy and is expressed as

$$E_C = E_{EXACT} - E_{HF} \tag{2.44}$$

where E_C is the correlation energy, E_{EXACT} is the exact energy of the system within the basis set and E_{HF} is the calculated HF energy for a given basis set.

The correlation energy is a significant contribution and is too important to be neglected. Therefore HF theory alone is often not accurate enough to describe systems

and so a post-Hartree-Fock method (post-HF) such as second-order Møller-Plesset perturbation theory (MP2) or coupled cluster (CC) method must be employed. An alternative approach is a method known as density functional theory (DFT) which is discussed in greater detail in a later section 2.5.

The HF method outlined above seeks to find energies and wavefunctions that correspond to a closed shell ground state molecule. The resulting Slater determinant applies to molecules where the electrons fill the lowest energy levels first, doubly occupying the molecular orbitals with opposing spin electrons. Systems are in an electronically excited state when one or more electrons have been moved into a higher energy molecular orbital than its ground state occupancy. These electronically excited systems require different quantum mechanical calculation methods which are detailed in section 2.6.1.

2.3.3 Restricted and Unrestricted Hartree Fock Theory

There are two main methods of applying HF theory to a molecule. The first is to perform a restricted HF calculation (RHF). This studies closed-shell Slater determinants and restricts electrons of different spins (α and β) to the same set of spatial orbitals. The second method is known as an unrestricted HF calculation (UHF). This uses separate spatial orbitals for electrons with different spins leading to two sets of molecular orbitals (α and β) for a system. UHF is frequently used to describe open shell systems in conjunction with methods which describe excited states such as time dependent density functional theory (TDDFT, section 2.6.1) and the maximum overlap method (MOM, section 2.6.3). The HF method, alongside many other quantum chemical methods such as post-HF and DFT rely on the use of basis functions.

2.4 Basis Sets

Molecular orbitals can be described as a linear combination of atomic orbitals that are modelled by basis functions (equation 2.36). These are mathematical functions that vary in their accuracy and overall computational cost when used to model the molecular orbitals. A set of basis functions is known as a basis set.

There are several types of basis functions that are used in quantum chemical calculations, including Slater functions, Gaussian functions as well as plane waves. The most widely used basis functions for molecular calculations are Slater functions and Gaussian functions.

2.4.1 Slater and Gaussian Type Functions

Slater functions provide an accurate representation of the radial distribution of electrons that take the form $e^{-\alpha r}$ where α is a constant and r is the radial distance. Although Slater functions provide an accurate description of molecular orbitals, their use is limited due to the high computational demand they incur when evaluating the two electron integrals. Gaussian functions provide a convenient form for the evaluation of the two-electron integrals and are more commonly used. Gaussian functions have the form $e^{-\alpha r^2}$ and can provide an approximation to the more accurate Slater functions. However, a single function can differ significantly from the equivalent Slater function at short and long range and on its own is a poor representation of the radial behaviour of a molecular orbital.

By introducing the idea of contracted basis functions it is possible to model a Slater type function more accurately using a linear combination of Gaussian functions. A contracted Gaussian basis function can be written as

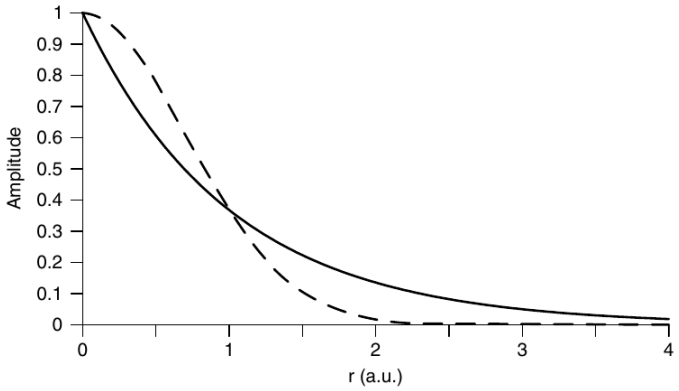


Figure 2.1: Behaviour of e^x where $x = r$ (solid line, STO) and $x = r^2$ (dashed line, GTO)

$$\chi(r) = \sum_P c_P e^{-\alpha_P r^2} \tag{2.45}$$

where c_P and α_P are constants relating to each Gaussian function. It is possible to illustrate how contracted basis functions more accurately model Slater orbitals, by comparing a Slater basis function with a single Gaussian (STO-1G) and the contracted three Gaussian (STO-3G) basis functions for a single hydrogen atom. The equations of each function are as follows where Φ denotes an atomic orbital arising from the basis functions.

$$\Phi_{Slater} = \left(\frac{\zeta^3}{\pi}\right)^{\frac{1}{2}} e^{-\zeta r} = 0.779 e^{-1.24r} \tag{2.46}$$

$$\Phi_{STO-1G} = \left(\frac{2\alpha_1}{\pi}\right)^{\frac{3}{4}} e^{-\alpha_1 r^2} = 0.370 e^{-0.417 r^2} \tag{2.47}$$

$$\Phi_{STO-3G} = 0.445 \left(\frac{2\alpha_1}{\pi}\right)^{\frac{3}{4}} e^{-\alpha_1 r^2} + 0.535 \left(\frac{2\alpha_2}{\pi}\right)^{\frac{3}{4}} e^{-\alpha_2 r^2} + 0.154 \left(\frac{2\alpha_3}{\pi}\right)^{\frac{3}{4}} e^{-\alpha_3 r^2} \tag{2.48}$$

$$\Phi_{STO-3G} = 0.0835 e^{-0.169 r^2} + 0.268 e^{-0.624 r^2} + 0.278 e^{-3.425 r^2} \tag{2.49}$$

By plotting the linear combination of Gaussian functions (STO-3G, described in the previous equations) it can be shown that they provide a much better approximation in comparison to the single Gaussian function (STO-1G) shown in figure 2.2.

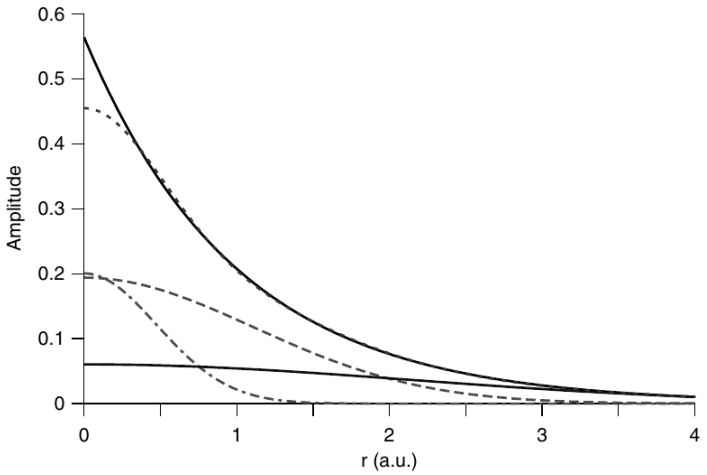


Figure 2.2: The radial behaviour of various basis functions in atom-centered coordinates. The bold solid line at top is the STO ($\zeta = 1$) for the hydrogen 1s function; for the one-electron H system, it is also the exact solution of the Schrödinger equation. Nearest to it is the contracted STO-3G 1s function that has been optimised to match the STO. It is the sum of the many single Gaussian functions. Note that from 0.5 to 4.0 a.u., the STO-3G orbital matches the correct orbital closely. However, near the origin there is a notable difference and, were the plot to extend to very large r, it would be apparent that the decay of the STO-3G orbital is more rapid than the correct orbital.

It is widely accepted that most chemical properties arise from the interactions of the valence orbitals and the core orbitals play a less significant role. This led to the development of split valence basis sets by Pople[47]. A split valence basis set describes the valence space with more basis functions than the core orbitals.

An example of a split-valence basis set is 3-21G. This basis set uses one basis function consisting of a linear combination of three Gaussian functions for the core orbitals and splits the valence orbitals into two parts, the inner and outer shell. The inner shell uses a contraction of two Gaussian functions whilst the outer shell consists of a single Gaussian function. The outer shell function has a smaller value of

α compared to the inner shell which leads to a longer range drop-off of the function allowing the valence orbitals to be more diffuse and have a longer spatial range. The basis set 3-21G is an example of a *double zeta (ζ)* basis set. Another popular variant is the *triple zeta* split-valence basis set which separates the valence shell into three parts.

2.4.2 Polarized and Diffuse Functions

Split-valence basis sets can be further improved by adding polarisation functions to the valence orbitals. Polarisation functions allow the electron distribution of the valence orbitals to become more polarised. This is achieved by the addition of functions that describe orbitals of higher angular momentum for the valence orbitals of atoms larger than helium. For Cartesian basis functions this means the six spatial d-orbitals are included. Double polarisation includes the p-shell spatial orbitals for hydrogen and helium atoms. This can be a significant addition to the calculation as hydrogen atoms are ubiquitous in chemical environments. Diffuse functions that have a small exponent are also often used. These add basis functions for the valence shell with an increased principle quantum number. Similarly to polarized functions a diffuse function is applied to functions with atoms larger than helium and doubly diffuse adds 2s functions to hydrogen and helium.

2.4.3 Dunning Basis Sets

Another popular type of basis sets are collectively known as the correlation consistent Dunning basis sets[48]. These contain functions that are optimised for configuration interaction singles and doubles wavefunctions (CISD, section 2.6.5) which is an accurate wavefunction based method for directly calculating properties of a system. Correlation consistent basis sets follows the nomenclature of cc-pVXZ where cc-pV denotes *correlation consistent-polarised valence,* meaning the basis set is correlation

consistent and contains valence shells (as with split-valence basis sets). XZ denotes whether the basis is a double-ζ, triple-ζ or higher order. These higher order basis sets are often too computationally demanding for use on larger system, limiting most calculations to use a triple-ζ basis set as its highest level of theory.

Functions for a cc-pVXZ basis set are added in shells. Carbon (a relatively small yet abundant atom) with a triple-ζ (cc-pVTZ) basis set contains 4s orbitals, 3p orbitals, 2d orbitals and 1f orbital. It is clear that the amount of orbitals to be evaluated is already large and will only increase with larger atoms shown in table 2.1. For these reasons these type of calculations are usually reserved for situations when high level of accuracy is required.

Atoms	$cc-pVDZ$	$cc-pVTZ$	$cc-pVQZ$
H-He	2s1p	3s2p1d	4s3p2d1f
Li-Ne	3s2p1d	4s3p2d1f	5s4p3d2f1g

Table 2.1: A comparison of the orbitals present for the first and second row atoms using popular correlation consistent basis sets.

These basis sets can be further expanded with diffuse functions which are known as *augmented* correlation consistent basis sets and take the nomenclature of aug-cc-pVXZ. These augmented basis sets add one set of diffuse functions for every shell with angular momentum present in the basis set. Using aug-cc-pVTZ with carbon there would be diffuse s, p, d and f functions.

One significant advantage of using these correlation consistent basis sets is that they scale to the complete basis set limit (CBS). This is effectively the energy of the system using an infinite basis set, and can be quantified for HF calculations with the Dunning basis set[49]:

$$E_{HF,\infty} = \frac{x^5 E_{HF,x} - y^5 E_{HF,y}}{x^5 - y^5} \tag{2.50}$$

where $E_{HF,\infty}$ is the CBS energy and x and y are the highest angular momentum of

two sequential basis sets (2 for $cc-pVDZ$, 3 for $cc-pVTZ$ and so on).

2.5 Density Functional Theory

Density functional theory (DFT) is currently the most popular method for quantum chemical calculations. Many post-HF methods quickly become intractable when studying larger systems at a higher level of theory. DFT provides a significant improvement on the accuracy of HF calculations without a large increase in computational cost. DFT is not a wavefunction based method like HF theory, but is based on what its name suggests, the one electron density. The electron density of a N-electron system is given by equation 2.51, where $\psi_i(\mathbf{r})$ describes a set of molecular orbitals and $\rho(\mathbf{r})$ is the electron density probability of finding an electron within the volume $d\mathbf{r}$.

$$\rho(r) = \sum_{i}^{N} |\psi_i(\mathbf{r})|^2 \tag{2.51}$$

The integral of the electron density gives the number of electrons in the system N:

$$\int \rho(\mathbf{r}) d\mathbf{r} = N \tag{2.52}$$

One distinct advantage of using electron density compared with wavefunction based methods is that the electron density is an observable property using techniques such as X-ray diffraction.

Perhaps the most important aspect of DFT is the way the cost of the calculation scales with the size of the system. As electron density is a function of position in three-dimensional space it only has three variables, X, Y and Z. This means the amount of variables in a density based formalism does not change with the size of the system. This differs from wavefunction based methods which rely on three spatial

coordinates and one spin coordinate per electron resulting in $4N$ variables in an N electron system.

2.5.1 Hohenberg-Kohn Theorems

In 1964 it was proved by Hohenberg and Kohn[50] that all properties for molecules with a nondegenerate ground state could be determined by the ground state electron probability density, ρ_0,

$$\rho_0(\mathbf{r}) \rightarrow E_0 \tag{2.53}$$

The Hohenberg-Kohn theorem states that any ground state property is a functional of the ground state density.

The theorem proved by Hohenberg and Kohn showed that the external potential and number of electrons in a system can be determined from a knowledge of ρ_0. Every electron density is unique and has its own unique external potential. Therefore the ground state electron density can uniquely determine the electronic Hamiltonian.

The expression for the energy can be written in the following form

$$E_0 = \int \rho_0(\mathbf{r}) V(\mathbf{r}) d\mathbf{r} + \bar{F}[\rho_0] \tag{2.54}$$

A way of calculating the unknown term $\bar{F}[\rho_0]$ was formalised in 1965 by Kohn and Sham[51] and is the form most DFT calculations use today.

2.5.2 Kohn-Sham DFT

Kohn-Sham DFT (KS-DFT) involves using a fictitious reference system of N-electrons, denoted with an s, which is analogous to the system which is being studied except with non-interacting electrons, with each electron experiencing the same external potential $V_s(\mathbf{r})$.[51] The aim of KS-DFT is to find an external potential such that

the electron density of the reference system ρ_s is equivalent to the electron density of the real system ρ_0. The Hamiltonian of this reference system eliminates the electron-electron repulsion terms as there is no electron interaction. This allows the simplification of equation 2.26 to the expression in equation 2.55 where $\hat{h}_i^{ks}$ is the one electron KS Hamiltonian.

$$\hat{H}_s = -\frac{1}{2}\sum_i^N \nabla_i^2 + \sum_i^N V_s(\mathbf{r}) = \sum_i^N \hat{h}_i^{ks} \tag{2.55}$$

The reference system can then be related to the real system by rewriting the electronic Hamiltonian as is seen in equation 2.56. In this form, λ is a parameter ranging between zero and one which at zero completely describes the reference system and at one describes the true system. $V_\lambda(\mathbf{r})$ is the external potential needed to make the ground state electron density of the reference system equal to that of the true molecular ground state.

$$\hat{H}_\lambda = \hat{T} + \sum_i^N V_\lambda(\mathbf{r}) + \lambda\hat{V}_{ee} \tag{2.56}$$

The ground state wavefunction is the antisymmetric product of the lowest energy Kohn-Sham spin orbitals of the reference system as it only contains non-interacting orbitals. Finding these orbitals will in turn find the electron density of the true system. The unknown part of the energy functional, $\bar{F}[\rho_0]$, can be written in terms of the difference between the reference and real system, where subscript s denotes a reference term.

$$\Delta\bar{T}[\rho_0] \equiv \bar{T}[\rho_0] - \bar{T}_s[\rho_0] \tag{2.57}$$

The difference in electron-electron repulsion $\Delta\bar{V}_{ee}[\rho_0]$ can be defined as the deviation of the true energy from the Coulomb repulsion $\bar{J}[\rho_0]$

$$\Delta\bar{V}_{ee}[\rho_0] \equiv \bar{V}_{ee}[\rho_0] - \bar{J}[\rho_0] \tag{2.58}$$

where

$$\bar{J}[\rho_0] = \frac{1}{2}\int\int \frac{\rho_0(\mathbf{r}_1)\rho_0(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.59}$$

All the reference and Coulomb terms in the above equations can be calculated exactly. Equations 2.57 and 2.58 can now be written in terms of real energy functions

$$\bar{T}[\rho_0] \equiv \Delta\bar{T}[\rho_0] + \bar{T}_s[\rho_0] \tag{2.60}$$

$$\bar{V}_{ee}[\rho_0] \equiv \Delta\bar{V}_{ee}[\rho_0] + \bar{J}[\rho_0] \tag{2.61}$$

Then substituting these expressions for the kinetic energy and electron-electron repulsion functionals into the overall energy equations, it is possible to construct an equation for the ground state energy of a molecule as a functional of the electron density of that system.

$$E_0 = E[\rho_0] = \bar{T}_s[\rho_0] + \bar{V}_{ne}[\rho_0] + \bar{J}[\rho_0] + \Delta\bar{T}[\rho_0] + \Delta\bar{V}_{ee}[\rho_0] \tag{2.62}$$

The expression given in equation 2.62 now leaves the only unknown terms as $\Delta\bar{T}[\rho_0]$ and $\Delta\bar{V}_{ee}[\rho_0]$. These are grouped together into what is known as the exchange-correlation functional E_{xc}, which groups together all unknown terms in the energy expression.

$$E_{xc}[\rho_0] = \Delta\bar{T}[\rho_0] + \Delta\bar{V}_{ee}[\rho_0] \tag{2.63}$$

This now allows equation 2.62 to be rewritten in the following form to give the full expression for the energy of a system.

$$E_0 = \int \rho_0(\mathbf{r})V(\mathbf{r})d\mathbf{r} + \bar{T}_s[\rho_0] + \frac{1}{2}\int\int \frac{\rho_0(\mathbf{r}_1)\rho_0(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 + E_{xc}[\rho_0] \tag{2.64}$$

The KS equations are used to find the energy of the system and are derived from the second Hohenberg-Kohn theorem which state that DFT is variational. It is shown that the electron density of the reference system ρ_s is equal to that of the real system ρ_0 and that the electron density can be expressed in terms of KS orbitals.

$$\rho_0 = \rho_s = \sum_i^N |\psi_i^{KS}(1)|^2 \tag{2.65}$$

Where ψ_i^{KS} are the KS spatial orbitals and N is the number of occupied KS orbitals. As the orbitals can be doubly occupied with different spin electrons, N is the number of electrons in the system for closed-shell systems. KS-DFT works by attempting to minimise the energy of the individual Kohn-Sham orbitals using a one electron Schrödinger-like equation to deduce the energy

$$\hat{h}_i^{KS}(1)\psi_i^{KS}(1) = \varepsilon_i^{KS}(1)\psi_i^{KS}(1) \tag{2.66}$$

Where $\hat{h}_i^{KS}$ is the one-electron Hamiltonian, ψ_i^{KS} are the KS orbitals and ε_i^{KS} is the energy level of the electron using KS-DFT. The one electron Hamiltonian is given by substituting the density expression in equation 2.65 into equation 2.64 and then varying E_0 with respect to the KS spatial orbitals

$$\hat{h}_i^{KS} = -\frac{1}{2}\sum_i^N \bigtriangledown_i^2 - \sum_A \frac{Z_A}{r_{iA}} + \int \frac{\rho_0(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2 + v_{xc}(1) \tag{2.67}$$

The term $v_{xc}(N)$ is the exchange correlation potential (for an electron number N) which is equivalent from a derivative of the unknown exchange correlation functional $E_{xc}[\rho]$ and is defined as

$$v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} \tag{2.68}$$

In comparison to HF theory, the KS energy in equation 2.64 is exact as long as the exchange correlation functional $E_{xc}[\rho]$ is also exact. The exchange correlation

functional is a collection of unknown terms and cannot be calculated exactly. This means that the DFT energy will always be slightly different to the true energy of a system. How much that energy deviates depends on both the size of the basis set and the choice of the exchange correlation functional.

2.5.3 Exchange and Correlation Functionals

There are several different approaches to approximating the exchange correlation functional. The simplest model is known as the local density approximation (LDA) and was suggested first by Kohn and Sham[51]. LDA functionals assume that if the electron density, ρ, varies very slowly with position then the $E_{xc}[\rho]$ will relate to an integral of electron density and exchange-correlation energy E_{xc}^{UEG}, at a uniform electron density, ρ.

$$E_{xc}^{UEG} = \int e_{xc}^{UEG}(\rho) d\mathbf{r} \tag{2.69}$$

This is based on the uniform electron gas model and is quite a simplistic view of electron behaviour since it assumes the electron density has zero gradient around a point. A variety of the LDA applies different spatial orbitals ψ_{α}^{KS} and ψ_{β}^{KS} to electrons with different spins α and β within the uniform gas model. This method is known as the local spin density approximation (LSDA) and is analogous to using unrestricted HF theory as opposed to spin restricted HF theory hence can be used with systems that contain unpaired electrons.

As a real molecular system does not have a uniform electron distribution, LDA and LSDA functionals offer poor approximations for the exchange correlation functional for these systems. A way to improve the model is to use a generalised-gradient approximation (GGA) functional where both electron density and the gradient of the electron density at point $\mathbf{r}$ are used to form the exchange correlation functional. Many GGA functionals exist and an example of two popular functionals are the PBE functional (Perdew-Burke-Ernzenhof,[52]) and the BLYP functional (Becke-Lee-

Yang-Parr,[53]). The latter separates the exchange and correlation into two parts (equation 2.70).

$$E_{xc} = E_x + E_c \tag{2.70}$$

With regards to the BLYP functional, Becke's exchange functional (B) gives the exchange energy (E_x) whilst LYP accounts for the correlation energy (E_c).[54,55]

A popular form of the exchange-correlation functional used throughout DFT is known as a hybrid functional. These functionals include a fraction of exact HF exchange energy, which can be taken from the equation for total HF energy (equation 2.30) and expressed as a sum of the exchange integrals (equation 2.29).

$$E_x = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} K_{ij} \tag{2.71}$$

Substituting the HF orbitals in equation 2.29 with the KS orbitals gives the exact HF exchange energy for a system of non-interacting electrons

$$E_x^{HF} = -\sum_{i=1}^{N}\sum_{j=1}^{N}\left\langle \psi_i^{KS}(1)\psi_j^{KS}(2)\left|\frac{1}{r_{12}}\right|\psi_i^{KS}(2)\psi_j^{KS}(1)\right\rangle \tag{2.72}$$

One of the most popular hybrid exchange functional in use at the moment is the B3LYP functional[54,55]. This functional has found a wide variety of applications in many fields and has proved to give accurate results for many systems. B3LYP includes components of exact HF exchange, LDA exchange and GGA exchange along with correlation terms. All these terms have been fitted according to empirical parameters as follows:

$$E_{xc}^{B3LYP} = a_0 E_x^{HF} + (1 - a_0 - a_x)E_x^{LDA} + a_x E_x^{B88} + (1 - a_c)E_x^{VWN} + a_c E_x^{LYP} \tag{2.73}$$

where the parameters $a_0 = 0.20$, $a_x = 0.72$ and $a_c = 0.81$ which when substituted

into equation 2.73 forms the following expression

$$E_{xc}^{B3LYP} = 0.20E_x^{HF} + 0.08E_x^{LDA} + 0.72E_x^{B88} + 0.19E_c^{VWN} + 0.81E_c^{LYP} \quad (2.74)$$

Within equation 2.74, E_x^{HF} is the term that describes the HF exact exchange, E_x^{LDA} describes the LDA contribution and E_x^{GGA} is the Becke component of the exchange functional. For the correlation terms, E_{xc}^{VWN} is the correlation functional according to Vosko, Wilk and Nusair[56] and E_c^{LYP} is the Lee-Yang-Parr correlation functional as seen earlier for the BLYP exchange correlation functional.

2.5.4 Long-range corrected functionals

The forms of functionals previously described work well for most cases. However, there are certain applications where they do not perform well. Regular functionals are known to fail to accurately describe: polarisability in long chain molecules, calculation of Rydberg states using time dependent density functional theory (TDDFT) and predicting charge transfer excitations where excitations are separated by a significant distance (several Å or more). This last failure of regular functionals is explored further in Chapter 3 where charge transfer (CT) in proteins is studied with TDDFT.

The failure of standard hybrid functionals to calculate these energies accurately is due to CT states requiring high percent of HF exchange whereas popular hybrid functionals such as B3LYP contain only 20% HF exchange. In TDDFT the excitation energy for a CT transition is essentially the energy difference between the orbital energies which due to the approximate exchange in the functional is often underestimated. Consequently B3LYP underestimates the charge transfer energy. This underestimation in energy has been seen for systems with separations as low as 2Å.[57]

There are two ways to model long-range HF exchange within an exchange correlation functional. This first method involves preserving the chosen functional at a short-range, while incorporating a larger percent of HF exchange at the long range and is a combination of the two in the medium range. Functionals that take advantage of this strategy are known as 'Coulomb-attenuated' or 'long-range corrected' functionals with CAM-B3LYP (Coulomb-attenuated method B3LYP) and ω-B97 being examples of widely used long range corrected functionals.[58,59] Equation 2.75 shows how the electron-electron Coulomb operator is partitioned in the evaluation of the exchange energy.

$$\frac{1}{r_{12}} = \frac{1 - \mathrm{erf}(\omega r_{12})}{r_{12}} + \frac{\mathrm{erf}(\omega r_{12})}{r_{12}} \tag{2.75}$$

Where erf is the error function, the left hand term is a short-range term that decays to zero with increasing interelectronic radius r_{12}. The left hand term deals with the long range interaction and stay relatively constant beyond a point. ω is a parameter that controls the extent of the two terms at a particular separation. For CAM-B3LYP, the above term is expanded to include parameters α and β:

$$\frac{1}{r_{12}} = \frac{1 - [(\alpha + \beta)\,\mathrm{erf}(\omega r_{12})]}{r_{12}} + \frac{(\alpha + \beta)\,\mathrm{erf}(\omega r_{12})}{r_{12}} \tag{2.76}$$

Where $\alpha = 0.19$, $\beta = 0.46$ and $\omega = 0.33/r_{Bohr}$. This has shown to give good improvements for energies of Rydberg states and CT states over standard B3LYP using TDDFT.

An alternative method for calculating long-range exchange accurately is to develop new functionals which are parametrised to deal with exact HF exchange at long range. One such functional that is parametrised specifically for CT is the M06-HF functional[60].

2.5.5 Dispersion Corrections within DFT

Another area that DFT with standard exchange-correlation functionals fails to model accurately is dispersion forces. This is due to the fact that many exchange-correlation functionals are functionals of the local electron density (LDA) or the gradient of the local electron density (GGA), which results in only local effects influencing the behaviour of the electrons in a system. Dispersion forces arise when an induced dipole on one molecule or atom induces a dipole on another molecule in the system. This creates instantaneous dipoles between the two species, which results in a weak attraction force. This is an inherently non-local effect. As the interactions in standard DFT use local electron interaction these dipoles do not form and therefore dispersion forces do not occur. There are several different approaches to overcoming this problem within DFT. One approach to modelling dispersion is the use of functionals that do not contain an explicit dispersion term parameterised using systems where dispersion interactions are prominent. Examples of this approach include the Minnesota suite of functionals[61,62], which have proven successful for modelling some dispersion bound complexes. This thesis concentrates on the more common methods of correcting for dispersion, which involve an add-on term to the standard DFT energy

$$E_{DFT-D} = E_{KS-DFT} + E_{disp} \tag{2.77}$$

where E_{DFT-D} is the dispersion corrected energy, E_{KS-DFT} is the standard Kohn-Sham DFT energy and E_{disp} is the dispersion correction.

Among the most popular dispersion corrections to DFT (DFT-D methods) are Grimme's D2 and D3 empirical corrections[63–65]. These methods are similar, and include E_{disp} as a sum of pairwise interactions of atoms in a system. For the D2 correction (DFT-D2) E_{disp} takes the form

$$E_{disp}^{D2} = -s_6 \sum_{A}^{N} \sum_{B<A}^{N} \frac{C_6^{AB}}{R_{AB}^6} f_{dmp}(R_{AB}) \tag{2.78}$$

where A and B are atoms in an N-atom system, R_{AB} is the distance between atoms A and B, s_6 is a global scaling parameter, $f_{dmp}(R_{AB}$ is the damping function which prevents the dispersion energy becoming infinitely low at small R_{AB} separations. C_6^{AB} is a term which combines atomic parameters C_6^A and C_6^B and takes the form

$$C_6^{AB} = \sqrt{C_6^A C_6^B} \tag{2.79}$$

The damping function is of the form

$$f_{dmp}(R_{AB}) = \frac{1}{1 + exp\left[-d\left(\frac{R_{AB}}{R_{AB}^0} - 1\right)\right]} \tag{2.80}$$

and includes two further parameters, d is the global scaling parameter for the damping function and R_{AB}^0 is the sum of the van der Waals radii of atoms A and B. The values R_{AB}^0 and C_6^{AB} have been derived from experiment by Grimme and global scaling parameters assessed for a range of functionals to empirically determine optimum results using those functionals. For PBE, s_6 is 0.75 whereas for B3LYP s_6 is 1.05.

Grimme also parameterised the slightly improved DFT-D3 method, where higher orders of C^{AB} are included in addition to a term evaluating three-body interactions. However most D3 calculations do not include these three body parts and the improvements comes by considering the chemical environment. The expression for the dispersion interaction of two body term is

$$E_{disp}^{D3} = -\sum_{A}^{N} \sum_{B<A}^{N} \sum_{n=6,8,10...} s_n \frac{C_6^{AB}}{R_{AB}^6} f_{dmp}(R_{AB}) \tag{2.81}$$

which follows a similar format and parameters to the DFT-D2 method.

As well as the empirical D2 and D3 methods, it is possible to introduce an add-on term to DFT which is calculated directly as opposed to relying on empirical terms. One such method is the exchange-dipole model (XDM)[66,67]. This model calculates dispersion directly from the interaction between the instant dipole moment of the exchange hole in one atom of molecule in system and the induced dipole moment in the other atom or molecule. The exchange hole refers to the space around an electron where there is almost zero probability of finding another electron. The XDM dispersion correction has the form

$$E_{disp}^{XDM} = -\sum_{A>B} \frac{C_6^{AB}}{\frac{R_{AB}^6 + kC_6^{AB}}{E_A^C + E_B^C}} \tag{2.82}$$

where k is a universal scaling parameter E_A^C and E_B^C are the absolute values of the correlation energies of atoms A and B, and C_6^{AB} takes the form

$$C_6^{AB} = \frac{\langle d_X^2 \rangle_A \langle d_X^2 \rangle_B \alpha_A \alpha_B}{\langle d_X^2 \rangle_A \alpha_B + \langle d_X^2 \rangle_B \alpha_A} \tag{2.83}$$

where $\langle d_X^2 \rangle_A$ is the exchange hole dipole moment of atom A, and α_A is the effective polarisability of atom A.

2.6 Excited State Methods

The methods discussed so far are appropriate for describing the electronic ground state of a system. The focus of this work is on the excited state properties of porphyrins since much of the interesting chemistry of these systems involves states where the electrons are not in their lowest energy configuration. The ability to accurately model excited state systems is critical to many areas of computational chemistry and several methods for describing excited states are described in the following sections.

2.6.1 Time Dependent Density Functional Theory

TDDFT is an extension of DFT that calculates electronic excitations and other time-dependent effects[68]. Similarly to DFT, TDDFT only relies on the electron density. TDDFT is less computationally expensive than wavefunction based methods for treating excited states, allowing larger systems to be studied more readily.

Extending the Hohenberg-Kohn theorem (section 2.5.1) to a time-dependent framework is described by the Runge-Gross theorem[69]. This states that *for a many-body time-dependent system evolving from an initial wavefunction that all properties can be calculated with the knowledge of the one-body density, and these properties are a functional of the density*, resulting in a time-dependent analogue to the standard HK theorems. The same KS theory for DFT described in section 2.5.2 can be applied to a time-dependent system obeying the time-dependent Schrödinger equation for a single electron with the following Hamiltonian

$$\hat{H}_s(t) = \hat{T} + \hat{V}_s(t) \tag{2.84}$$

where the subscript s refers to a KS non-interacting system, $\hat{T}$ is the kinetic energy and $\hat{V}_s(t)$ the time-dependent potential energy. This can be expanded to use known terms for the Hamiltonian components and placed in context of a time-dependent Schrödinger equation.

$$i\frac{\partial}{\partial t}\psi_i(\mathbf{r};t) = \left[-\frac{\nabla^2}{2} + v_{KS}(\mathbf{r};t)\right]\psi_i(\mathbf{r};t) \tag{2.85}$$

where the density of the time-dependent KS orbitals are

$$\rho_s(\mathbf{r};t) = \sum_{i=1}^{N} |\psi_i(\mathbf{r};t)|^2 \tag{2.86}$$

As before with regular KS DFT, the density of the reference system is the same as the density of the true system

$$\rho_s(\mathbf{r};t) = \rho(\mathbf{r};t) \tag{2.87}$$

The unknown part of these equations is the time-dependent potential, $v_{KS}(r,t)$. This consists of the sum of the external potential of the system v_{ext}, the Coulombic interaction v_J and the exchange correlation potential v_{xc}, with all these potentials being dependent on both position and time.

$$v_s(\mathbf{r};t) = v_{ext}(\mathbf{r};t) + v_J(\mathbf{r};t) + v_{xc}(\mathbf{r};t) \tag{2.88}$$

The most popular formalism of TDDFT involves linear response to the external potential formalised by Casida[70,71]. This calculates the poles in the response to time-dependent varying applied external field which correspond to the excitation energies.

2.6.2 Tamm-Dancoff Approximation

TDDFT can be used within the Tamm-Dancoff approximation (TDA)[72], where the excitation energies of a system are determined in the following eigenvalue equation

$$\mathbf{AX} = \omega\mathbf{X} \tag{2.89}$$

where $\mathbf{X}$ corresponds to the oscillator strengths, ω are the excitation energies and $\mathbf{A}$ is given by the following expression

$$A_{ia\sigma,jb\tau} = \delta_{ij}\delta_{ab}\delta_{\sigma\tau}(\varepsilon_{a\sigma} - \varepsilon_{i\tau}) + (ia\sigma|jb\tau) + (ia\sigma|f_{xc}|jb\tau) \tag{2.90}$$

where ε_i are orbital energies and

$$(ia\sigma|jb\tau) = \int\int \frac{\Psi^*_{i\sigma}(\mathbf{r}_1)\Psi_{j\tau}(\mathbf{r}_2)\Psi_{b\tau}(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.91}$$

$$(ia\sigma|f_{xc}|jb\tau) = \int \Psi^*_{i\sigma}(\mathbf{r}_1)\Psi_{a\sigma}(\mathbf{r}_1)\frac{\partial^2 E_{xc}}{\partial\rho_\sigma(\mathbf{r}_1)\partial\rho_\tau(\mathbf{r}_2)}\Psi_{j\tau}(\mathbf{r}_2)\Psi^*_{b\tau}(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 \qquad (2.92)$$

where E_{XC} is the exchange correlation functional. This allows excitation energies to be calculated for a system.

2.6.3 Maximum Overlap Method

One approach to calculating excited states is to explicitly model the state of interest within an SCF calculation. The Maximum Overlap Method[73] is a useful alternative to TDDFT. By forcing the SCF calculation to converge to higher energy roots it is possible to study excited states using KS-DFT. This is advantageous as the orbitals are optimised directly for the excited state of interest. This allows excitation energies to be determined in a ΔSCF approach where the excitation energy is the difference between the ground and excited state SCF energies.

The MOM employs a simple algorithm for finding the excited state solutions to the KS and HF SCF equations. The MOM uses the shape of the molecular orbitals in order to determine the next orbital selection in the SCF procedure. Instead of describing electronic arrangement according to the Aufbau principle, the MOM algorithm uses an approach where it maximises the overlap between the occupied orbitals preventing the excited state from collapsing on successive SCF iterations. This prevents the variational collapse to the ground state and allows the SCF to find the nearest energy solution rather than the lowest.

At each SCF iteration the MOM creates a new set of occupied orbitals which are based on those that overlap most with the old occupied orbitals. The overlap matrix, O, of the orbitals is defined as

$$\mathbf{O} = (\mathbf{C}^{old})\mathbf{S}\mathbf{C}^{new} \tag{2.93}$$

where $\mathbf{C}^{old}$ is the old coefficient matrix in the SCF process, $\mathbf{C}^{new}$ is the new coefficient matrix and $\mathbf{S}$ is the basis function overlap matrix. The general term O_{ij} is the overlap between the i-th old orbital and the j-th new orbital in the system. The projection p_j of the j-th orbital onto the old occupied space is expressed as:

$$p_i = \sum_{j}^{N} O_{ij} \tag{2.94}$$

where

$$O_{ij} = \sum_{\alpha=1}^{M}\sum_{\beta=1}^{M} C_{\alpha i}^{old} S_{\alpha\beta} C_{\beta j}^{new} \tag{2.95}$$

N is the number of electrons, and M is the number of basis functions. Then N orbitals with the largest projections p_j should be occupied at the next SCF iteration.

The MOM requires an initial 'guess' of orbitals and a starting electronic configuration of the occupied and virtual space. They are typically generated by modifying the occupancies of the molecular orbitals obtained from a calculation performed on the ground state. In order to optimise the excited state it is necessary to identify the orbitals involved in the transition prior to the MOM calculation. If the excitation is a result of a $\pi \rightarrow \pi^*$ transition (commonly associated with porphyrin excitations) then these orbitals must be identified with a pre-MOM single point calculation. The success of the MOM relies on the quality of the initial guess for the calculation. If the virtual orbitals are of poor quality then the calculation is more likely to collapse to a lower energy state of the same symmetry.

The MOM has several advantages over existing low-cost excited state methods. One of these its ability to target very high energy states and ignore low lying states which are not of interest. It can be seen as an inexpensive method in order to

optimise excited states particularly of large systems which is the overall central focus of this work. The computational cost of the MOM is the same as that of the corresponding ground-state method per excited state. Another advantage of this approach is that the memory requirements remain constant and do not scale with the number of states. However, being a single-determinant method, MOM is expected to give poor results for excited states with genuine multi reference characteristics.

Upon a successful MOM calculation the obtained results are accurate for a wide range of systems although MOM is known to fail in certain cases. Excitation energies for open shell singlet states involving excitation to valence orbitals are often underestimated. This is due to the open-shell singlet state being described by a single determinant to give a spin-mixed state. Excitation energies can be significantly improved by applying a post-SCF spin-purification correction suggested by Ziegler[74].

$$E = 2E_S - E_T \tag{2.96}$$

where E is the true (spin-purified) singlet energy, E_S the spin-mixed (single determinant) singlet energy and E_T the energy of the corresponding triplet state. This is rationalised by knowing that the separation between the triplet and spin-mixed singlet states is roughly the exchange energy K between the two unpaired electrons, as well as the spin-mixed singlet and spin-pure singlet energy separation being K. This gives the expression

$$E = E_T + 2K \tag{2.97}$$

which can be shown to be $2E_S$ minus E_T when E_S and E_T are rearranged in terms of E_T and K. In this work calculations which have been spin corrected are referred to as spin-mixed MOM (SM-MOM), whereas calculations using the spin-purification post-SCF method are known as spin-pure (SP-MOM).

2.6.4 Single Excitation Configuration Interaction

In Single Excitation Configuration Interaction (CIS) the excited state wavefunction is expressed as a linear combination of all possible singly excited determinants,

$$\Psi_{CIS} = \sum_{ia} a_i^a \Psi_i^a \tag{2.98}$$

where the coefficients a_{ia} can be obtained by diagonalising the many-electron Haamiltonian, $\mathbf{A}$, in the space of all single substitutions. The appropiate matrix elements are

$$A_{ia,jb} = \left\langle \Psi_i^a | H | \Psi_j^b \right\rangle = \left[E_{HF} + \varepsilon_a - \varepsilon_j \right] \delta_{ij} \delta_{ab} - (ja||ib) \tag{2.99}$$

According to Brilloun's theorem single substitutions do not interact directly with a reference HF determinant, so the resulting eigenvectors from the CIS excited state represent a treatment roughly comparable to that of the HF ground state. The excitation energy is simply the difference between the HF ground state energy and the CIS excited state energies, and the eigenvectors of $\mathbf{A}$ correspond to the amplitudes of the single-electron promotions.

2.6.5 Configuration Interaction

Configuration interaction (CI) improves the HF wavefunction by adding additional determinants that describe the promotion of electrons from occupied to virtual (unoccupied) orbitals. The total wavefunction can be described as

$$\Psi = c_0 D_0 + c_1 D_1 + c_2 D_2 + ... + c_i D_i \tag{2.100}$$

where D_0 is the HF determinant, D_1 through to D_i are the determinants of the system with electrons promoted to virtual orbitals, and c_1 through to c_i are constants. This can lead to the expression

$$\Psi_{CI} = \Psi_{HF} + \sum_{I} a_I \Psi_I \tag{2.101}$$

where Ψ_{CI} is the CI wavefunction, Ψ_{HF} is the HF wavefunction and $\sum_I a_I \Psi_I$ corresponds to the sum of all excitation wavefunctions. CI calculations need a reasonably-sized basis set to give reliable results. Hence full CI is prohibitively expensive and usually restricted to very small systems due to the large number of possible excitation wavefunctions in a many body electron system.

2.6.6 Complete Active Space

Another method which is computationally less demanding than full CI is known as complete active space self consistent field (CASSCF).[75] This method splits the molecular orbitals into three categories: core orbitals, active orbitals and virtual orbitals. CASSCF seeks to calculate all combinations of electronic excitations for only the active space. This is analogous to the full CI; however, the active orbital space is significantly reduced from the full molecular system. In the CAS framework, core orbitals are always occupied and virtual orbitals always empty, leaving the active space to deal with the highest occupied and lowest unoccupied orbitals in the system.

An extension to the regular CASSCF procedure is CASPT2[76], which introduces dynamic correlation effects into the regular CASSCF calculation, by addition of a second order perturbation similar to Møller-Plesset perturbation theorem for ground states.

2.7 Polarizable Continuum Models

Many electronic properties will be different in solution than in gas phase. There are several possible methods to address this problem. One approach describes the solvent environment by including them explicitly in the calculation which inherently increases the cost significantly. Clearly this approach is inappropriate for most

cases and so a polarizable continuum model (PCM) is a commonly used method to model solvation effects. PCM provides realistic description of molecular shape by constructing the solute cavity from a union of atom-centered spheres. In addition, the exact electron density of the solute (rather than a multipole expansion) is used to polarize the continuum. Electrostatic interactions between the solute and the continuum manifest as an induced charge density on the cavity surface, which is discretized into point charges for practical calculations. The surface charges are determined based upon the solute's electrostatic potential at the cavity surface, hence the surface charges and the solute wavefunction must be determined self-consistently. The use of PCM is investigated in Chapter 4.

2.8 *Ab Initio* Molecular Dynamics

Ab initio molecular dynamics (AIMD) is a technique that combines part quantum-mechanical *ab-initio* calculations within classical molecular dynamic (MD) simulations. AIMD uses a quantum mechanical approach in order to calculate energies and forces from first principles, as opposed to from a classical empirical forcefield of regular classical MD. In this work, TDDFT is used to calculate the excitation energies from the AIMD simulations. The specific AIMD method used is known as Born-Oppenheimer molecular dynamics (BOMD)[77]. This technique uses classical molecular dynamics trajectories from the Born-Oppenheimer potential energy surface generated by a particular theoretical chemical model, such as using parameters associated with DFT. This method calculates the forces on the nuclei for the MD simulation 'on-the-fly'. A BOMD calculation consists of an energy calculation along with a gradient calculation at each molecular dynamics time step. Snapshots of the trajectories can be extracted as single point coordinates and treated with excited state methods such as the MOM or TDDFT described earlier in order to give the excited state properties of a system.

AIMD calculations can be quite expensive. Hence quantum chemistry packages

take advantage of specific algorithms in order to accelerate calculations. The speed at which BOMD at the HF and KS level is calculated can be greatly increased by taking previous information from the previous time step to construct a good initial guess for the new molecular orbitals of Fock matrix, thus hastening SCF convergence. In QChem, a Fock matrix extrapolation procedure[78], based on a suggestion by Pulay and Fogarasi[79], is available for this purpose. The Fock matrix elements $\mathbf{F}_{\mu\nu}$ in the atomic orbital basis are oscillatory functions of the time t, and Q-Chem's extrapolation procedure fits these oscillations to a power series in t:

$$\mathbf{F}_{\mu\nu}(t) = \sum_{n=0}^{N} c_n \, t^n \tag{2.102}$$

The $N + 1$ extrapolation coefficients c_n are determined by a fit to a set of M Fock matrices retained from previous time steps. Fock matrix extrapolation can significantly reduce the number of SCF iterations required at each time step, but for low-order extrapolations, or if the SCF convergence is not satisfied, a systematic drift in the total energy may be observed. Benchmark calculations testing the limits of energy conservation[57] demonstrate that numerically exact classical dynamics (without energy drift) can be obtained at significantly reduced cost.

AIMD can be performed on a single molecular system or combined with classical molecular mechanics to simulate its behaviour in a solvent or membrane. These types of calculations are known as quantum mechanics/molecular mechanics (QM/MM) calculations[80–82], and allow accurate *ab initio* methods to be applied to the molecule of choice, while still modelling its surroundings albeit in an approximate way.

2.9 General Theory of Molecular Spectroscopy

The simplest appearance of absorption spectra is found for isolated atoms in the gas phase. These spectra exhibit very narrow ranges of wavelength where light

absorption occurs, which was first shown by Bohr. These lines reflect the pattern of energy quantization in the atoms. A key to his theory was Einstein's suggestion that the energy of light is quantized, whereby the energy E_{ph} of the light quanta (photons) is determined by the frequency, ν of the light

$$E_{ph} = h\nu \tag{2.103}$$

for a wave with frequency, ν, and wavelength, λ the following equation is true

$$\nu\lambda = c \tag{2.104}$$

where c is the speed of light, the photon energy can be also expressed by the wavelength

$$E_{ph} = \frac{hc}{\lambda} \tag{2.105}$$

Light can only be absorbed if the photon energy exactly matches the energy difference between the two possible energy state n and m of the atom

$$\Delta E_{nm} = \frac{hc}{\lambda} \tag{2.106}$$

In the case of an absorption spectrum, it is usually the energy spacing between the ground state, the normal state of the atom, and any of its excited states. Molecular energy level diagrams (like the ones in Chapter One) are a reflection of this.

In molecules the atoms can vibrate. They can do so in $3N$-6 modes. The vibrational energy of each of these modes is quantized in multiples of $h\nu_{vib}$, where ν_{vib} is the vibrational frequency of the mode. Their energy level schemes represent uniformly spaced ladders with step widths much smaller than typical electronic excitation energies, but extending up to typical bond breakage energies which are normally larger than the spacing between the electronic levels.

2.9.1 Franck-Condon Principle

The Franck-Condon Principle concerns the intensities of vibronic transitions, or the absorption or emission of a photon. It states that when a molecule is undergoing an electronic transition, such as ionization, the nuclear configuration of the molecule experiences no significant change. This is due in fact that nuclei are much heavier than electrons and the electronic transition takes place faster than the nuclei can respond. When the nucleus realigns itself with with the new electronic configuration, the theory states that it must undergo a vibration.

If the vertical transition from ground to excited electronic state occurs from a vibrational wave function that gives a probability distribution of finding the nuclei in a given region of space it is possible to determine the probability of a given vibrational level from the overlap integral $S_{v',v}$ which gives the overlap of the vibrational wave function in the ground and excited state. The v' quantum numbers refer to the ground state and the v quantum numbers refer to the excited state. The transition probability can be separated into electronic and nuclear parts using the Condon approximation. This is the assumption that the electronic transition occurs on a time scale short compared to nuclear motion so that the transition probability can be calculated at a fixed nuclear position.

In figure 2.3, the nuclear axis shows a consequence of the internuclear separation and the vibronic transition is indicated by the vertical arrows. This figure demonstrates three things: An absorption leads to a higher energy state, fluorescence leads to a lower energy state, and the shift in nuclear coordinates between the ground and excited state is indicative of a new equilibrium position for nuclear interaction potenial. The fact that the fluorescence arrow is shorter than the absorption indicates that it has less energy, or that its wavelength is longer.

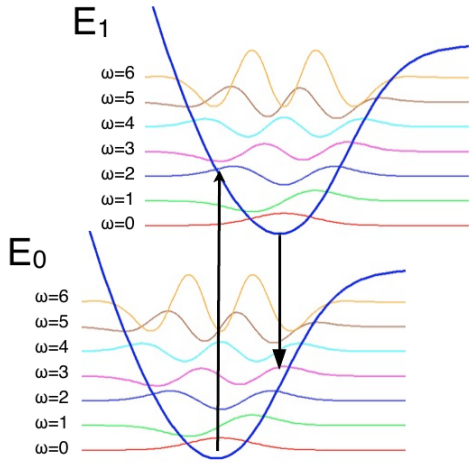


Figure 2.3: Franck–Condon principle energy diagram. Since electronic transitions are very fast compared with nuclear motions, vibrational levels are favored when they correspond to a minimal change in the nuclear coordinates.

This change in vibration is maintained during a state termed the rapid electronic excitation. The resulting Coulombic forces produce an equilibrium as shown in the figure for the nuclei termed a turning point. The turning point can be mapped by drawing a vertical line from the minimum of the lower curve to the intersection of the higher electronic state. This procedure is deemed a vertical transition.

The Franck-Condon Principle explains the relative intensities of vibronic transitions by relating the probablity of a vibrational transition to the overlap of the vibrational wave functions. It states that the probability of a vibrational transition occurring is weighted by the Franck-Condon overlap integral:

$$P_{i\to f} = |\langle \psi^*_{final} | \boldsymbol{\mu} | \psi_{initial} \rangle|^2 = | \int \psi^*_{final} \boldsymbol{\mu} \psi_{initial} d\tau|^2 \tag{2.107}$$

Within the Franck-Condon approximation, the nuclei are considered 'fixed' during electronic transitions. Thus, electronic transitions can be considered vertical transitions on electronic potential energy curves.

The Franck-Condon Principle has both a classical and quantum application. Classically, the Franck–Condon principle is the approximation that an electronic transition is most likely to occur without changes in the positions of the nuclei in the molecular entity and its environment. The resulting state is called a Franck–Condon state, and the transition involved, a vertical transition. The quantum mechanical formulation of this principle is that the intensity of a vibronic transition is proportional to the square of the overlap integral between the vibrational wavefunctions of the two states that are involved in the transition.

The Franck-Condon principle is based on the Born-Oppenheimer approximation, which allows the separation of the electronic $\Psi_{\mathrm{el}}(q)$ and nuclear $\Psi_{\mathrm{nuc}}(Q)$ wave functions, such that the total wave function may be written as,

$$|\psi_{total}(Q,q)\rangle = |\psi_{nuc}(Q)\rangle|\psi_{el}(q;Q)\rangle \tag{2.108}$$

Since the transition operator, $\hat{\mu}(q)$, is dependent only on the electronic component, the nuclear components can be separated from the transition moment integral that dictates the probability of the transition occuring:

$$\langle\psi^*_{total,f}|\hat{\mu}|\psi_{total,i}\rangle = \langle\psi^*_{nuc,f}|\langle\psi^*_{el,f}|\boldsymbol{\mu}|\psi_{el,i}\rangle|\psi_{nuc,i}\rangle \tag{2.109}$$

$$= \underbrace{\langle\psi^*_{nuc,f}|\psi_{nuc,i}\rangle}_{\text{nuclear overlap}}\langle\psi^*_{el,f}|\boldsymbol{\mu}|\psi_{el,i}\rangle \tag{2.110}$$

If the nuclear overlap integral is zero for this transition, then the transition will not be observed, irrespective of the magnitude of the electronic factor.

The nuclear overlap for the zero-zero transition S_{00} can be calculated quite simply using the definition of the Gaussian form of the harmonic oscillator wavefunctions.

The zero-point wavefunction in the ground electronic state is

$$|\psi(R)\rangle = \Big| \left(\frac{\alpha}{\pi}\right)^{1/4} e^{-\alpha(R-R_e)^2/2} \Big\rangle \tag{2.111}$$

The zero-point wavefunction in the excited electronic state is

$$|\psi(R)\rangle = \Big| \left(\frac{\alpha}{\pi}\right)^{1/4} e^{-\alpha(R-Q_e)^2/2} \Big\rangle \tag{2.112}$$

where $\alpha = \dfrac{\sqrt{mk}}{\hbar}$, R_e is the equilibrium bond length in the ground electronic state and Q_e is the equilibrium bond length in the excited electronic state.

The nuclear overlap integral is

$$S_{00} = \langle \psi^*_{nuc,f} | \psi_{nuc,i} \rangle = \sqrt{\frac{\alpha}{\pi}} \int_{-\infty}^{\infty} e^{-\alpha(R-R_e)^2/2} e^{-\alpha(R-Q_e)^2/2} dR \tag{2.113}$$

The exponent in equation 2.113 can be expanded as

$$S_{00} = \sqrt{\frac{\alpha}{\pi}} \int_{-\infty}^{\infty} e^{-\alpha(2R^2 - RR_e - 2RQ_e + R_e^2 + Q_e^2)/2} dR \tag{2.114}$$

and we use

$$(R_e + Q_e)^2 = R_e^2 + Q_e^2 + 2R_eQ_e \tag{2.115}$$

and

$$(R_e - Q_e)^2 = R_e^2 + Q_e^2 - 2R_eQ_e \tag{2.116}$$

to substitute and complete the square inside the integral. We can express

$$R_e^2 + Q_e^2 = \frac{1}{2}[(R_e + Q_e)^2 + (R_e - Q_e)^2]. \tag{2.117}$$

Thus, the integral in equation 2.114 is

$$S_{00} = \sqrt{\frac{\alpha}{\pi}} e^{-\alpha(R_e - Q_e)^2/4} \int_{-\infty}^{\infty} e^{-\alpha\{R - 1/2(R_e + Q_e)\}^2} dR \tag{2.118}$$

The integral is a Gaussian integral. It can be shown that if we let $z = \sqrt{\alpha}\{R - 1/2(R_e + Q_e)\}$ then $dz = \sqrt{\alpha}dR$ and the integral becomes

$$S_{00} = \sqrt{\frac{\alpha}{\pi}} e^{-\alpha(R_e - Q_e)^2/4} \frac{1}{\sqrt{\alpha}} \int_{-\infty}^{\infty} e^{z^2} dz \tag{2.119}$$

this integral has been solved already, from a table of integrals, equation 2.119 becomes

$$S_{00} = e^{-\alpha(R_e - Q_e)^2/4} \tag{2.120}$$

We would follow the same procedure to calculate that overlap of the zeroth level vibration in the ground to the first excited vibrational level of the excited state: S_{01}.

To calculate the overlap of zeroth ground state level ($v = 0$) with the first excited state level ($v' = 1$) we use the Hermite polynomial $H_1(x) = 2x$ for describing the excited state wavefunction). Here $x = \sqrt{\alpha}(R - Q_e)$.

$$S_{01} = \langle \psi^*_{nuc,f} | \psi_{nuc,i} \rangle \tag{2.121}$$

with the zero-point wavefunction in the ground electronic state is

$$|\psi(R)\rangle = \Big| \left(\frac{\alpha}{\pi}\right)^{1/4} e^{-\alpha(R - R_e)^2/2} \Big\rangle \tag{2.122}$$

The first excited-state wavefunction in the excited electronic state is

$$|\psi(R)\rangle = \Big| \left(\frac{\alpha}{\pi}\right)^{1/4} \sqrt{\alpha} 2(R - Q_e) e^{-\alpha(R - Q_e)^2/2} \Big\rangle \tag{2.123}$$

The overlap of zeroth ground state level with the first excited state level (equation 2.121) is then

$$S_{01} = \frac{1}{\sqrt{2}} \sqrt{\frac{\alpha}{\pi}} \int_{-\infty}^{\infty} e^{-\alpha(R - R_e)^2/2} \sqrt{\alpha} 2(R - Q_e) e^{-\alpha(R - Q_e)^2/2} dR \tag{2.124}$$

and

$$S_{01} = \sqrt{\frac{2\alpha^2}{\pi}} e^{-\alpha(R_e - Q_e)^2/4} \int_{-\infty}^{\infty} (R - Q_e) e^{-\alpha\{R - 1/2(R_e + Q_e)^2\}} dR \tag{2.125}$$

The same substitutions can be made as above so that the the final result is

$$S_{01} = \sqrt{\frac{\alpha^2}{2}} (R_e - Q_e) e^{-\alpha(R_e - Q_e)^2/4} dR \tag{2.126}$$

We could continue and calculate that overlap of the zeroth level in the ground state with all the higher light vibrational levels: S_{02}, S_{03}, etc. Each term corresponds to a transition with a different energy since the vibrational levels have different energies. The absorption band then has the appearance of a progression (a Franck-Condon progression) of transitions between different levels each with its own probability.

2.10 Methods for predicting electron transfer

The transport of charge via electrons and the transport of excitation energy via excitons are two processes of fundamental importance in diverse areas of research. Characterization of electron transfer (ET) and excitation energy transfer (EET) rates are essential for a full understanding of, for instance, biological systems (such as respiration and photosynthesis) and optoelectronic devices (which interconvert electric and light energy).

The electronic coupling factor is critical in determining transfer rates. Although ET and EET are different processes, many strategies for calculating the couplings share common themes. The electronic coupling factor is an off-diagonal Hamiltonian matrix element between the initial and final diabatic states in the transport processes. ET coupling is essentially the interaction of the two molecular orbitals where the electron occupancy is changed.[83]

2.10.1 Generalised Mulliken Hush

Solving the Schrödinger equation yields a set of eigenstates, but the coupling is the Hamiltonian element in a charge-localized space. Generalized Mulliken-Hush (GMH) and its variants use additional operators to define this charge localised space and calculate the off-diagonal Hamiltonian matrix element.

The GMH scheme is a generalization from the Mulliken-Hush expression, which uses absorption spectra to determine coupling values for optical ET.[84] In GMH,[85] the two charge-localized states are assumed to be the states that have a zero transition dipole moment between them.

The ET coupling, V_{if}, is then calculated with two charge-localized states $|i\rangle$ and $|f\rangle$, or in the eigenstate space

$$V_{if} = \frac{\mu_{12}\Delta E_{12}}{\sqrt{(\mu_1 - \mu_2)^2 + 4{\mu_{12}}^2}} \tag{2.127}$$

where μ_{12} is the transition dipole moment, ΔE_{12} is the energy difference, and μ_1 and μ_2 are the permanent dipoles, all of which are calculated between the two eigenstates of the system.

2.10.2 Fragment Charge Difference

The fragment charge difference (FCD) scheme is similar to the GMH. In FCD,[83] a charge difference operator is employed. The molecule is partitioned into two fragments, one for the donor and the other for the acceptor. A 2×2 donor -acceptor charge difference matrix, Δq, is defined with its element Δq_{mn}:

$$\Delta q_{mn} = \int_{\mathbf{r}\epsilon D} \rho_{mn}(\mathbf{r}) d\mathbf{r} - \int_{\mathbf{r}\epsilon A} \rho_{mn}(\mathbf{r}) d\mathbf{r} \tag{2.128}$$

where the diagonal elements Δq_{mn} are calculated from the diagonal one particle

density $\rho_{mm}(\mathbf{r})$ and the off diagonal Δq_{mn} is the corresponding quantity from the transition density, $\rho_{mm}(\mathbf{r})$.

$$\rho_{mn}(\mathbf{r}) = N \int \cdots \int d\mathbf{r}_2 \cdots d\mathbf{r}_N \Psi_m(\mathbf{r}, \mathbf{r}_2 \cdots \mathbf{r}_N) \Psi_n^*(\mathbf{r}, \mathbf{r}_2 \cdots \mathbf{r}_N) \tag{2.129}$$

where N is the number of electrons in the system. Therefore, the matrix for charge difference can be calculated in eigenstate space.

$$\Delta q = \begin{pmatrix} \Delta q_1 & \Delta q_{12} \\ \Delta q_{12} & \Delta q_2 \end{pmatrix} \tag{2.130}$$

In general Δq is not diagonal because transition density may exist between the two eigenstates. The charge-localized states are defined as the linear combination of the eigenstates that diagonalizes Δq or, equivalently, maximizes charge separation. The coupling is obtained after a transformation of the Hamiltonian

$$V_{if} = \frac{(E_1 - E_2)|\Delta q_{12}|}{\sqrt{(\Delta q_1 - \Delta q_2)^2 + 4{\Delta q_{12}}^2}} \tag{2.131}$$

Both GMH and FCD are useful and flexible computational schemes for ET couplings. With entry-level excited-state models, they can be used for large molecules for practical purposes. They can also be used with advanced quantum models for a precise account.[86]

The expression in equation 2.132 is the fragment excitation difference (FED) scheme.[85] Similar to FCD, it is convenient to use a population analysis for evaluating such as the Mulliken population analysis. FED is used for calculating EET where x is the excitation difference calculated in the same way as Δq (equation 2.130).

$$V_{if} = \frac{(E_1 - E_2)|\Delta x_{12}|}{\sqrt{(\Delta x_1 - \Delta x_2)^2 + 4{\Delta x_{12}}^2}} \tag{2.132}$$

Chapter 3

Quantum Chemical Calculations of Tryptophan→Heme Electron and Excitation Energy Transfer Rates in Myoglobin

The development of optical multidimensional spectroscopic techniques has opened up new possibilities for the study of biological processes. Recently, ultrafast two-dimensional ultraviolet spectroscopy experiments have determined the rates of tryptophan → heme electron transfer and excitation energy transfer for the two tryptophan residues in myoglobin.[87] Here we show that accurate prediction of these rates can be achieved using Marcus theory in conjunction with TDDFT. Key intermediate residues between the donor and acceptor are identified, and in particular the residues Val68 and Ile75 play a critical role in calculations of the electron coupling matrix elements. Our calculations demonstrate how small changes in structure can have a large effect on the rates, and show that the different rates of electron transfer are dictated by the distance between the heme and tryptophan residues, while for excitation energy transfer the orientation of the tryptophan residues relative to the

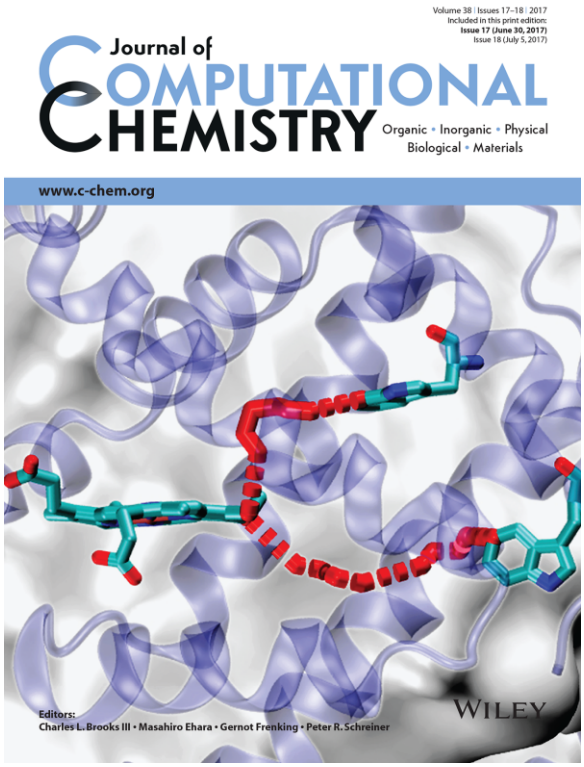


Figure 3.1: The work presented in this chapter is published in the Journal of Computational Chemistry[4] and features on the front cover of the Journal for June 2017.

heme is important.

3.1 Introduction

Electron transfer (ET) and the transfer of excitation energy are fundamental processes in biological systems. The efficient and controlled movement of electrons is one of the primary regulation mechanisms in biology and critical for the existence of living organisms,[88,89] and excitation energy transfer (EET) is important in light harvesting systems.[90] This has motivated the development of experimental and computational approaches to characterise the mechanisms of ET and EET processes. However, these studies are challenging due to the fast time-scale of ET and EET and the complexity of biological systems. Two-dimensional ultraviolet (2D-UV) spectroscopy has emerged as a powerful technique for the study of biological systems, since many biological chromophores, such as tryptophan, absorb in this region of the spectrum.[91] Alongside these experimental techniques complementary

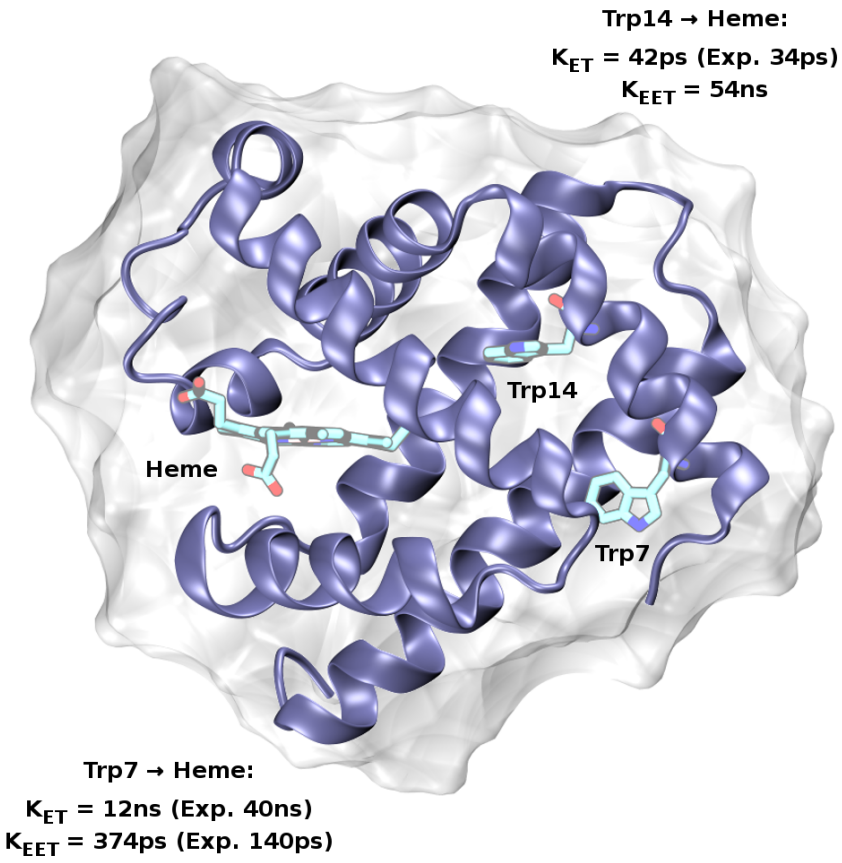


Figure 3.2: The rates of tryptophan → heme electron transfer and excitation energy transfer in myoglobin are simulated based upon Marcus theory in conjunction with quantum chemical calculations and reveal the structural factors affecting the rates of these processes.

theoretical approaches have been developed,[92] which can play an important role in interpreting spectra measured in experiment. Here we focus on tryptophan→heme ET and EET in myoglobin. Myoglobin, an oxygen carrier in muscle tissue, comprises a single polypeptide chain of 153 amino acids arranged in eight α-helices with an iron porphyrin active site and has been described as the hydrogen atom of biology and a paradigm of complexity.[93]

Myoglobin has two tryptophan residues (Trp7 and Trp14), and the fluorescence decay times of these residues have been measured in several myoglobin complexes and are approximately 120 ps for Trp7 and 20 ps for Trp14.[94–97] Recently, Chergui and co-workers studied myoglobin with ultrafast 2D-UV spectroscopy providing in-

sight into the ET and EET processes from the tryptophan residues to the heme.[95] The tryptophan residues are located in an α helix separated from the heme by the E helix, of which several amino acids (Val68, Leu69, Thr70, Gly74, Ile75 and Leu76) lie within the direct path to the heme. The experiments showed that Trp14, which is closer to the heme, decays predominantly by ET, whereas Trp7 relaxes by EET. The rate of ET was quantified, with the relaxation time for Trp14 and determined to be $\leq$40 ps, with a much slower time of >40 ns for Trp7. Furthermore, a time of 140 ps was measured for EET from Trp7. The distance between the tryptophan and heme suggests the EET occurs via the Förster mechanism. Subsequent work on deoxy-myoglobin indicated that a similar ET process is present in ligated ferrous myoglobins.[96] It has been suggested that the ET pathway involves the Leu69 residue which is in van der Waals contact with the Trp14 and Val68 residues,[96] although other work has found a glutamic acid residue to be important.[98]

It is important that computational modelling of ET and EET develops alongside advances in experiments in order to predict rates of ET and EET with sufficient accuracy to elaborate on the detailed underlying mechanisms. In particular, computational modelling of these processes can provide an understanding of the different ET and EET properties of the two tryptophan residues. The theoretical treatment of ET usually follows Marcus theory[99–101] wherein for weak coupling the rate is expressed as

$$k_{ET} = \frac{|V_{DA}|^2}{\hbar}\sqrt{\frac{\pi}{\lambda k_B T}}\exp\left[-\frac{(\lambda + \Delta G)^2}{4\lambda k_B T}\right] \tag{3.1}$$

where ΔG is the change in free energy between the final and initial states, λ is the reorganisation energy and corresponds to the energy for ET in the absence of a change in structure and V_{DA} is the diabatic electron coupling matrix element between the two electronic states. Within this formalism the rate can be estimated once the key quantities, ΔG, λ and V_{DA} have been determined, and this approach has been successfully applied to many systems.[102–104] ΔG and λ are further ex-

plained in figure 3.3. Many heme proteins, such as cytochrome c, are involved in biologically important electron transfer reactions. In contrast to cytochrome c, the heme group in myoglobin is not covalently bound and metal substitution is easier to effect, which has motivated many experiments on myoglobin as an exemplar of electron transfer in heme proteins.[105] The quantitative determination of the rates of ET and EET in myoglobin provides an opportunity to assess computational models for determination of these rates. This can establish the important criteria for calculation of these rates which can then be applied in studies of ET and EET of biologically important processes.

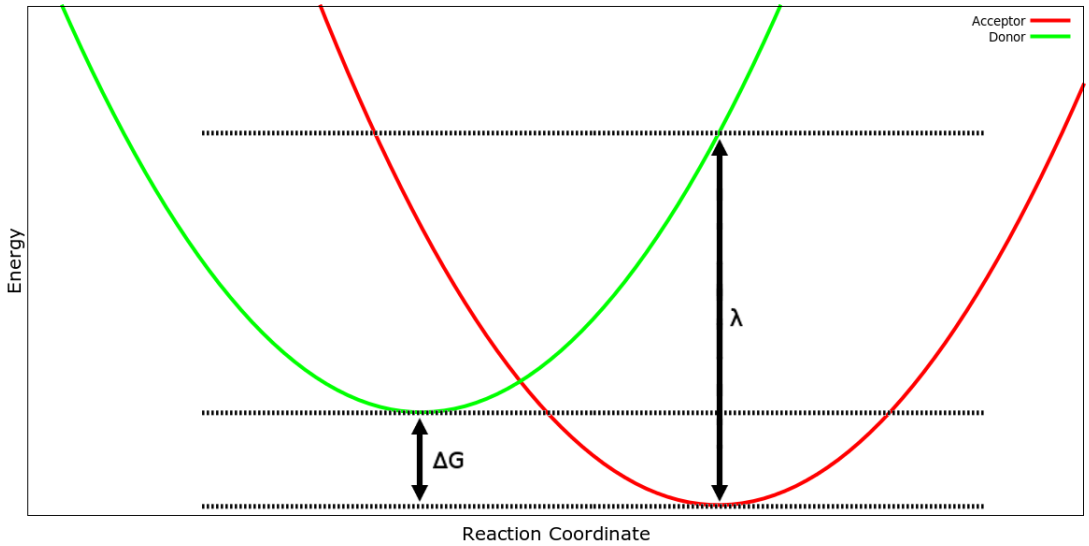


Figure 3.3: Marcus parabolas give a visual representation of the values ΔG and λ. ΔG is the free energy difference between the excited state of the donor acceptor. λ is difference in energy for the excited state and the optimised excited state.

A range of computational methods of varying degrees of sophistication have been proposed for the study of ET, and comprehensive review articles on this subject are available.[106–108] One relatively simple but successful approach is the pathway tunnelling model of Beratan and co-workers.[109,110] This model provides a framework to characterise the influence of the protein structure on V_{DA} in biological ET reactions. We use this approach here as a starting point for more sophisticated TDDFT

calculations and as a benchmark. The model assumes a donor-acceptor complex mechanism and the ET is mediated by consecutive interactions between atoms connecting the donor and acceptor. The steps taken are characterised by decay factors or "penalties" associated with covalent bonds, hydrogen bonds, or through-space jumps. Covalent-bond mediated steps are deemed to have a low tunnelling barriers and hence assigned a low penalty, whilst higher penalties are applied to those with high tunnelling barriers such as a through-space jump. Using a graph-search algorithm[111] to determine all potential pathways, the overall penalty for ET is the product of each penalty throughout every step[112]

$$V_{DA} = A \prod_i \epsilon_i^C \prod_j \epsilon_j^H \prod_k \epsilon_k^S \tag{3.2}$$

where ϵ_i^C is the penalty for propagation through covalent bond i, ϵ_j^H is the penalty for propagation through a hydrogen bond j and ϵ_k^S is the penalty for the through-space jump k. A is the pre-factor, and in this work we take A=1 eV.[113] Using this approach it is possible to identify the strongest ET pathways between the donor and the acceptor, estimate partial electronic couplings mediated by each pathway, calculate the importance of individual protein groups for mediating ET, and ultimately determine the most dominate ET pathway between the donor and acceptor.

The direct calculation of the electron coupling with quantum chemical methods provides a potentially more accurate means for determining the coupling strength. There is a wide range of approaches to this problem including different schemes to characterise the diabatic states[83,106,114,115] along with the choice of electronic structure method. The coupling elements for ET can be evaluated using the Generalised Mulliken-Hush scheme (GMH)[116]

$$V_{DA}^{ET} = \frac{(E_2 - E_1)|\vec{\mu}_{12}|}{\sqrt{(\vec{\mu}_{11} - \vec{\mu}_{22})^2 + 4|\vec{\mu}_{12}|^2}} \tag{3.3}$$

where $\vec{\mu}_{11}$, $\vec{\mu}_{22}$ and $\vec{\mu}_{12}$ represent the permanent and transition dipole moments for

the two states.

An advantage of a quantum chemistry based approach is that it is also possible to study the rates of EET, and an overview of the theoretical treatment of EET is available.[117] The coupling elements associated with EET can be calculated through the fragment excitation difference (FED) approach.[83]

$$V_{DA}^{EET} = \frac{(E_2 - E_1)\overline{\Delta x}_{12}}{\sqrt{(\Delta x_{11} - \Delta x_{22})^2 + 4\overline{\Delta x}_{12}^2}} \tag{3.4}$$

where Δx represents elements of the excitation difference matrix. Here we are concerned with EET occurring over long range through dipole-dipole coupling in a Förster mechanism. Computational methods for the calculation of EET via a Dexter mechanism, which requires overlap of the wavefunctions of the donor and acceptor, have been described elsewhere.[118]

Evaluating the terms in equation 3.1 requires excited state electronic structure methods, including structural optimisation of the relevant states. This is commonly done using methods such as single excitation configuration interaction (CIS) or TDDFT.[106,119–125] The study of large biological systems is challenging, since the size of the system usually precludes a full treatment with quantum chemical methods, necessitating some further approximations. Despite this inherent complexity, several studies have reported calculations of ET in proteins based upon Marcus theory.[126] One approach has been to compute coupling matrix elements for structures extracted from a molecular dynamics simulation.[127–129] Recent work has studied the ultrafast ET in cryptochromes based upon GMH coupling strength computed using TDDFT.[130] Furthermore, the absorption and fluorescence spectra of tryptophan residues embedded in the protein environment have been studied using TDDFT calculations.[131,132] Studies of EET of large biological systems have also been reported.[133–135] This includes application of the FED approach with TDDFT

to study the electronic energy transfer pathways in cyanobacteria phycocyanin[134] and cyanobacteria allophycocyanin.[135]

In this study we aim to provide insight into the structural factors that affect the rates of both ET and EET in the protein myoglobin. The availability of experimental data for the rates of tryptophan → heme ET and EET allows the accuracy of TDDFT based calculations to be assessed. Treating ET and EET in an equivalent way, and achieving good agreement with experiment observations, provides a sound basis for the physical basis for the ET and EET rates observed in experiment to be explored with confidence.

3.2 Methodology

The donor-acceptor coupling elements, V_{DA}, for the both Trp14→heme and Trp7→heme ET processes were computed employing the pathways tunnelling method based upon the coordinates of the crystal structure (PDB: 1YMB) using the software of Balabin *et al.*[111] The electron coupling was also evaluated using TDDFT with the GMH scheme for electron transfer and the FED approach for excitation energy transfer. The TDDFT calculations, using the long-range corrected CAM-B3LYP exchange-correlation functional,[58] were performed on reduced structural models based upon the crystal structure coordinates. The location of hydrogen atoms were constructed using the IQMOL software. Key intervening residues that were identified based upon the pathways tunnelling calculations were included in the TDDFT calculations. To explore the sensitivity of the computed coupling to fluctuations in structure, further couplings were computed for a range of crystal structures (PDB: 1BZR, 2MB4, 4MBN, 1YMB) including those recorded in time-resolved serial femtosecond crystallography experiments (PDB: 5CN4 at 0, 50 and 150 ps).[136]

In the 2D-UV experiment, ET or EET transfer occurs following the initial elec-

HOMO → LUMO	CIS	M06-HF	ωB97-X	CAM-B3LYP	Exp.[a]
^{14}Trp → ^{14}Trp	4.39	4.29	4.73	4.12	∼4.20
^{7}Trp → ^{7}Trp	4.20	3.99	4.52	4.03	∼4.20
Heme → Heme	3.60	2.84	2.96	2.37	1.9 - 2.2
^{14}Trp → Heme	3.52	3.64	3.78	4.06	
^{7}Trp → Heme	3.44	3.55	3.74	3.93	

Table 3.1: Computed excitation energies in eV with CIS and TDDFT with different exchange-correlation functionals. [a]Experimental data from references[1,5,6]

tronic excitation of the tryptophan. Here we study the electron transfer from the tryptophan residues to the porphyrin ring to form an Fe(II)-porphyrin π-anion radical. Subsequent ET from the ring to iron to form the ferric heme is expected to be sufficiently fast that it does not affect the overall rate. A schematic of the electronic states involved is given in Figure 3.4. It shows that to describe the ET and EET processes requires calculations on the ground state, local excited state of the tryptophan, local excited state of the heme and a tryptophan→heme charge transfer (CT) excited state. The accurate description of the electronic structure of metal-porphyrin systems is a challenge for computational methods, where the prediction of the correct ground state electronic configuration is debated.[137–139] Multiconfigurational perturbation theory based calculations in conjunction with large basis sets represent the most reliable methods, however, the computational cost of these methods makes them unsuitable for the current study and the calculations presented here are based upon DFT and TDDFT. These calculations find the lowest energy spin state of the iron porphyrin ring to be a triplet state, which is computed in our work to lie about 1.8 eV lower in energy than the singlet state at the TDDFT level of theory. This is consistent with previous DFT studies.[139] The excited states arise from local excitations between the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) of the tryptophan and a CT excitation from the HOMO of tryptophan to the heme. These orbitals correspond to π and π^* orbitals localised on the tryptophan residues or the heme. Table 3.1 shows excitation energies for the relevant states computed with CIS and a range of exchange-correlation functionals that are suitable for describing CT excitations for

a model system that includes the two tryptophan residues and the heme in their geometries taken directly from the crystal structure. The CAM-B3LYP functional predicts excitation energies for the HOMO → LUMO transitions of Trp14 and Trp7 to occur at 4.12 eV and 4.03 eV, respectively, which are consistent with the value of 4.3 eV measured for tryptophan in the gas phase.[5] The HOMO → LUMO transition localised on the heme lies at 2.37 eV. This transition corresponds to the Q band in porphyrin related compounds, which lies in the region 1.9 - 2.2 eV.[1,6] CAM-B3LYP performs best of the functionals considered, particularly for Q band excitation of the heme.

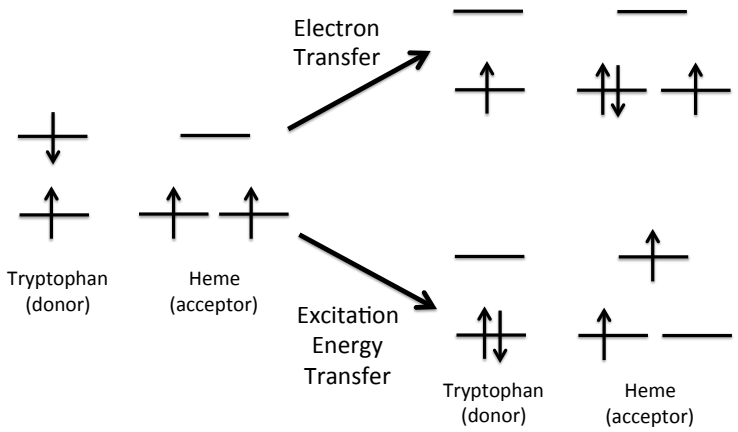


Figure 3.4: Schematic of the electronic configurations for ET and EET processes from the S_1 state of a tryptophan residue to the heme.

All calculations were performed with Q-CHEM[140] and used the 6-31G* basis set. Calculations using larger basis set 6-311G* did not show significant difference in the computed values for the smaller systems studied here, and hence the less demanding basis set was used. This observation is consistent with previous simulations of the electronic spectra of porphyrin.[141]

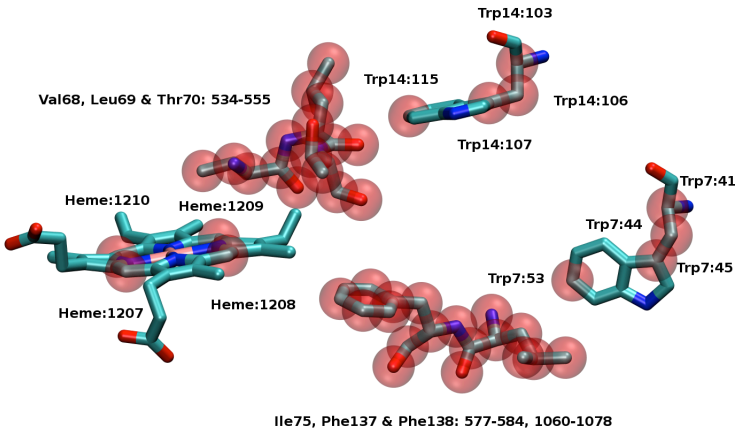


Figure 3.5: Atoms highlighted in red are held fixed in position during the optimisation of the structure of the heme, intervening residue and tryptophan.

Residue	Index in PDB (1YMB)
Trp7	41,44,45,53
Trp14	103, 106, 107, 115
Val68	534-540
Leu69	541-548
Thr70	549-555
Ile75	577-584
Phe137	1060-1067
Phe138	1068-1078
Heme	1207,1208,1209,1210

Table 3.2: Atoms frozen in the geometry optimisations.

One of the most computationally challenging aspects of the application of Marcus theory is the structural optimisation of the initial and final states. These states are electronically excited states, and in our calculations we assessed optimising the structures using several methods: CIS, TDDFT and the maximum overlap method (MOM).[73] The S_1 excited state geometries for the heme predicted by the three excited state methods, CIS, TDDFT and MOM, show a largest RMSD between any two structures of 0.04 Å. MOM provides accurate predictions for excited state structures[142] and was chosen as the preferred method of excited state optimisation, since it was the most straightforward to apply. When optimising the structure of high-lying excited states with TDDFT or CIS, in many instances the order of the roots changed between optimisation cycles making it necessary to monitor the nature of the excited states constantly to ensure that the correct state was being optimised.

An unconstrained optimisation of the reduced model of the protein would yield an unphysical structure. Consequently, it is necessary to constrain the optimisation in some way in order to maintain biologically realistic geometry for the residues within the protein while allowing for some relaxation of the structure. These constraints represent the physical constraints that the surrounding protein environment would impose on the fragments. During the optimisation a subset atoms, shown in Figure 3.5 and listed in Table 3.2, were kept frozen in position. Under the constraint of these frozen atoms, the two tryptophans are kept a fixed distance from the heme consistent with the crystal structure, but allowed sufficient freedom for relaxation in the optimisation process. The constraints also restrict the macrocyclic of the heme from becoming unduly distorted while allowing porphyrins to dome and ruffle in their excited states.[143,144] The methodology described does not take into account the effects of entropy in ΔG and neglects the role of solvent. These effects may be significant, but to describe them accurately at a quantum chemical level is beyond our current capabilities.

3.3 Results and Discussion

Figure 3.6 shows the strongest ET pathways predicted by the pathways tunnelling model for ET between the tryptophan residues and the heme. The computed pathways for the two tryptophans pass through different parts of the E helix, where key amino acids are Val68 for Trp14 and Ile75 for Trp7. The computed values of the electron coupling term ($|V_{DA}|$) are 8.82×10^{-3} eV and 4.27×10^{-4} eV for Trp14 and Trp7, respectively. These values are consistent with the observation from experiment that ET from Trp14 occurs on a faster timescale than for Trp7.

Tables 3.3 and 3.4 show the computed coupling elements for structural models that incorporate different components of the ET pathway that were identified by

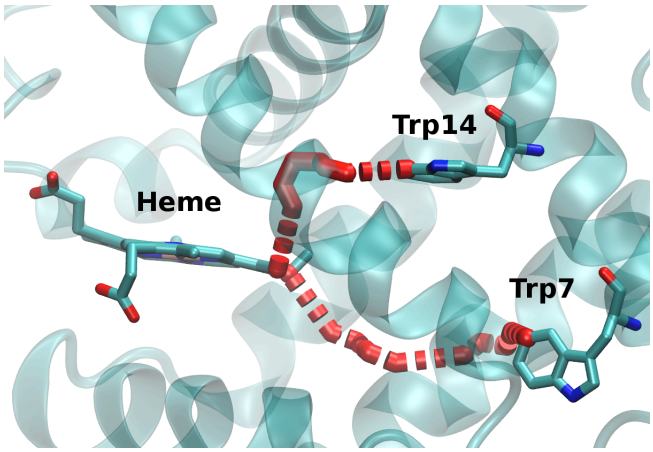


Figure 3.6: Dominant electron transfer pathways for Trp14 and Trp7 to heme calculated by the pathway tunnelling model. For details of the residues see the main text.

the pathways tunnelling model calculation. The TDDFT coupling strengths are computed using the GMH scheme, and correspond to the initial state being the S_1 excited state of the tryptophan residue. The tunnelling pathways model used here makes no distinction between the strength of coupling between the donor in different electronic states, and we will examine the significance of this later. For some of the reduced structural models the dominant pathway is broken resulting in a vacuum tunneling pathway with no significant coupling. The simplest structural model includes only the donor tryptophan and heme. However, intervening residues are likely to affect the computed value for the coupling. We study the importance of the intervening residues by including a single residue (Val68 for Trp14 and Ile75 for Trp7), three residues (Val68, Leu69 and Thr70 for Trp14; Ile75, Leu137 and Phe138 for Trp7) and the full E helix with the additional Leu137 and Phe138 residues between the tryptophan and heme. These residues are shown in Figure 3.7. The calculations show that the inclusion of just a single key residue, Val68 for Trp14 and Leu137 for Trp7, (denoted Trp + Heme + 1AA in the table) has a significant affect on the computed coupling, with a three- to five-fold increase in the strength of coupling. This indicates that Val68 and Ile75 are important in facilitating ET to the heme. The inclusion of further structural elements leads to much smaller additional increases in the coupling.

Model	$\lvert V_{DA}^{TDDFT}\rvert$	$\lvert V_{DA}^{Beratan}\rvert$
Full Protein	-	8.82×10^{-3}
Trp14 + Heme + E helix	7.92×10^{-3}	8.72×10^{-3}
Trp14 + Heme + 3AA (Average)	8.14×10^{-3}	7.71×10^{-3}
Trp14 + Heme + 3AA	6.92×10^{-3}	6.16×10^{-3}
Trp14 + Heme + 1AA	6.91×10^{-3}	3.65×10^{-3}
Trp14 + Heme	2.29×10^{-3}	-

Table 3.3: Computed pathways model ($|V_{DA}^{Beratan}|$) and TDDFT ($|V_{DA}^{TDDFT}|$) coupling values in eV for Trp14 $\rightarrow$ heme electron transfer for different structural models. 3AA and 1AA indicate three and one intervening amino acid residue included in the calculation, see text for details. The average values are evaluated using seven different crystal structures with the Trp14 + Heme + 3AA structural model.

Model	$\lvert V_{DA}^{TDDFT}\rvert$	$\lvert V_{DA}^{Beratan}\rvert$
Full Protein	-	4.27×10^{-4}
Trp7 + Heme + E helix	4.94×10^{-4}	3.69×10^{-4}
Trp7 + Heme + 3AA (Average)	6.62×10^{-4}	5.91×10^{-4}
Trp7 + Heme + 3AA	4.87×10^{-4}	3.29×10^{-4}
Trp7 + Heme + 1AA	4.82×10^{-4}	-
Trp7 + Heme	9.46×10^{-5}	-

Table 3.4: Computed pathways model ($|V_{DA}^{Beratan}|$) and TDDFT ($|V_{DA}^{TDDFT}|$) coupling values in eV for Trp7 $\rightarrow$ heme electron transfer for different structural models. 3AA and 1AA indicate three and one intervening amino acid residue included in the calculation, see text for details. The average values are evaluated using seven different crystal structures with the Trp7 + Heme + 3AA structural model.

The computed TDDFT $|V_{DA}|$ for the structural models with three intervening residues are 6.92 $\times 10^{-3}$ eV and 4.87 $\times 10^{-4}$ eV for Trp14 and Trp7, respectively. When the donor is the tryptophan in its ground (S_0) state, the corresponding values computed with the same structural model are 6.75 $\times 10^{-3}$ eV and 4.10 $\times 10^{-4}$ eV. Thus, there is an increase in $|V_{DA}|$ in the excited state compared to the ground state. The calculation of $|V_{DA}|$ for the excited state adds significantly to the cost of the TDDFT calculation, since it requires higher energy roots in the TDDFT calculation to be computed. These results suggest that using the coupling value for the ground state is a reasonable approximation for the excited state, but is likely to underestimate the value for $|V_{DA}|$.

The calculated coupling strengths can be sensitive to changes in the structure.

Average values for the coupling computed over seven different crystal structures reported in the literature with the Trp + Heme + 3AA structural model are also shown. The magnitude of the computed couplings vary by at most a factor of two, and the average values are reasonably close to the single structure values. For all structures the qualitative difference between the computed couplings for the two tryptophan residues is observed, and average values of 8.14 $\times 10^{-3}$ eV and 6.62$\times 10^{-4}$ eV are obtained for the TDDFT calculations. The ratio $|V_{DA}|$(Trp14):$|V_{DA}|$(Trp7) is found to be 12.3 from the TDDFT calculations, consistent with the value of 13.0 from the Beratan model.

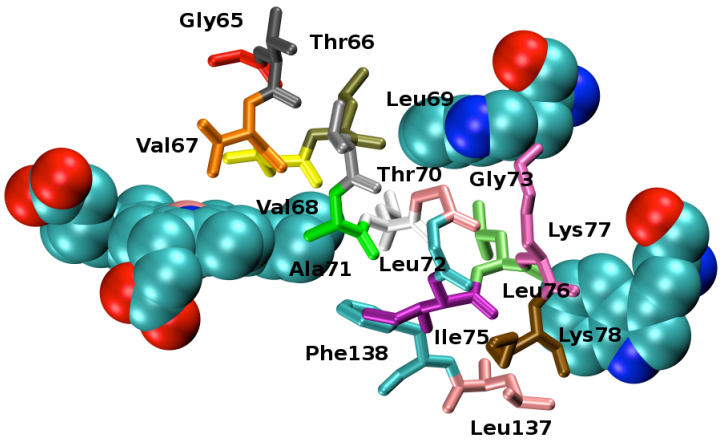


Figure 3.7: The reduced myoglobin system consisting of heme, Trp14, Trp7 and the E-Helix. Different colours represent each amino acid.

It is common to describe the strength of electronic coupling for ET as an exponential dependence on the distance between the donor and acceptor[145]

$$V_{DA}(r) = V_{DA}^{0}(r_0) \exp\left[-\frac{\beta}{2}(r - r_0)\right] \tag{3.5}$$

where r_0 is the van der Waals contact distance, and β is a parameter reflecting the effectiveness of the protein in mediating ET and typically ranges from 1.10 to 1.65

	Trp14(S_0)	**Trp14**(S_1)	**Trp7**(S_0)	**Trp7**(S_1)
$\lvert V_{DA}^{ET}\rvert$	6.75 $\times 10^{-3}$	6.92 $\times 10^{-3}$	4.10 $\times 10^{-4}$	4.87 $\times 10^{-4}$
$\lvert V_{DA}^{EET}\rvert$	2.73 $\times 10^{-4}$	3.13 $\times 10^{-4}$	6.25 $\times 10^{-3}$	7.08 $\times 10^{-3}$

Table 3.5: Computed TDDFT (CAM-B3LYP/6-31G*) coupling values in eV for ET and EET for the Trp + Heme with three intervening amino acid residues model.

Å^{-1} for condensed phase systems and from 3 to 5 Å^{-1} for electron tunnelling across a vacuum.[88] Using the values of $|V_{DA}^{TDDFT}|$ for Trp14 and Trp7 computed here gives a value for β of 0.8 Å^{-1}. This is close to typical values for this parameter for condensed phase systems, and suggests that the protein is effectively mediating ET and the slower rate of ET for Trp7 is associated largely with its greater distance from the heme (22.6 Å compared with 15.9 Å for Trp14 for the 1YMB crystal structure).

$|V_{DA}|$ for EET have been computed using the FED scheme with three intervening residues (Trp + Heme + 3AA) for both the S_0 and S_1 initial states of tryptophan. These values and the corresponding values for ET are summarised in Table 3.5. The key change for EET compared with ET is that $|V_{DA}^{EET}|$ for Trp7 is larger than the value for Trp14. If the intervening residues are removed from the calculation, the computed couplings are 5.68 $\times$ 10^{-3} eV and 2.56 $\times$ 10^{-4} eV for Trp7 and Trp14, respectively. This represents only a modest change in the relative coupling strengths, whereby the strength of coupling for Trp7 remains about 23 times larger than for Trp14. This suggests that the intervening residues do not play a key role in the qualitative difference in the rates of EET between the two tryptophan residues. Similar to ET, there is an increase in the strength of coupling for the S_1 state compared to the S_0 state for EET.

Another potentially important factor that determines the strength of the coupling is the orientation of the tryptophan residues with respect to the heme. To investigate this, the strength of coupling has been computed with the Trp7 rotated such that it has the orientation of Trp14 and conversely Trp14 was rotated to have the orientation of Trp7. This is illustrated in Figure 3.8. This increases the coupling

of "Trp14" from 2.56 $\times$ 10^{-4} eV to 2.85 $\times$ 10^{-2} eV, and decreases the coupling of "Trp7" from 5.68 $\times$ 10^{-3} eV to 7.80 $\times$ 10^{-5} eV, and thus accounts for the qualitative difference in the EET of the two residues. There is an increase in the modified coupling strengths indicating that distance from the heme does play a role, but the effect of orientation is orders of magnitude greater. This sensitivity to the orientation of the donor residues is consistent with EET occurring by a Förster mechanism.

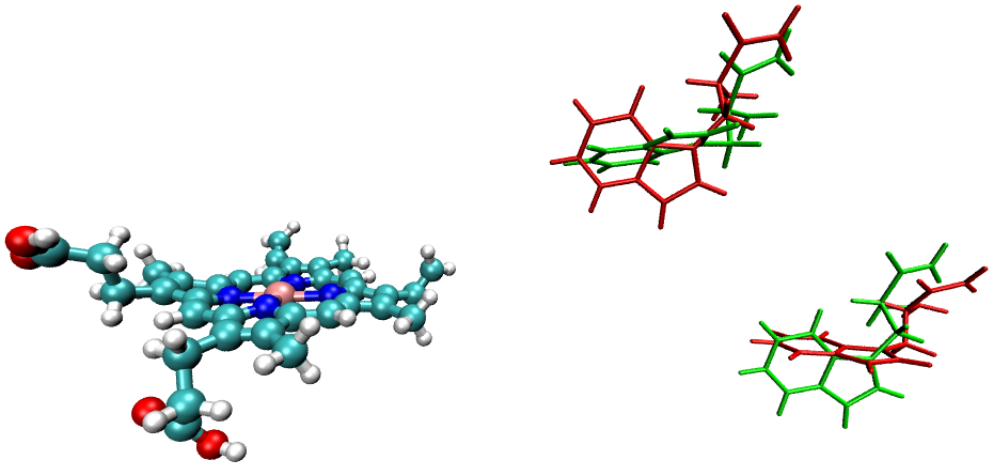


Figure 3.8: Modified heme and tryptophan system. Original tryptophan orientations shown in green and modified orientations shown in red.

In order to compute the rates ET and EET and allow a direct comparison with experiment, the reorganisation energy (λ) and change in free energy (ΔG) need to be determined (equation 3.1). This requires structural optimisation of the excited states corresponding to the final and initial states of the ET and EET processes. Based upon optimisations of the excited states using the MOM approach with CAM-B3LYP/6-31G*, we computed values of ΔG and λ of -0.16 eV and 0.69 eV for Trp14 and -0.06 eV and 0.55 eV for Trp7 for ET, and values of ΔG and λ of -1.46 eV and 0.83 eV for Trp14 and -1.38 eV and 0.72 eV for Trp7 for EET. The calculated values of λ are of similar magnitude to those reported for other related systems.[146]

System	τ_{ES}	τ_{GS}	τ_{hybrid}	Exp.
ET: Trp14	42	60	47	34
ET: Trp7	12000	32000	15000	40000
EET: Trp14	54000	70000	-	-
EET: Trp7	374	480	-	140

Table 3.6: Calculated relaxation times in ps. For τ_{ES} the $|V_{DA}|$ is computed using TDDFT for the S_1 state, for τ_{GS} $|V_{DA}|$ is computed using TDDFT for the S_0 state and for τ_{hybrid} $|V_{DA}|$ is evaluated using the pathways tunnelling model.

Through combining the computed ΔG and λ with the coupling strengths $|V_{DA}|$, the rates and relaxation times for ET and EET ($\tau_{ET} = 1/k_{ET}$ and $\tau_{EET} = 1/k_{EET}$) can be evaluated, and three sets of relaxation times are given in Table 3.6. Overall, the calculation that most closely corresponds to experiment evaluated the relaxation time through combining the computed ΔG and λ with $|V_{DA}|$ computed with TDDFT for the S_1 state, and this relaxation time is denoted τ_{ES}. The remaining two relaxation times correspond to where ΔG and λ are combined with $|V_{DA}|$ computed with TDDFT for the S_0 state of tryptophan (denoted τ_{GS}) and with $|V_{DA}|$ evaluated using the pathways tunnelling model (denoted τ_{hybrid}). τ_{GS} also reproduce the experimental rates well and shows that the additional computational effort to evaluate the excited state coupling strengths could be avoided. The computed relaxation times reproduce the key observations made in the experiment. For ET the relaxation time for Trp14 is much faster than for Trp7, and for EET the relaxation time for Trp7 is much faster than for Trp14. The qualitative description of the computed rate is heavily influenced by V_{DA}, and it is possible to account for the experimental observations based solely on comparing coupling values, $k_{ET} \propto |V_{DA}|^2$. Addressing some of the approximations made in the calculations primarily the structural models used, neglect of entropy in ΔG and the neglect of solvent could lead to a more precise quantitative agreement with experiment.

3.4 Conclusions

The rates of tryptophan→heme electron transfer and excitation energy transfer in myoglobin have been studied using a combination of DFT and TDDFT. These rates have been measured in recent 2D-UV spectroscopic experiments by Chergui and co-workers[95] providing an opportunity to assess the accuracy of different computational models and probe structural factors that affect the rates. Application of the tunnelling pathways model shows that the important intermediate residues for electron transfer are Val68 and Leu69 for Trp14 and Ile75 for Trp7, and inclusion of these residues is important in TDDFT calculations of the coupling matrix elements. Both the pathways tunnelling model and TDDFT calculations correctly predict diabatic electron coupling matrix elements consistent with the rate of ET for Trp14 being greater than for Trp7. The predicted rate is greater for an initial S_1 electronic state of the tryptophan donor compared to the ground state. With TDDFT it is possible to extend the study to consider EET, and the calculations correctly predict that the rate for EET is greater for Trp7.

Marcus theory calculations using the computed electron coupling elements for ET and EET combined with λ and ΔG evaluated from quantum chemical calculations of the appropriate excited states gives relaxation times in good agreement with experimental measurements. Subsequent analysis of the structure shows that the different rates of electron transfer from the two tryptophan residues can be associated with the distance between the heme and tryptophan residues, while for excitation energy transfer the orientation of the tryptophan residues relative to the heme is important.

Chapter 4

Absorption and Fluorescence Spectroscopy of Porphyrins

The absorption and fluorescence spectroscopy of a series of porphyrin based systems have been studied. The range of systems has been selected in order to investigate the influence of both the electronic and geometric configuration on the photophysical properties. The origin of the bathochromic shift in the absorption and fluorescence spectra of substituted porphyrins is attributed to both steric distortions of the ring and electronic effects of the substituents. Three DFT based approaches have been used to model and calculate these properties. The approach using the MOM predicted the largest discrepancy from experimental results whilst TDDFT calculated shifts within 0.05 eV of experimental values. Finally a third method labelled as a 'hybrid' approach has been used, where the MOM is employed to optimise excited state geometries and single point TDDFT calculations are used to evaluate the vertical excitation energies. This approach improves on the excitation energies predicted by the MOM but does not improve on the values that a full TDDFT calculation produced. However, this 'hybrid' approach is computationally less demanding. There is distinct trade-off between accuracy and feasibility of calculations, where this 'hybrid' method of MOM and TDDFT becomes beneficial and useful.

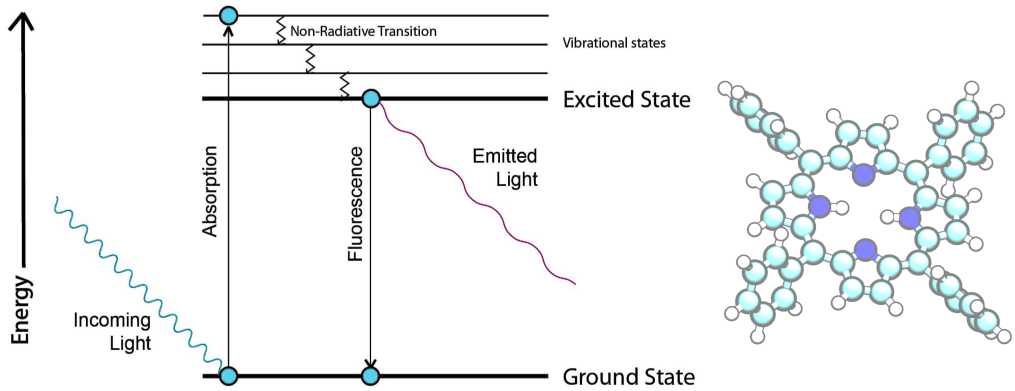


Figure 4.1: Modified Jablonski diagram showing in a schematic way the energies of the various states of the porphyrin molecule.

4.1 Introduction

The ability to perform quantum mechanical calculations for very large chemical systems has made significant progress over the last 50 years due to the exponential growth of computational resources.[147–149] However, the time required to complete the calculations can be lengthy, taking days to even months to complete, and the financial costs of operating the large computational resources can become prohibitive. A 2015 report[150] found that the top petascale computing platforms consume several gigawatts of electricity every year, a significant drain on financial resources. Even modestly-sized computational clusters can cost a substational amount a year to power. Thus, using these high performance computing platforms in the most efficient manner, both in terms of power and in the throughput of computations within a particular power constraint has become necessary in scientific computing[150].

Porphyrins and porphyrinoid based compounds have assumed extraordinary importance as the potential use of these species in both biological and non-biological systems, could have an impact in a plethora of applications across many scientific fields which are discussed in detail later in this section.[151,152] Porphyrins are ubiquitous throughout nature and attract widespread attention due to their ability to accumulate in many kinds of cancerous cells as well as having numerous important

magnetic and optical properties. Porphyrins are molecules, whose physicochemical properties can be easily adjusted by modifications of the electronic distribution on the aromatic ring through peripheral substitutions.[153] Studies of the properties of these compounds and their multichromophoric derivatives have grown rapidly as new and more sophisticated techniques become readily available.[150] However due to the sheer size of the porphyrinoid compounds and complicated electronic structure of metalloporphyrins it is challenging to study these systems from a theoretical viewpoint.

Porphyrins are macrocyclic organic molecules which have an extensive system of delocalized π electrons. They can be modified by connecting different peripheral substitutes, changing the central metal or expanding the size of the macrocycle. The geometric configuration of porphyrin complexes are modulated by the addition of substituents which also impact on the electronic energy levels. Depending on the extent of the disruption to macrocyclic network this can lead to heavily distorted non-planar porphyrin structures seen in figure 4.2. The influence of the substituents, the central metal ions and structural configuration are reflected in the photophysical properties.[154]

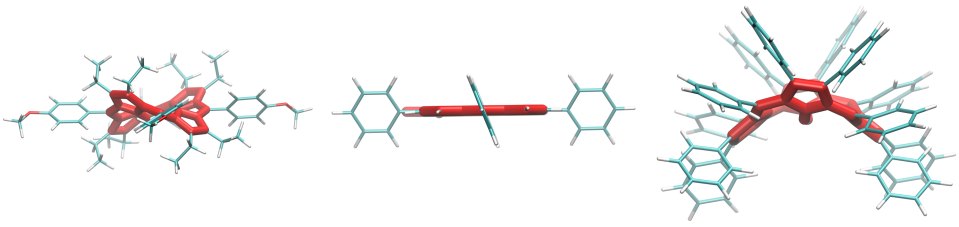


Figure 4.2: The different configurations of porphyrins. Left: Saddled, Middle: Planar, Right: Domed.

The accepted interpretation[155] is as follows: there is strong absorbance in the near-UV called the Soret or B band and denotes the $S_0 \rightarrow S_2$ transition. There are three other absorption bands in the deep UV known as the N, L and M bands. In

the visible region of the spectrum there are four weak bands called the Q bands arising from $S_0 \rightarrow S_1$ transitions. The fluorescence emission spectrum consists of two strong emission bands at λ Q_x.[156] On closer inspection there are also other smaller fluorescence bands at wavelengths shorter and longer than the main bands.[157,158] The spectra in figure 4.3 is typical of many free-base porphyrins. When studying absorption and fluorescence it is common to make use of the Stokes shift. This is the difference in wavelength between positions of the band maxima of the absorption and emission spectra of the same electronic transition.[159]

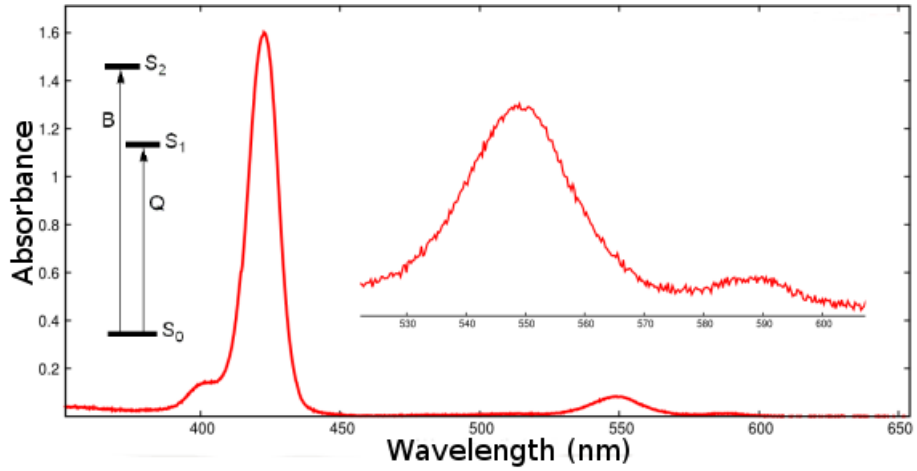


Figure 4.3: Absorption spectra of metalloporphyrin. The 'B' band occurs around 400 nm and is attributed to an $S_0 \rightarrow S_2$ transition. The 'Q' band occurs around 500 nm and is a result of an $S_0 \rightarrow S_1$ transition. Due to the relative weaker intensity it is highlighted in the figure.[1]

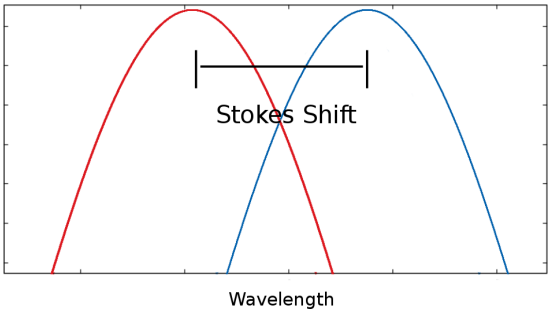


Figure 4.4: The Stokes shift which is the difference between the emission (blue) and absorption (red) maximum.

The chemistry behind porphyrins is directly mediated by the dominant presence

of electron transfer reactions. A rationale for this is provided with Gouterman's four orbital model which describes the electronic transitions from the highest occupied molecular orbital (HOMO) and HOMO-1 to the lowest unoccupied molecular orbital (LUMO) and LUMO+1 as π to π^* ring transitions (Figure 1.7, Chapter 1), demonstrating similar properties to the quintessential aromatic benzene.[156] A justification towards treating porphyrin this way is that with the addition of peripheral substituents the route of electron delocalisations is altered which shows similarities to the effects seen in benzene.[160] The additional functional groups alter the polarity and spread of electron density which has a significant impact on the physical properties modifying the geometric and electronic behaviour.[161,162]

As porphyrins are such versatile and useful molecules they attract significant attention and many studies have focused on their applications which has been extensively reviewed.[151] Photodynamic therapy (PDT) is a commonly applied method for the treatment of *in vitro* cancerous behaviour. Successful application depends on three requirements being satisfied: the presence of oxygen, a photosensitizer and a light source emitting waves coherent with the absorption maximum relative to the utilised photosensitizer.[163] Separately all these components are nontoxic; however, when combined they have potential to destroy cells deemed cancerous which provides a significant advantage over other contemporary treatments such as chemo- and radiotherapy. The opportunities extend even further than oncological applications and PDT is being appreciated in a diverse selection of medical fields.[164]

The applications of porphyrins are not limited to medicine. Porphyrins are popular within organic research as they are successful catalysts for the oxidations of many organic compounds. This is due to the fact that they posses a unique structure which can faciliate electron transfer which is further emphasized by the addition of a transition metal positioned in the centre of the porphyrin unit. Porphyrin based compounds also attract significant attention for the use in electronic

engineering as molecular wires which are well studied due to them being very effective conductors.[35,165] The dimerised porphyrin system is also popular in the use of dye-sensitized solar cells (DSSCs). Two porphyrin monomers connected via a bridge allows efficient electron transfer to occur.[166] The choice on the linking bridges further expands the potential of porphyrins.[167]

The distortion of the macrocycle from a series of sterically hindered porphyrins results in unusual physical and optical properties. These are known as 'doming', 'saddling' and 'ruffling' which are all recognised features of porphyrin (shown in figure 4.2).[168,169] These distorted geometries result in differing bathochromic shifts of the Q-band in both the absorption and flouresence spectrum. This has generated the discussion of whether the properties are a result of the distorted ring geometry or are they attributable to the electronic effects of the substitutents which are known to have directing effects on the aromatic ring.[170]

A range of porphyrin based systems has been studied in order to account for all possible substituent configurations.[7] Figure 4.5 describes the eight systems studied. These were chosen so that any feature that will have significant impact on the energies of the excited states and the photophysical properties would be highlighted. Both the free base and metallo variant of each system are studied in order to observe the effects of both altering the symmetry and also disrupting the conjugation network with a transition metal.[171]

As discussed porphyrins have been the subject of many experimental studies.[152] They have also been studied computationally but due to the sheer size of some of the systems it is difficult to study at a high enough level of theory without access to large scale high performance computing. The free base porphyrin monomer has received significant attention from computational studies where almost every level of theory has been used with varying complexities of basis sets and functionals.[172]

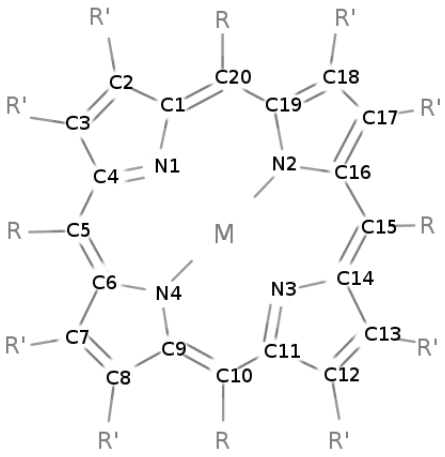


Figure 4.5: Core porphyrin unit with labelled atomic indices.

System	M	R	R'
H_2TPP	H_2	C_6H_5	H
H_2DPP	H_2	C_6H_5	C_6H_5
H_2OETPP	H_2	C_6H_5	CH_3CH_2
H_2T(t-Bu)P	H_2	$(CH_3)_3C$	H
ZnTPP	Zn	C_6H_5	H
ZnDPP	Zn	C_6H_5	C_6H_5
ZnOETPP	Zn	C_6H_5	CH_3CH_2
ZnT(t-Bu)P	Zn	$(CH_3)_3C$	H

Table 4.1: Descriptors for M, R and R' for atomic labels in figure 4.5.

Often studies either focus on how cheap a calculation can be done whilst still maintaining respectable accuracy or they focus on getting the highest level of theory possible. Table 4.3 reports the current status of theoretical porphyrin work relevant to this study. The amount of work reported rapidly declines as the size of the porphyrin system increases so theoretical work reporting accurate exication energies for deca-substituted porphyrin system is rare even more so for emission energies which requires optimising the molecule in the excited state, which are very demanding and resource intensive calculations.

With the benefit of computational studies, the HOMO's and the LUMO's can be studied. This aims to provide useful insights as to why porphyrins display such features and how they may be exploited. As experimental methods advance it is important that accurate methods for modeling electronically excited states de-

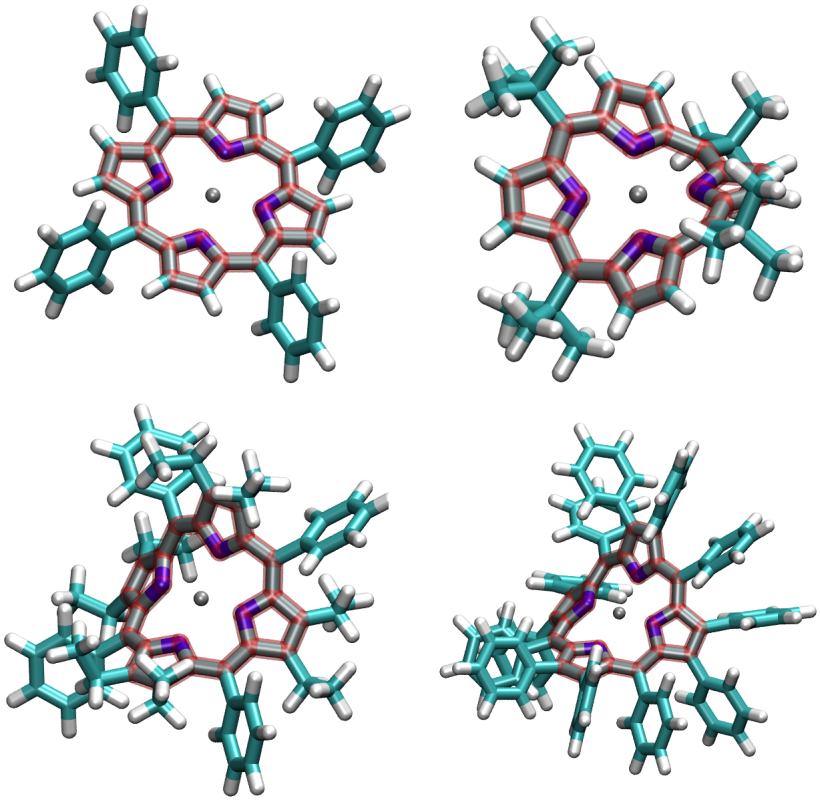


Figure 4.6: The metal versions of the porphyrin structures with the base unit highlighted. (Top Left: Zinc Tetraphenylporphyrin (ZnTPP) Top Right: Zinc-tetra-tert-Butyl Porphyrin (ZnT(t-Bu)P) Bottom Left: Zinc-octaethyl-tetrephenylporphyrin (ZnOETPP) Bottom Right: Zinc-dodecaphenylporphyrins (ZnDPP).

Energy	Band
2.16	Q_x (1-0)
2.56	Q_y (1-0)
3.33	B
3.65	N
4.25-4.67	L

Table 4.2: Experimentally observed excitation energies of free base porphyrin in eV.

B3LYP	CAM-B3LYP	CIS	CIS(D)	SAC-CI level 1	SAC-CI level 3	RICC2	SCS-RICC2
2.26	2.20	2.39	2.46	1.67	1.69	2.30	2.13
2.42	2.42	2.50	2.61	2.06	2.18	2.50	2.51
3.30	3.55	4.44	4.13	3.41	3.41	4.38	3.45
3.44	3.65	4.62	4.45	3.57	3.53	4.72	3.78
3.79	4.26	5.23	4.97	4.13	4.19	5.23	4.38
4.05	4.79	5.60	5.30	4.68	4.75	6.54	4.66

Table 4.3: Calculated excitation energies of free base porphyrin using various levels of theory. Results presented are gas phase and in eV.

velop alongside. Two DFT based methods TDDFT and MOM[73] are used. Both of these DFT methods possess their respective strengths and weaknesses discussed fully in Chapter 2. TDDFT has the ability to study several excited states in a single calculation but may be subject to computational limitations particularly for large porphyrin based compounds. The MOM on the other hand offers a cheaper alternative for studying excited states however it is hindered only handling a single exctied state. Also the intial and excited state electronic configuration must be known prior to any calculation hence a standard MOM input is more complicated than the corresponding TDDFT input.

Porphyrins are dynamic systems and their geometric configurations can significantly alter their overall function and efficacy. Due to the complex electronic structures of porphyrins classical dynamics is not sufficient for describing the geometry of the system as even the ground state is subject to much internal charge transfer that dictates the orientation. In order to describe this behaviour it is necessary to use *ab initio* Molecular Dynamics (AIMD).[173]

ab initio quantum mechanical methods may be used to calculate the potential energy of a system on the fly which is necessary for simulation of properties such as excited state dynamics. A significant advantage of using *ab initio* methods is the ability to study reactions that involve breaking or formation of covalent bonds, which correspond to multiple electronic states.[174]

4.2 Methodology

Ground state structures were optimized using unrestricted KS-DFT employing the BPW91 functional and the 6-31G* basis set. Larger basis sets (figure 4.9) were investigated on the smaller computationally more accessible free base porphyrin system. Excited state geometries and energies were obtained using TDDFT, MOM and a

hybrid approach that takes the coordinates of the excited state geometry predicted by MOM and calculates the excitation energies using TDDFT. This is a computationally less demanding and is rationalised by the lack of difference in the predicted structures (RMSD of 0.05Å, Figure 4.7). Calculations have been performed using Q-Chem in both gas phase and solvent (methyltetrahydrofuran (MeTHF)) models (Polarizable continuum model (PCM), ε of 32.61).

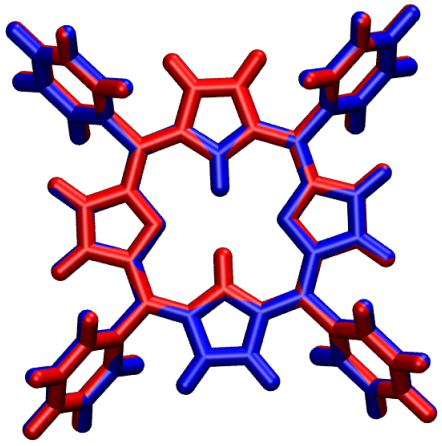


Figure 4.7: A comparison of the excited state geometries obtained using TDDFT (Red) and MOM (Blue). The RMSD between the two optimised systems is 0.05 Å.

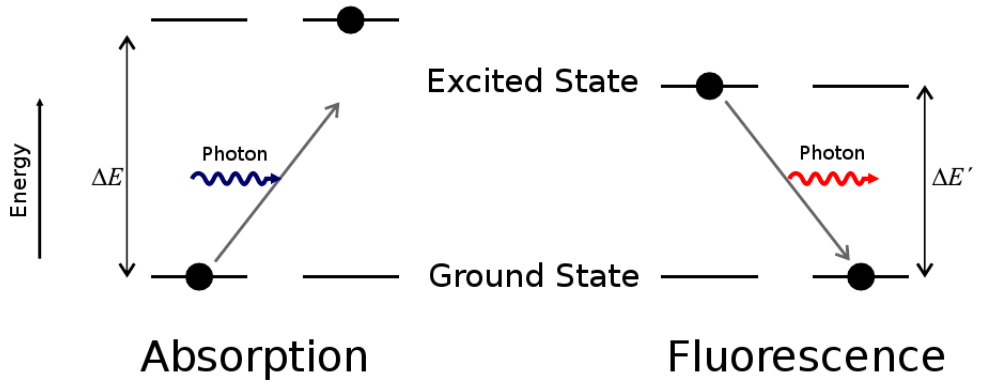


Figure 4.8: A simplified diagram of an electronic interactions at both absorption and fluorescence.

Figure 4.8 highlights the key difference between absorption and fluorescence and gives an overview of the basic principles of these processes. The fluorophore is

described by two electronic states which are separated by ΔE. Absorption occurs when an incident photon hits the target molecule in its ground state. The photon is absorbed during this process and the corresponding energy excites an electron. Fluorescence occurs when the excited molecule returns to its ground state. In most cases the emitted light has a longer wavelength and hence a lower energy than the absorbed radiation. This results in a shift known as the Stokes shift.

$$\Delta\lambda_{Stokes} = \lambda^{max}_{fluorescence} - \lambda^{max}_{absorption} \tag{4.1}$$

In order to calculate absorption a single point vertical excitation energy calculation must be performed. This models an electron from a ground state system being promoted to an excited state configuration as shown in figure 4.8 and predicts an energy. For the fluorescence calculation an excited state optimisation must be employed in order to calculate energy which is the most challenging and computationally demanding part of the calculations. The methods previously discussed are all capable of performing these tasks although they vary in terms of computational practicalities.

The MOM allows excited states to be studied with KS-DFT as the algorithm invoked by the MOM forces the SCF procedure to converge to an excited state solution as opposed to relaxing to its ground state configuration. In order to facilitate this, an initial set of orbitals are generated from a standard single point calculation for the ground state. Here the ground state orbitals were analysed and electronic configuration excited state was specified by moving an α electron from HOMO to the LUMO. Using the MOM procedure variational collapse is prevented in the SCF calculation. This approach is advantageous over TDDFT as the cost of calculation remains consistent with ground state optimizations rather than calculation scaling with respect to the number of excited states to be analysed. It also allows for the orbitals of the excited state of interest to be directly calculated using a ΔSCF approach where excitation energies are simply the difference between the ground and

excited state energies.

This method produces accurate excited state energies for numerous states without any further corrections. However it is known for the MOM to underestimate excitation energies for open-shell singlet states that involve excitations to the valence orbitals. This is attributed to the fact that the open-shell singlet is described with a single determinant resulting in a spin mixed state. Fortunately the energies can be significantly improved by applying the post-SCF spin-purification correction proposed by Ziegler.[74]

$$E = 2E_S - E_T \tag{4.2}$$

where E is the energy of the true or spin-purified singlet state, E_S is the energy of the spin-mixed (single determinant) state, and E_T is the corresponding triple state energy.

Optimised porphyrin structures were studied with AIMD to assess structural changes. The significant advantage of AIMD simulations is their predictive nature, since they are based on quantum mechanics. Developing an empirical interaction potential is no longer needed, which is usually fitted to reproduce some experimental data in classical MD simulations. These are either not available or regarded as unreliable due to the complicated electronic configurations of free base and metalloporphyrins. AIMD was performed using Born-Oppenheimer Molecular Dynamics (BOMD) at constant temperature of 298 K.

The simulation was run for a sufficiently long enough timeframe until patterns emerged. Three known porphyrin configurations emerged in the trajectory: planar, ruffled and domed, oscillating evenly between the geometries. The structures were then sampled at every 20 snapshots which accounted for all detected configurations but significantly reduced the amount of required calculations. The vertical exci-

tation energies of these structures were then calculated using TDDFT. A distinct trade-off between the accuracy of AIMD is the very high computational costs associated with it limiting the systems to be studied to relatively small in comparison to classical MD. Through use of parallel high performance computing these limiting factors become less apparent.

4.3 Results and Discussion

This chapter looks at predicting absorption and fluorescence spectroscopy using *ab initio* methods. A great variety of functionals and basis sets are available for studying these systems with a large choice of levels of theory. The aim is to find an optimum balance between calculating values within a high level of agreement with experimental data but also utilise an affordable level of computational resources. As can be seen in figure 4.9, the results from the more demanding calculations do not significantly effect the accuracy of the final excitation energies with the shifts between energies being around 0.1 eV. The results obtained from the BPW91/6-31G* calculation are already close to results obtained experimentally.

Figure 4.4 shows the λ_{max} values for both the Q and B band for free base porphyrin. Each functional predicted energies that were quite close to the experimental values, but the most accurate results were given by BPW91.[175]

Functional	λ_{max} **Q Band**	λ_{max} **B Band**
B3LYP	2.26	3.30
CAM-B3LYP	2.20	3.55
BLYP	2.14	3.24
M06	2.38	3.50
M11	2.34	3.71
PBE	2.06	3.13
BPW91	2.16	3.33
Exp.[a]	2.16	3.33

Table 4.4: λ_{max} values in eV of the Q and B band for free base porphyrin calculated using a variety of functionals. [a]Experimental data from Ref.[1]

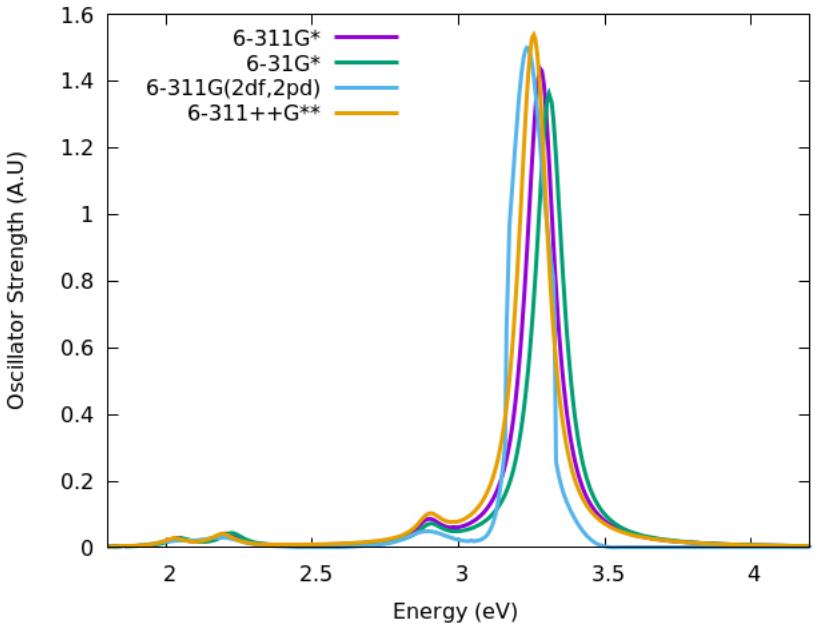


Figure 4.9: A comparison of basis sets for predicting excitation energies for free base porphyrin.

The influence of solvent was also investigated with the PCM solvent model. This was chosen as it provides a realistic description of molecular shape, typically by constructing the solute cavity from a union of atom-centered spheres. In addition, the exact electron density of the solute (rather than a multipole expansion) is used to polarize the continuum. Electrostatic interactions between the solute and the continuum manifest as an induced charge density on the cavity surface, which is discretized into point charges for practical calculations. The surface charges are determined based upon the solute's electrostatic potential at the cavity surface, hence the surface charges and the solute wavefunction must be determined self-consistently. The excitation energies obtained from the PCM calculation (Dichloromethane ε 8.93 and MeTHF ε 32.61) were not significantly different (greatest shift of 0.1 eV) from the gas phase energies but did increase computational cost and were therefore not included in further calculations.

Two DFT based methods and combination of both have been used to calculate excitation and emission energies using 6-31G* basis and BPW91 exchange correla-

tion functional. It becomes clear from the results shown in tables 4.5 and 4.6 that calculating the energies and excited state geometries with TDDFT yields results in closest agreement with experimental values. Hence referring to the TDDFT method as the most appropriate method for calculating absorption and emission is justified. Using solely the MOM produces results which are most distant from the reference values which suggests it as the least accurate method with regards to the three approaches used. Finally the method named as the 'hybrid' approach sits in the middle between these two results. This method takes the geometries predicted via the MOM and calculates excitation energies using TDDFT. This is made possible by the fact that the two excited state geometries predicted are similar in structure with an RMSD of 0.05 Å, as shown in figure 4.7. There is an element of utility to this approach and should not be discounted for its failure to improve on the TDDFT result, the highest level of theory employed in this study.

It is important to consider the overall computational demand that is associated with each approach. The method that employs only TDDFT is the most accurate but it also is the most computationally demanding and indeed does struggle when modelling the larger systems such as the ten substituent zinc based porphyrin system (ZnDPP). As porphyrins contain degenerate and near degenerate states a significant amount of state switching occurs when optimising the higher excited states. It is only with the availability of such great computing resources that results like this are possible. The merit in the 'hybrid' approach is that it significantly reduces the computational expense making larger calculations feasible with only a minor impact on the overall accuracy. Hence it should be treated as an appropriate alternative solution as it can be applied as a method with a broader application.

In order to simulate a physical representation of the biologically active porphyrin, an AIMD simulation was performed. Several well recognised and distinct geometric configurations are seen in the simulations; the planar, the ruffled, the saddled

and the domed configurations. The simulation was allowed to run for 1 ps deemed sufficiently long as step a trend of oscillations between the different configurations was seen which deemed that enough simulation time has occurred to map any observable features. This produced 1000 snapshots as the timestep was set to write new coordinates after every 0.242 fs. Every twentieth snapshot was sampled and vertical excitation energies using TDDFT were then calculated from the trajectory produced by the AIMD. These results are shown in the figure 4.10. A shift is observed in both the Q and the B band whilst the λ_{max} peak at around 370 nm has a relatively broad range in the spectrum spanning around 50 nm but still remains close to the experimental value shown in blue.

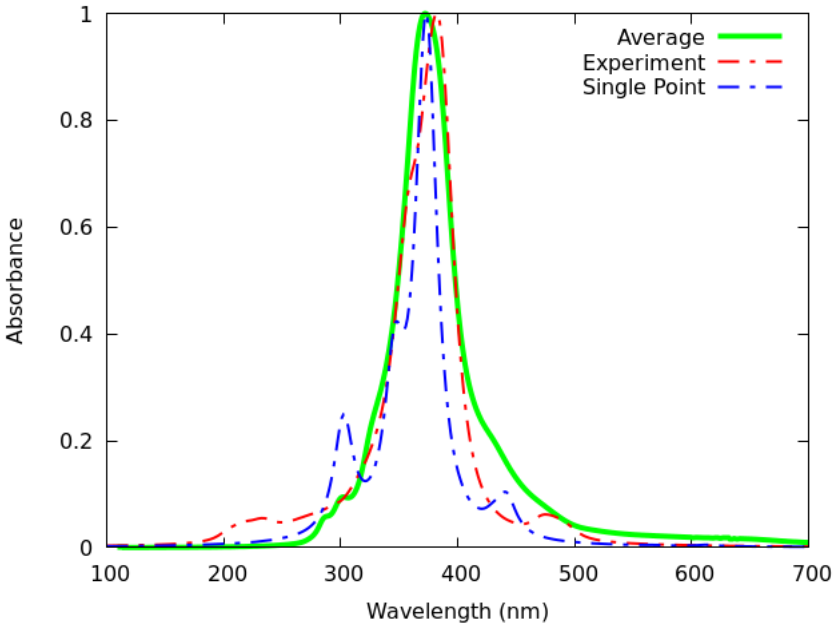


Figure 4.10: A comparison of absorption spectra for average of the results from the AIMD simulations (Green), Single point gas phase (Blue) and Experiment (Red)[1].

In figure 4.10 the average values of all snapshots have been calculated and presented as a single continuous spectrum and compared to both the experimental and gas phase spectra. From this it is evident that averaging over confirmation does not yield significantly different results from the single structure calculations and allows confidence in proceeding with the less demanding single point calculations on the

optimised structures.

	MOM		Hybrid		TDDFT	
System	E_{abs}	E_{ems}	E_{abs}	E_{ems}	E_{abs}	E_{ems}
H_2TPP	2.002	1.875	2.047	1.992	2.047	1.983
H_2DPP	1.490	1.336	1.677	1.515	1.677	1.575
H_2OETPP	1.721	1.521	1.795	1.629	1.795	1.672
H_2T(t-Bu)P	1.801	1.571	1.924	1.785	1.924	1.857
ZnTPP	2.152	1.989	2.219	2.140	2.219	2.160
ZnDPP	1.589	1.220	1.738	1.505	1.738	1.563
ZnOETPP	1.745	1.552	1.850	1.603	1.850	1.646
ZnT(t-Bu)P	1.917	1.467	2.046	1.498	2.046	1.954

Table 4.5: Excitation energies (eV) of the Q band absorption max and the emission maximum

	Optimization Method			
System	**SP-MOM**	**"Hybrid"**	**TDDFT**	**Exp.**[a]
H_2TPP	0.127	0.055	0.064	0.009
H_2DPP	0.154	0.162	0.102	0.112
H_2OETPP	0.200	0.165	0.122	0.095
H_2T(t-Bu)P	0.230	0.140	0.067	0.060
ZnTPP	0.163	0.079	0.059	0.021
ZnDPP	0.369	0.203	0.175	0.169
ZnOETPP	0.193	0.247	0.204	0.206
ZnT(t-Bu)P	0.451	0.347	0.092	0.161

Table 4.6: Calculated Stokes Shifts in eV. [a]Experimental data from Ref.[7].

As can be seen in tables 4.5 and 4.6 the shifts in absorption and emission values predicted closely match experimental results. In figure 4.11 the correlation between theory and experiment is observed. The MOM method has an R^2 value of 0.33 which indicates less correlation to experimental values in comparison to the 'Hybrid' and TDDFT approaches which have R^2 values of 0.73 and 0.74 respectively. It is expected that these two methods would have similar values and further support the use of this 'hybrid' approach as a low cost alternative to the full TDDFT excited state optimisation. The tert-butyl systems yield results which match experiment the closest which is interesting as these complexes have the lowest amount of symmetry and are the most distorted of the porphyrin structures studied. Perhaps the broken symmetry allows for further degrees of freedom in the simulation which is not seen

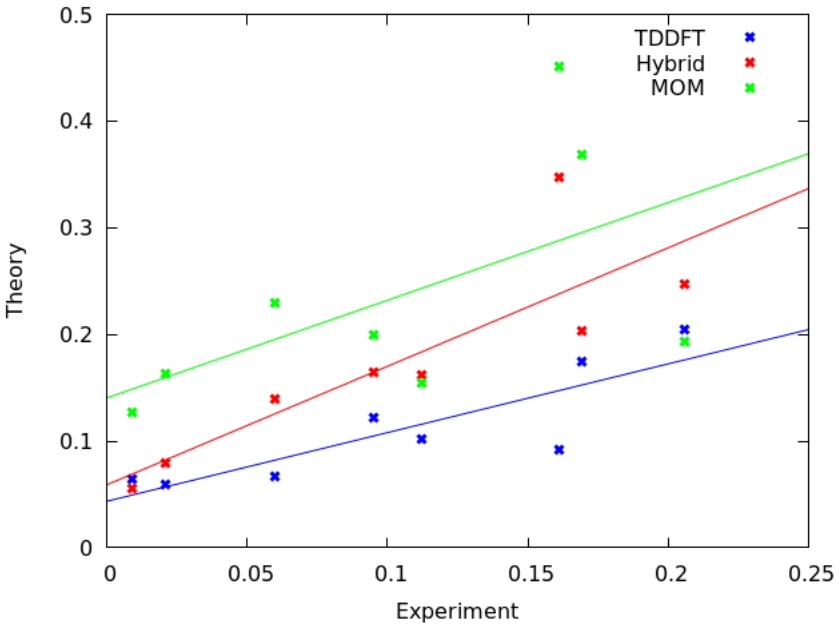


Figure 4.11: A comparison of stoke shifts (in eV) calculated using different methods against experiment. MOM $R^2 = 0.33$, 'Hybrid' $R^2 = 0.73$ and TDDFT $R^2 = 0.74$.

in the other systems. It is seen in table 4.5 that when additional phenyl rings are attached as substituents the overall excitation energy is lowered. This is attributed to the greater electron distribution and added conjugation in the system.

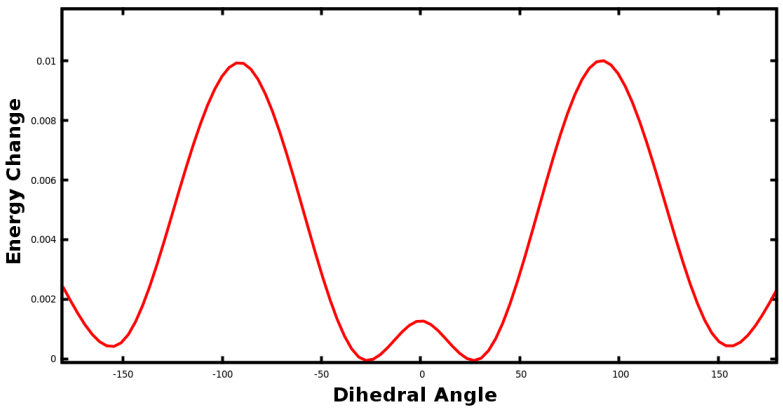


Figure 4.12: Potential Energy Surface diagram for a zinc centred porphyrin dimer where the dihedral angle between the two porphyrin units is rotated through 360°. Energy units are hartrees.

As mentioned a dimerised porphyrin system opens up a host of new opportunities for light harvesting and UV-optimisation. Figure 4.12 shows a potential energy

surface for the optimised structures of the zinc centred porphyrin dimer. The lowest energy configuration is when the two porphyrin monomers exist at a dihedral angle of 27° and its symmetry based counter part -27°. Another minimum is noticed in the energy surface at +/- 153° which would be the same orientation as the 27° system but flipped resulting in a slightly higher energy. The highest energy configuration occurs when the dihedral angle is at 90°. By examining the systems at an electronic level and observing the orbitals of the systems, shown in figure 4.14, it becomes clear why the energy levels are different. The lowest energy configuration shows the charge as an even distribution across the entire system whereas the highest energy configuration shows the charge distributed locally to a single monomer leading to a higher energy. This is seen throughout the whole scan of the system with orbitals becoming more localised to a single monomer as the energy of the system increases.

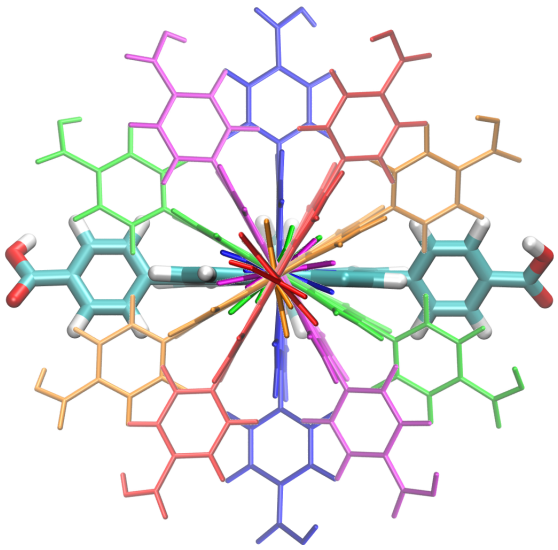


Figure 4.13: A side view showing a porphyrin dimer with dihedral angle between the two monomers being rotated through 360°

The interest in porphyrin dimers is heavily linked with its application towards dye sensitised solar cells. Having a 'dynamic' dimer which maximises absorption at both low and high light intensities is a clear benefit. The UV spectrum for the rotated dimer shows a broad range of absorption going far into the infrared region spanning around 600 nm (figure 4.15). Having this range of light absorp-

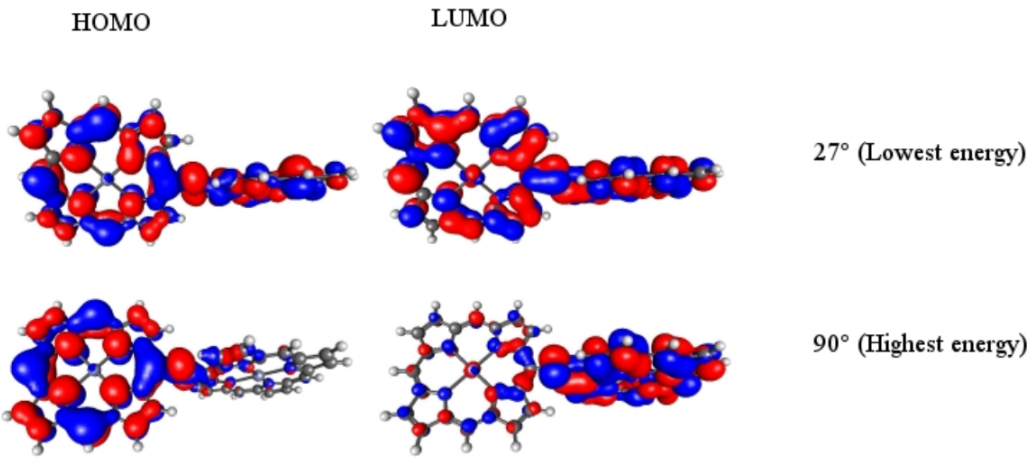


Figure 4.14: The HOMO and LUMO of a zinc porphyrin dimer at the lowest energy (27°) configuration and the highest energy (90°) configuration.

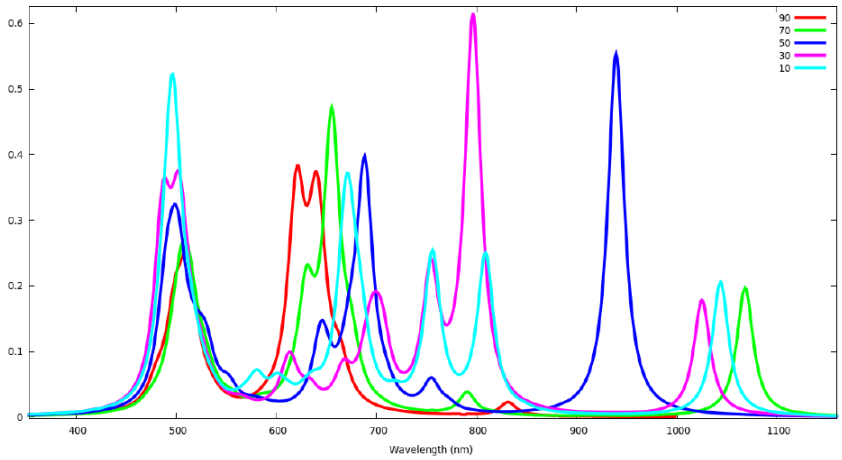


Figure 4.15: The UV spectra of a porphyrin dimer where the two monomers are at different dihedral angles. The graph is 'noisy' but shows how the dimer can be manipulated to absorb light over a large range in frequency.

tion is particularly advantageous towards light harvesting devices. Being able to exploit and manipulate the dihedral angle throughout the day to accommodate for both low and high light intensities can significantly improve performance in DSSC's.

Figure 4.16 shows the HOMO's and LUMO's of all systems studied. As seen the orbitals do not differ greatly between the metallated and free base systems apart from some of the charge being distributed on the zinc. The ruffled system, the tert-butyl complex, shows the most significant difference between the orbitals that suggests that the geometry may disrupt the electron distribution. The ruffled structure

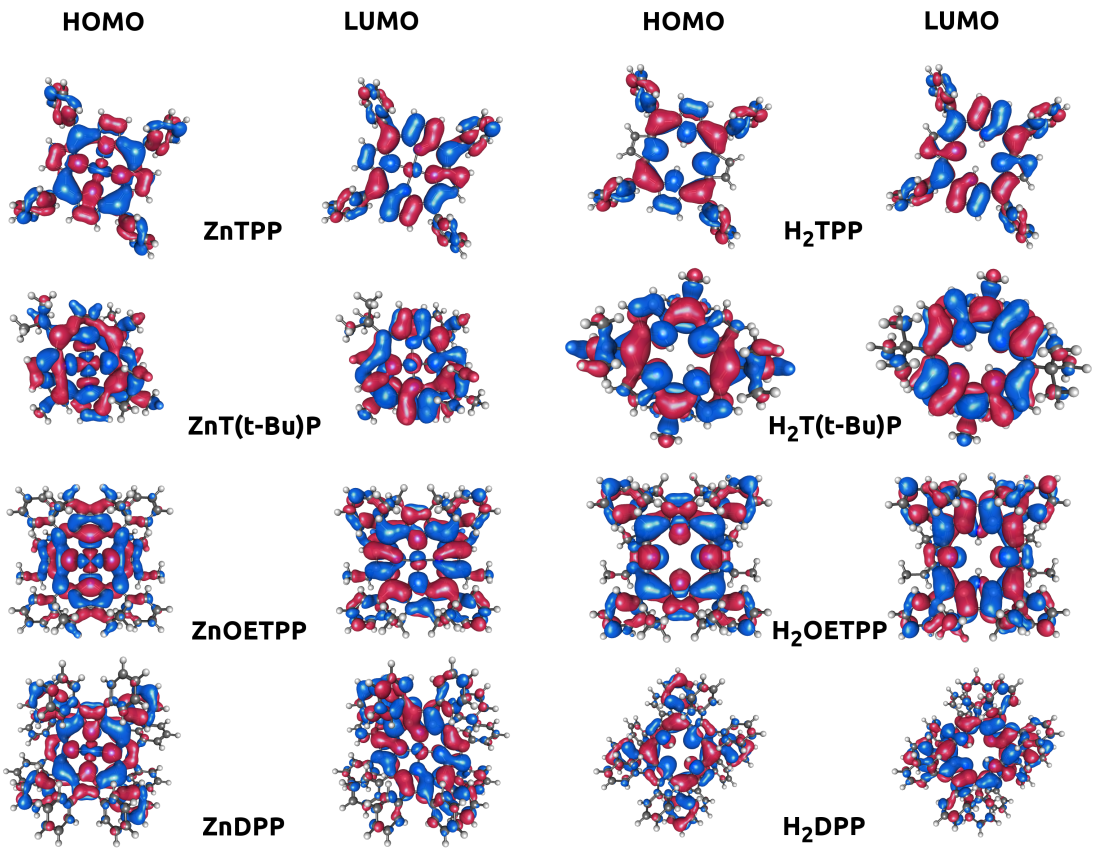


Figure 4.16: Orbital diagrams of the HOMO and LUMO of each of the zinc (left) and free base (right) substituted porphyrin complexes.

messes up the network which is not displayed in the planar or saddled configuration.

Porphyrins are photo-active molecules that are heavily influenced by distortions native to the macrocycle occurring at the centre of the unit. Hence there is a distinct difference between the S_0 and the S_1 states which provides important information with regards to function. Studies of the saddle shaped dodecasubstituted free base porphyrins such as H_2DPP and H_2OETPP have shown that they display a relatively significant shorter lifetime (ca. 400-800 ps)[7] of the lowest excited singlet state in comparison to the planar porphyrin based compounds (10-15 ns)[7]. This can be attributed to the ease that the distortions in the geometry facilitate both internal crossing and intersystem crossing. Structural parameters for both states of all the studied porphyrin complexes have been calculated and key features are shown in figures 4.17 and 4.18. Expected absorption spectra for geometric characteristics

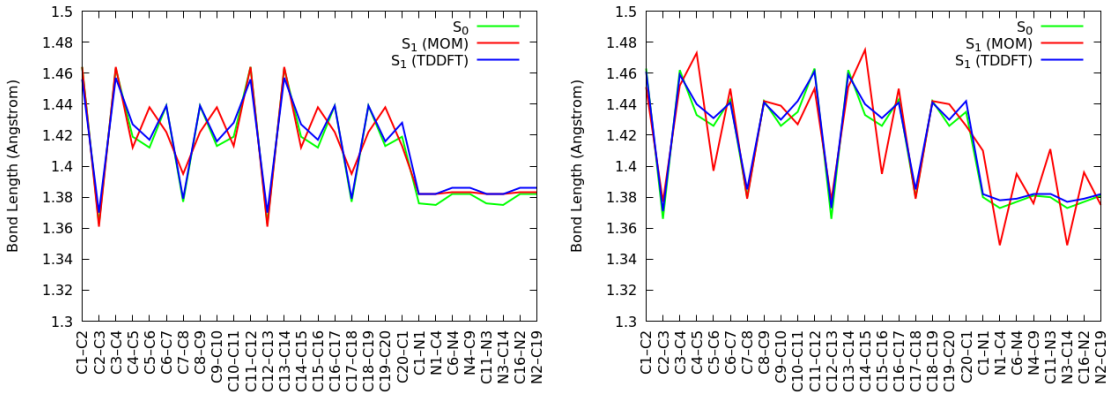


Figure 4.17: Optimised structural parameters of H_2TPP (left) and H_2T(t-Bu)P (right) of the S_0 and S_1 states for both the MOM and TDDFT methods. For atomic indices see figure 4.5.

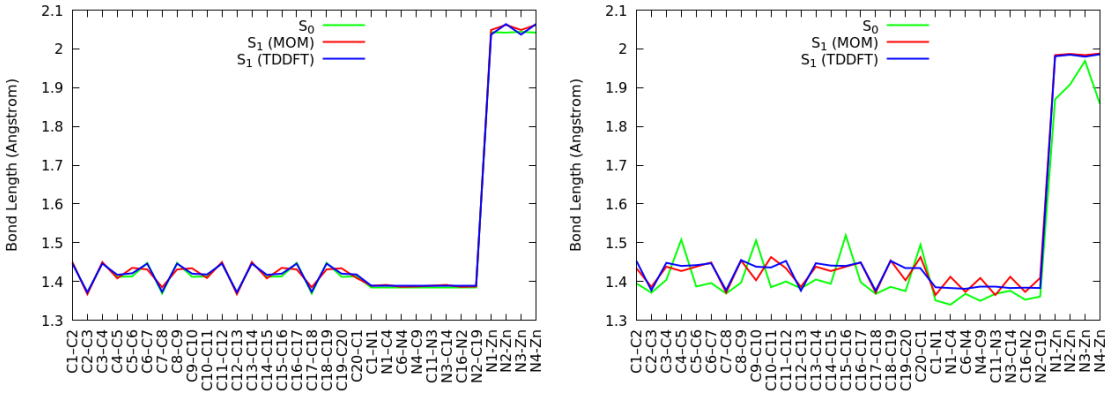


Figure 4.18: Optimised structural parameters of ZnTPP (left) and ZnT(t-Bu)P (right) of the S_0 and S_1 states for both the MOM and TDDFT methods. For atomic indices see figure 4.5.

such as ruffling and saddling of the macrocycle are observed and consistent with literature values and crystal structure data.

Ground state and excited state equilibrium geometries are reported for eight porphyrin systems that have been optimized in the gas phase. Although solvated systems have been studied the effects are negligible to the overall final geometry and energy of the excited state at this level of theory. The tetra phenyl porphyrin systems (ZnTPP and H_2TPP) displayed a planar configuration in both the S_0 and S_1 states with minor perturbations noticed in the geometries. ZnDPP, ZnOETPP

and their free base equivalents showed a saddled configuration again with minor differences between the bond lengths and dihedral angles in the S_0 and S_1. Finally the H_2T(t-Bu)P systems had a distinct ruffled feature in its porphyrin unit. Even though as shown in both figures 4.17 and 4.18 small pertubations exist between the ground and excited state the overall structure is still prominent seen in figure 4.2. As a result of this it is possible to say that the geometric configuration is dictated by the substituents attached to the core porphyrin unit and is not a result of photo-excitation.

At the calculated energy minimum, H_2T(*t*-Bu)P and its zinc based counterpart, displayed C_1 symmetry. H_2T(*t*-Bu)P is known to relax into this 'ruffled' configuration where the pyrrole groups on the porphyrin unit twist in order minimise any steric repulsion from the *tert*-butyl side chains. The same final ruffled configuration of ZnT(*t*-Bu)P was seen as its lowest energy configuration. It was thought that the metal porphyrin variant may exhibit different behaviour as a result of the zinc stabilising the structure so much that it may favour a saddled shape conformer. This feature was observed in the steps of the geometry optimisations. However, it did find its lowest energy configuration in this ruffled geometry which was confirmed by subsequent frequency calculations which are discussed further in Chapter 5.

Porphyrins are versatile molecules as they can be 'custom' tuned in accordance with the required function. Altering the porphyrin system so that it absorbs light at a different wavelength becomes extremely useful in applications such as PDT. The range of sampled porphryin based systems have been specifically selected in order to emphasize which feature has the most significant impact on the absorption of light. As can be seen in both tables 4.6 and 4.5 there are apparent shifts between the systems. Although there is an effect between the free base and the metal centred system which affects the E_{abs} and E_{ems} the influence of the substituent on the energies seems to be a much larger factor.

The deviation in the results presented in table 4.6 suggest that it is the contributions of the metal which are the larger factor whereas the opposite is true in when comparing results in table 4.5. The metal contributes to the bigger Stokes shift but the substituent effects are responsible for the bigger impact on the individual energies. This can be attributed to the electronic energy levels being disrupted as the molecular orbital network (seen in figure 4.16) is altered to a much greater degree with the changing substituents. Substituents on porphyrins have significant impact on the geometric features and cause structures to display features such as doming and ruffling. This disruption is also observed with the metal centre but it is seen consistently between all the metal based systems so any shift accredited to the change is seen throughout each system. When optimising the excited state this feature is emphasised more resulting in a larger shift in energies.

4.4 Conclusion

A series of free base and zinc based porphyrins have been studied with DFT. The range encompassed a selection of various geometric configurations in order to sample as many possible variations of the porphyrin structure. It was shown that by use of TDDFT porphyrin systems can be studied in detail and a rationale for what accounts for the behaviour could be deduced. Furthermore the excitation energies of both absorption and emission can be calculated with a strong correlation to experimentally derived results. The results demonstrate that less demanding methods (the "hybrid" approach) can also still be used whilst providing an appropriate level of accuracy.

The combined results demonstrate that non-planar saddled and ruffled configurations of porphyrins exhibit significant perturbations of the electronic spectra and the behaviour of the excited state. The perturbations of the excited state properties

are larger for the ruffled class, and the modified photo-physical properties occur whether the complex bears two central protons or a closed-shell metal ion. The strength of the shift increases between the free base and metal analogue which is indicative of a zinc to porphyrin charge transfer. This shift has been observed with analysis of the orbitals between the HOMO's and the LUMO's.

Overall the question remains of what is the significant factor that impacts the photo-activity of porphyrins. The results presented here suggest that the photo-physical properties are largely modulated by geometry which is dictated as a result of steric interactions that are governed by the substituents. The molecular orbitals show enough distortion to suggest the metal has an influence on the properties but ultimately geometry is the predominant factor.

The work presented here adds insight into what accounts for the photoactivity of porphyrins. However, fundamental questions still remain regarding the photophysical behaviour of the electronic and geometric perturbed porphyrins. A focus of the work in Chapter 3 is how the conformation of the two tryptophans with respect to the heme affects the rate of charge transfer. Here this study elaborates on this feature demonstrating how significantly porphyrin based systems can be modulated by conformation and electronic factors and the effect on the photophysical properties. The results clearly show that the properties of tetrapyrrole cofactors in biological systems may be strongly influenced by steric and electronic constraints imposed by a surrounding environment such as a protein. The results further enforce that once the porphyrin structures lose the planarity to become saddled/ruffled structures new opportunities for electronic pertubations such as photoexcitation become available. These controllable conformational changes may be exploited with application towards molecular electronic devices as well as towards protein engineering.

Chapter 5

Vibrationally resolved porphyrin spectra using the Franck-Condon and Herzberg-Teller approach

High resolution spectra of both free base and metallocentred complexes of porphyrin are calculated for the Q band region. Calculations on the vibronic structure of porphyrin are performed using DFT and TDDFT. Both Franck–Condon (FC) and Herzberg–Teller (HT) approaches have been used for estimation of frequency and intensity of vibronic bands in the simulated absorption spectra with respect to the $S_0 \rightarrow S_1$ electronic transition as the summation of contributions from both schemes characterise the electronic transitions and provide a high resolution description of the Q-band. This chapter shows that the first electronic transition into the singlet excited state is vibronically active in the Q_x region of the absorption spectrum, in good agreement with experimental data. The HT scheme appears to have more significant contributions and provides more insight into resolving the vibrationally active area of the absorption spectra. HT contributions to the electronic transition dipole moments are essential to assign the weak vibrational transitions and reproduce the experimental spectral profile. In order to provide a detailed account of the vibronic structure it is necessary to assign the vibrational transitions using both

HT and FC schemes. The importance of theoretical calculations are highlighted here and can help the general understanding of the photophysical properties of porphyrins.

5.1 Introduction

Porphyrins are of fundamental biological importance. The utility of these molecules has already been addressed and described in previous chapters. The fascinating properties are a direct result of the highly stable macrocyclic π system that affords it the optical stability as well as an efficient light absorber in both the visible and near-infrared region of the spectra. Significant effort has been devoted to characterising the electronic structures and spectra of these macrocyclic rings. By assessing vibronically active transitions it is possible to compute a highly resolved spectrum of porphyrin and related complexes. This information is not easily accessible as it requires careful interpretation and assignments of each spectral feature.[176]

UV spectra of metallophthalocyanines and metallotetrabenzoporphyrins were first given spectral assignments by Edwards, Gouterman, and co-workers[1]. These earlier band assignments were based on the Gouterman four-orbital model[6]. There have been ZnP studies proposing a symmetry forbidden n-π^* transition blue-shifted from the highest energy states of the Q band region.[177–180] Even though the additional assignments in the first four bands have been debated in literature, Gouterman and Edwards made a key observation in 1970. They noted that the Q band, in particular, would be subject to Franck-Condon (FC)[181,182] vibrational displacement, and point to an earlier publication by Gouterman and Fulton, in which the higher energy bands were believed to take on much more peculiar shapes, especially at the investigated temperature of 0.5 K.[183] The four orbital model is simplistic, but the framework is generally still reliable for the major transitions, and is used in the present work.

The study of molecular electronic spectra must consider vibronic effects. These are essential as the shape of both absorption and emission bands are a result of vibronic transitions. There is an abundance of experimental data available allowing the theoretical assignment of vibrational bands in highly resolved spectra to be performed with confidence[8]. Frequencies and structural features of porphyrins are difficult to study due to the short lifetimes of the excited states. By use of quantum mechanical methods it is possible to carry out accurate calculations of these bands and provide a comprehensive study of the vibronic transitions.

The Franck–Condon principle is a rule that explains the intensity of vibronic transitions.[181,182] These transitions are the simultaneous changes in electronic and vibrational energy levels of a molecule due to the absorption or emission of a photon of the appropriate energy. The principle states that during an electronic transition, a change from one vibrational energy level to another will be more likely to happen if the overlap between the two vibrational wave functions is significant.

Gouterman's four orbital model already describes that the two distinctive Q and B (soret) bands are derived from orbital transitions between the pair of HOMO's and the two LUMO's.[155] The strong configuration interaction between the four orbitals (HOMO-1, HOMO, LUMO, LUMO+1) accounts for the low lying states associated with the Q-band and the high lying state corresponding to the B band. The transition dipole moments associated with the configuration mixing result in an anti-parallel and a parallel electron transition which lead to a weak absorption (Q band) and a strong absorption (B band).

The transitions in the Q band are strictly forbidden. However, they are allowed as a result of vibronic effects. The FC approximation breaks down for these type of weakly allowed transitions. Hence, the requirement to introduce the HT effect.

Here we observe the phenomenon of intensity 'borrowing' which was described by Herzberg and Teller as a way to understand certain unexpected vibronic transitions.[184,185] It was pointed out that this could be understood if the dipole transition moment for the electronic transition depended on the nuclear co-ordinate associated with the vibration. In the Franck-Condon approximation one assumes that there is no such dependence.

This work describes an effective strategy for simulation and analysis of the vibronic absorption spectra for a series of free base and metallo centered porphyrin complexes. The HT scheme and FC approximation are employed in order to facilitate this. This allows for the nuclear vibrations (within the harmonic approximation)of the Q_x band to be solved at high resolution.

Jarzqcki *et al*[186] simulated the highly resolved infrared and Raman spectra of ZnP and demonstrated that the a_{2u} and e_u vibrations are IR active. Their results showed a strong a_{2u} C-H out of plane wagging around 850 cm^{-1} was accompanied by a weaker component in the IR spectrum. TDDFT was used to calculate the absorption spectra of metallo-porphyrins and their derivatives[187–192]. To simulate the spectral profile in theory a method known as the linear coupling model (LCM)[193] was used to compute optical spectra. The LCM approach postulates that, except for a displacement, the normal ground-state modes are identical to the excited-state modes, although this approach neglects the possible change of frequencies and the Duschinsky rotation.[194–196] Because this model undoubtedly oversimplifies the vibronic interactions for sizeable molecules, it cannot provide accurate quantitative descriptions of their electronic spectra.[197] A more comprehensive strategy is required to characterize the spectral features and provide a direct link between spectra and structural parameters. This chapter describes a strategy for simulating the absorption spectrum of free base and zinc centred porphyrin in the Q-band region using the Franck–Condon (FC) approximation with the Herzberg–Teller (HT) and Duschin-

sky contributions.

Obtaining a reliable description of the high resolution vibronic absorption spectrum of porphyrin requires the calculation of Franck–Condon factors for various peaks (in the vibronic bands), oscillator strengths and excitation energies. Many theoretical studies have reported the oscillator strengths and vertical excitation energies of low-lying excited states of ZnP.[198–200] A study by Ziegler *et al*[198] provides a comprehensive review of the calculated spectral properties of the ZnP system. In the present work, these parameters were re-calculated using the TDDFT method, and the validity of this approach was confirmed by comparing the results with available experimental and theoretical data.

The transition dipole moments of strong dipole-allowed transitions can be considered independent of the nuclear coordinates (Condon approximation). This approximation is usually satisfactory for transitions characterized by large oscillator strengths, giving rise to the so-called FC spectrum. However, the $S_0 \rightarrow S_1$ electronic transition of porphyrin is weak, and the FC approximation is often inadequate for weak dipole-allowed and dipole forbidden transitions. To fully characterize the vibronic fine band of the $S_0 \rightarrow S_1$ absorption spectrum for free base and zinc porphyrin, it is necessary to consider the HT and Duschinsky effects.

The equilibrium geometries and vibrational frequencies of the ground and first excited singlet states have been calculated using TDDFT. The $S_0 \rightarrow S_1$ absorption spectrum of H_2P and ZnP has been simulated using the FC approximation both with and without the HT contributions. The simulated spectra were compared with experimental observations and tentatively assigned fundamental vibrational modes. The resulting theoretical insights should be very useful in identifying the true vibrational transitions in the Q-band regions of large molecules.

5.2 Methodology

Optimized geometries of both the ground and excited states were obtained using the same methods as in the previous chapter. All vibrational wave numbers were found to be real indicating the location of a minimum on the hypersurface of the total energy for all systems. Vibronic absorption spectra taking into account the first order Franck–Condon and the Herzberg–Teller effects were simulated in the vacuum approximation using the optimized geometries, the vibrational frequencies at the S_0 and S_1 states, adiabatic electronic transition energy and transition dipole moment for the $S_0 \rightarrow S_1$ electronic transition for all compounds. All vibronic absorption spectra were simulated with Gaussian broadening. FC and HT approximations were calculated using FCClasses code[8] which implements a time-independent method for computing spectra between electronic states with negligible non adiabatic couplings. It is suitable for semi rigid molecules for which an harmonic approximation is reliable.

The vibrationally resolved spectra of porphyrin complexes including the HT contributions have been calculated. The spectra were simulated with a full width at half maximum (FWHM) of 0.0005 eV. The maximum number of integrals to be computed for each class was set to 10^6. The simulated spectrum taking into account both the zero-order (FC) and the first-order (HT) terms is referred to as the FCHT. In order to investigate the HT contributions separately, the Figures of pure HT contributions and the summation of HT and FC contributions were calculated and plotted.

5.3 Results and Discussion

Porphyrins are very important chromophores, involved in many biological processes such as photosynthesis. Free-base porphyrin (H_2P) is the major building block of chlorophyll, and it is therefore very interesting to study its absorption properties

of UV and visible light. These are nicely interpreted according to the four-orbital model of Gouterman which has been described in previous chapters.[8]

The bands and transitions attributed to the porphyrin complexes are largely modulated by the symmetry and point group of the molecule. As seen in the uv-vis spectrum metal porphyrins exhibit two bands; Q and B (Soret) band. For an idealised D_{4h} symmetry displayed by ZnP the two bands are actually composed of multiple degenerate polarised transitions along the x or y axis (Q_x, Q_y, B_x and B_y). As a result of the two hydrogens in the center of free base porphyrin belongs to the point group D_{2h}. The symmetry of the molecule is broken and the x and y polarised transitions are no longer degenerate hence four peaks are observed in the Q band. Even though symmetry has been broken, the Q bands are still very weak and a calculation of their spectra must take into account the HT effect. The following results present calculations for the absorption spectra of the Q_x band for both free base and zinc porphyrin.

For the FC spectrum to be an adequate approximation for simulating the absorption spectra the differences between vibrational frequencies of the S_0 and S_1 electronic states must be small. This is seen in tables 5.1 and 5.2 and shown in figures 5.1 and 5.5 where the frequency difference ($\Delta\omega$) is only relatively small ranging between 8.8 to 27 cm^{-1}.

Figure 5.2 shows the FC, HT and a combined FCHT spectrum in order to compare directly between intensities of FC and HT. From this it is clear of the significant contributions from the HT terms on the shape of the overall spectra. The FC spectra are very narrow and dominated by S_0 - S_0 transitions. This is dependent on both geometry and flexibility of the molecule. As the molecules in both ground and excited optimised states are rigid the displacements are only small. The other peaks present are due to transisitions to states with excited state vibrations and are

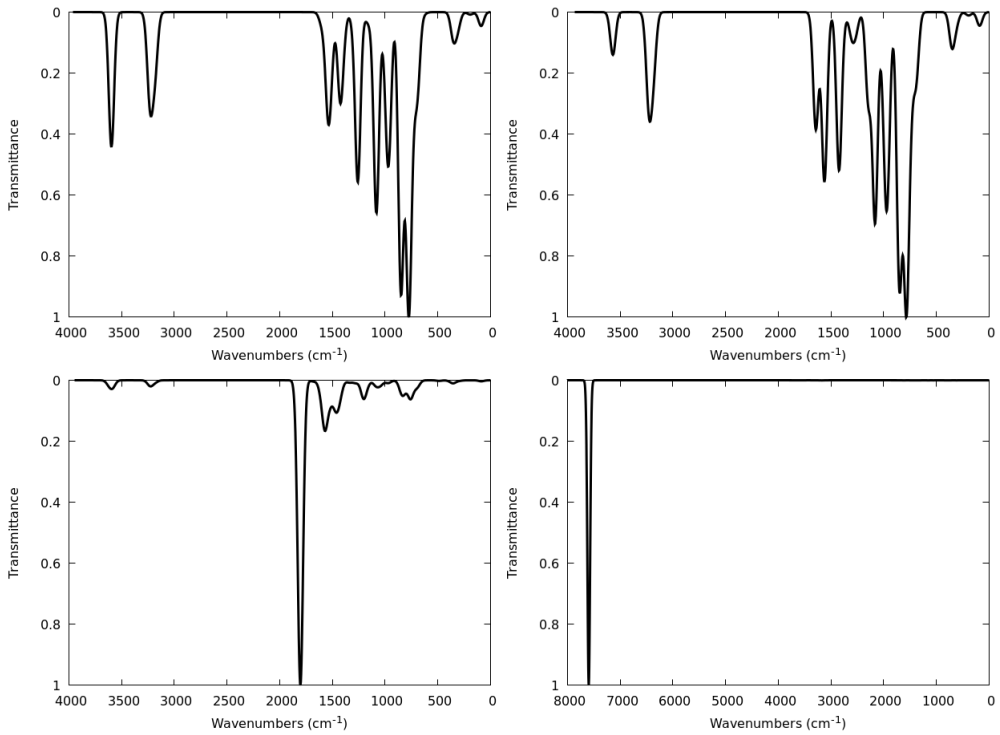


Figure 5.1: IR spectra for the first four excited state geometries responsible for the four peaks in the Q band for free base porphyrin. Top Left is excited state 1, Top Right is excited state 2, Bottom Left is excited state 3 and Bottom Right is excited state 4.

described in table 5.1.

The normal vibronic modes are a_{1g} or b_{1g} symmetry. The a_{1g} mode give rise to peaks which have both FC and HT contributions which are polarized along the x direction, while b_{1g} mode progressions result only from HT contributions polarized along y. Symmetry considerations show that HT terms polarized along x mainly reflect intensity borrowing from the B_x Soret band, while y-polarized HT terms derive from mixing with the B_y state whilst minor contributions may also be derived from the Q_y state. This feature has been observed experimentally[8,201] and is also seen in the results presented here.

The FC spectrum is shown Figure 5.2. The most significant peak corresponds to the 0-0 transition at 2.27eV whilst the second highest peak occurs at 2.31eV an

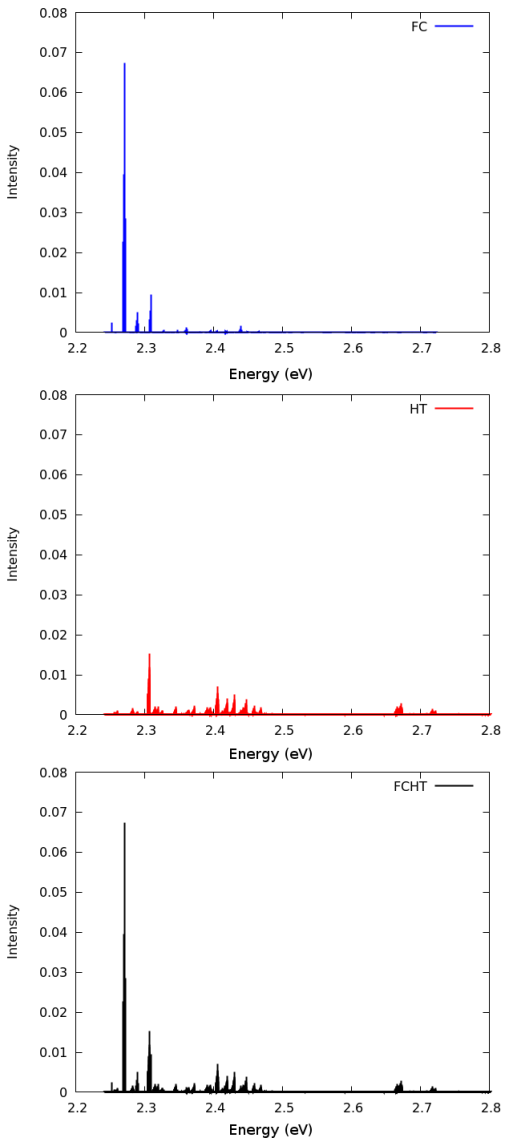


Figure 5.2: Computed stick spectra (298K) of H_2P using only FC contributions (top), HT mechanism (middle) and a combined model including both FC and HT contributions, FCHT (bottom).

is the result of the excitation of one quanta to the 13th vibrational mode. The HT spectrum calculates more vibrational frequencies over a broader range in comparison to the FC spectra however the intensity of these peaks are significantly reduced. This is due to the fact that the HT contributions are much stronger than the contributions from FC. By combining both results it is possible to create the FCHT spectra which gives a rich and insightful view towards the peaks in Q-band (figures 5.3 and 5.4) and furthers the understanding of the vibronic transitions associated with the $S_0 \rightarrow S_1$ transition. These frequency modes can be visualised and assign-

Mode (n)	State S_1 ω (cm^{-1})	δ	State S_0 m'$_1$[J]	m'$_2$[J]	m'$_3$[J]	$\Delta\ \omega$ (cm^{-1})
7	151.89	-0.347	7'[1.00]			-4.11
13	310.08	0.554	13'[1.00]			-2.23
32	729.79	0.094	32'[0.86]	33'[0.14]		-10.61
33	735.06	-0.170	33'[0.86]	32'[0.14]		-9.42
41	793.99	0.000	41'[1.00]			-4.25
53	977.93	0.427	54'[0.99]	63'[0.00]	32'[0.00]	-13.17
61	1085.68	-0.093	63'[0.99]	54'[0.00]	68'[0.00]	-5.28
68	1213.84	-0.101	68'[0.98]	78'[0.01]	84'[0.00]	-12.86
70	1266.58	0.000	71'[1.00]			-4.86
75	1369.5	0.000	75'[0.92]	76'[0.06]	79'[0.02]	-15.48
76	1387.82	0.200	78'[0.85]	80'[0.07]	84'[0.04]	-30.97
78	1420.59	0.000	79'[0.97]	75'[0.01]	76'[0.02]	-32.27
90	1606.43	-0.068	91'[0.95]	94'[0.05]	84'[0.00]	-25.88
91	1637.01	0.000	92'[0.99]	76'[0.00]	86'[0.00]	-35.82
94	1671.8	-0.038	94'[0.94]	91'[0.05]	78'[0.00]	-15.75

Table 5.1: Relevant normal modes (n) of the two electronic states for H_2P. δ is the dimensionless displacement. The three normal modes m'$_1$, m'$_2$ and m'$_3$ of S_0 (identified by a prime) on which the S_1 mode n is mainly projected, together with their weights $J(m'_x, n)^2$, where J is the Duschinsky matrix as demonstrated by Santoro et al[8]. Values = 0.000 refer to values <0.001. The frequency difference between n and the normal mode m'$_1$ ($\Delta\omega = \omega - \omega_{m'1}$).

ments can be made.

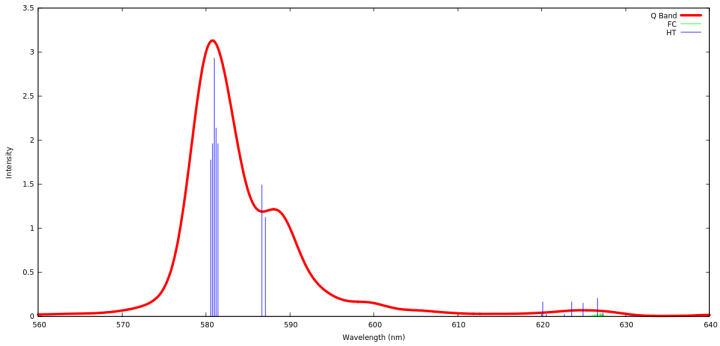


Figure 5.3: Q band of free base porphyrin with scaled HT and FC factors.

The calculated results for ZnP show that the FC mechanism, like H_2P, is also very weak. The simulated FC absorption spectral profile of the Q band is very narrow and is dominated by the 0–0 transition. This result seems to indicate that there are only small displacements between the equilibrium structures of the two

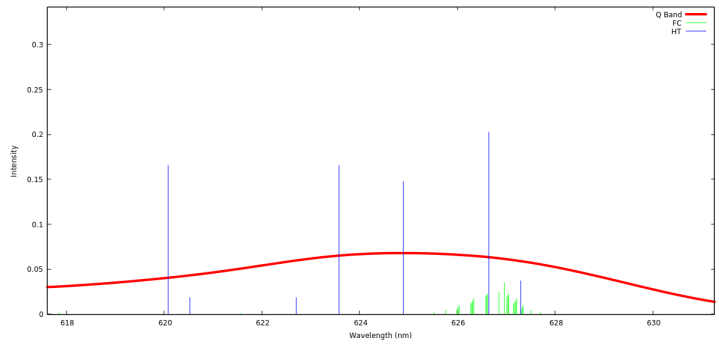


Figure 5.4: Q band of free base porphyrin with scaled HT and FC factors. This figure focuses on the 620-630 nm region.

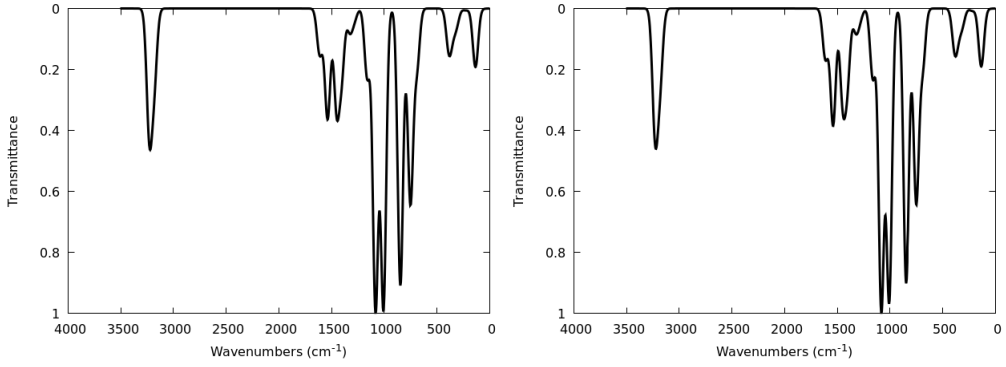


Figure 5.5: IR spectra for the first two excited state geometries responsible for the two peaks in the Q band for zinc centred porphyrin

electronic states. For all the normal mode frequencies reported in table 5.2 only the totally symmetric modes: 7, 18, 35, 55, 77, 83 and 86 have calculated FC factors that were greater than zero indicating that only symmetric normal modes contribute to the spectra obtained via the FC approximation. For the S_0 state with D_{4h} point group the dominate symmetry groups are a_{1g} and b_{1g} whilst the excited state has a reduced symmetry point group, D_{2h}, leading to the dominate symmetry group to be a_g.

A benefit of computational results is the ability to visualise results. Vibronic analysis reveals that mode 7 corresponds to the asymmetric N–Zn–N stretching vibration in which one N–Zn–N bond is elongated while the other is simultaneously shortened. This distortion in the centre leads to the reduction in symmetry from

	State S_1		State S_0			
Mode (n)	ω **(cm^{-1})**	δ	**m'$_1$[J]**	**m'$_2$[J]**	**m'$_3$[J]**	$\Delta\ \omega$ **(cm^{-1})**
7	173.9	0.346	6'[1.00]			-8.8
8	195.2	0.000	7'[0.97]	2'[0.02]		-3.0
16	357.5	0.000	17'[0.88]	16'[0.12]	24'[0.00]	-2.6
18	368.4	0.460	18'[1.00]			-1.0
35	729.2	0.117	34'[0.98]	35'[0.02]		-7.8
37	750.7	0.000	36'[0.99]			-4.2
39	776.5	0.000	38'[0.95]	33'[0.04]		-6.8
49	854.4	0.000	49'[0.99]	38'[0.01]		-12.7
55	1002.1	0.084	56'[0.71]	57'[0.27]	59'[0.01]	-15.8
56	1008.1	0.000	55'[0.91]	60'[0.07]	60'[0.01]	-4.1
57	1010.5	0.000	54'[0.85]	59'[0.12]	55'[0.01]	-1.7
63	1084.2	0.000	62'[0.96]	63'[0.02]	59'[0.01]	-3.2
64	1085.7	0.000	63'[0.94]	62'[0.02]		-4.7
67	1152.2	0.000	66'[0.97]	75'[0.02]		-54.9
69	1159.9	0.000	67'[0.53]	68'[0.46]		-16.1
70	1195.5	0.009	69'[0.99]			-3.1
72	1284.4	0.000	71'[0.91]	74'[0.05]		-18.2
77	1368.6	0.156	78'[0.95]	86'[0.02]	79'[0.01]	-14.2
80	1398.9	0.000	81'[0.70]	80'[0.26]	74'[0.01]	-16.6
83	1442.4	0.004	84'[0.99]			-21.5
84	1447.6	0.000	83'[0.76]	82'[0.24]		-15.6
86	1502.0	0.018	86'[0.91]	91'[0.04]	93'[0.03]	-27.2
87	1534.8	0.000	88'[0.68]	89'[0.27]	90'[0.02]	-15.6

Table 5.2: Relevant normal modes of the two electronic states for ZnP. δ is the dimensionless displacement. The three normal modes m'$_1$, m'$_2$ and m'$_3$ of S_0 (identified by a prime) on which the S_1 mode n is mainly projected, together with their weights $J(m'_x, n)^2$, where J is the Duschinsky matrix as demonstrated by Santoro et al[8]. Values = 0.000 refer to values <0.001. The frequency difference between n and the normal mode m'$_1$ ($\Delta\omega = \omega - \omega_{m'1}$).

D_{4h} to D_{2h} as the molecule transitions from the ground to the excited state. This suggests that this mode is the main contributor to the $S_0 \rightarrow S_1$ absorption spectrum of ZnP within the FC approximation.

The FCHT scheme is a much richer representation and matches experiment closer. The simulation gives a greater insight into the vibrational components of each vibronic transition. In order to assign vibrational modes this work compares experimental[201] and simulated spectra. Some of the calculated frequencies are slightly overestimated. An explanation for this is the difference of the optimised and ground

state geometries in the absence of a solvent compared to experimental values. It is notable that most of the main transitions in the FC and FCHT spectra were assigned as the fundamental 0–1 vibration of a single mode. This is reasonable because the displacements and frequency shifts between the simulated ground and excited states are very small.

So far all discussions have focused on the FCHT spectra at 0K. For steady state finite temperature spectroscopy a Boltzmann distribution of vibrational states of the initial electronic state must be accounted for. This can have significant impact on the final spectrum but dramatically increases the computational demand.[8]

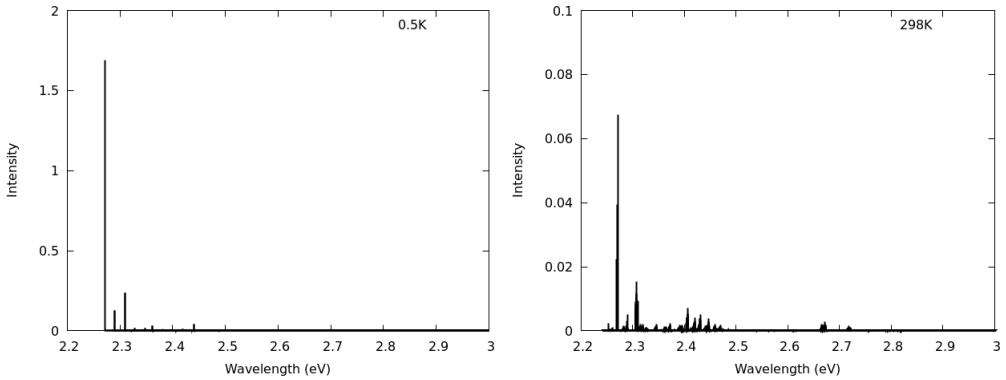


Figure 5.6: Comparisons of the FCHT absoprtion stick spectra calculated at 0.5K (left) and 298K (right).

The FCHT spectra for both ZnP and H_2P have been calculated at 0.5K and 298K shown in figure 5.6. The different temperature spectra introduce a broadening of the spectrum which otherwise could be simulated with a broader Gaussian convolution. The high-resolution stick spectra are on the contrary very different with the one at finite temperature being much more dense, as expected. These differences are diminished by convolution.

5.4 Conclusion

In this chapter the high resolution electronic absorption spectra of free base porphyrin has been studied using DFT based methods to determine the harmonic adiabatic vibronic structure of the first electronic transistion. The effects have been studied using the Herzberg-Teller scheme and Franck-Condon approach. The role of geometric configuration, electronic structure and presence of a metal have been studied vibronically. All significant and relevant vibronically-active vibrational modes have been identified and their corresponding intensities have been estimated. The theoretically derived results accurately describe the intensities of the vibrationally active peaks when directly compared to experimental spectra.

Previously discussed is the idea that Porphyrins can be treated as a large aromatic molecule with it being likened to quintessential aromatic benzene. Several studies have shown large aromatic molecules to be successfully resolved within the FC approximation leading to the assumption that the vibronic structure may be calculated within the same framework. From the results shown this is true to an extent however the vibronically induced transitions from the Herzberg-Teller mechanism are found to be overall more significant contribution than the FC induced lines. Not disregarding any results a combination of both FC and HT schemes helps to determine the vibrational spectra which is not modulated by the presence of a metal but is influenced greater by symmetry and temperature. The FCHT mechanism produces a richer vibronic structure than solely the FC approximation. The study reproduces detailed features of the highly resolved S_0 to S_1 absorption spectrum and is therefore expected to increase the understanding of the nature of electronic transitions in the Q-band for both porphyrin centred systems and large aromatic molecules.

Chapter 6

Conclusions

The work carried out in this thesis has shown a variety of techniques and methodologies related to quantum chemical calculations which focus on predicting excited states of porphyrins. This has enabled the successful calculation of rates of electron transfer in proteins, Chapter 3. It has provided useful insights into understanding how the absorption and emission spectra are affected by conformational changes of the porphyrin molecule, Chapter 4. Finally, the understanding of the nature of electronic transitions in the Q-band for both porphyrin centred systems and large aromatic molecules has been investigated and steps taken to share new insights on the highly resolved absorption spectra, Chapter 5.

Looking to the future, this work has laid out a reliable way to calculate absorption and emission of large molecules using feasible computational methods. It has also shown that electron and energy transfer can be calculated using simpler methods. This allows predictions to be made about the potential system such as *'will it be faster or slower?'* or *'will it be better or worse'*. The facile answers to these questions are vital and are attracting significant attention in the field of protein engineering where the study of radical catalysis is growing due to these newer ways to investigate.

6.1 Radical approaches towards protein engineering

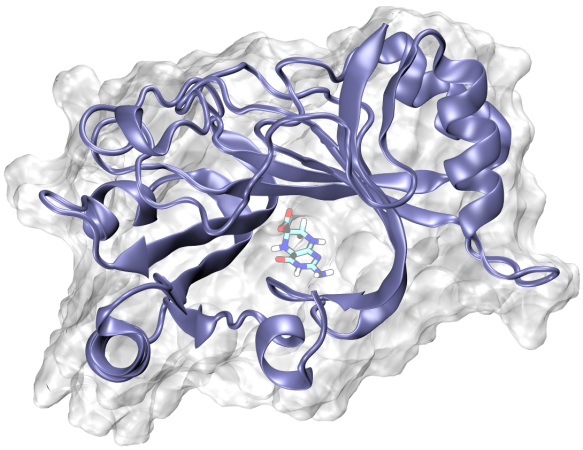


Figure 6.1: QueE protein with the radical SAM enzyme nestled in the active site.

Radical S-adenosylmethionine (SAM) dependent enzymes[202] are a relatively recently discovered enzyme family with great prospects for biotechnological applications. Their key catalytic communality is to deal with highly reactive radical intermediates during catalysis. The way the enzymes control these intermediates, which are particularly hard to control in classical synthesis, makes them essential in a broad set of challenging biosynthetic pathways that often lead to products of potential use in anti-viral, anti-cancer and antibiotic treatments.

A thorough knowledge of the reaction mechanisms involved in the biocatalysis of these enzymes presents the foundation for rational enzyme engineering and one of the bottlenecks for a more rapid access to rational enzyme design simultaneously. In the best case scenario everything from the catalytic reaction mechanism, the kinetics, to substrate binding, and dynamic effects is known in detail. Still, individual steps of the catalysis can be addressed by quicker methods, in order to get a first qualitative picture of how these steps can be influenced and can feed into the enzyme design process. This work focuses on potential transferable approaches to reduce complexity and computational effort and to focus on key features related to rSAM

enzyme catalysis for their use in enzyme design.

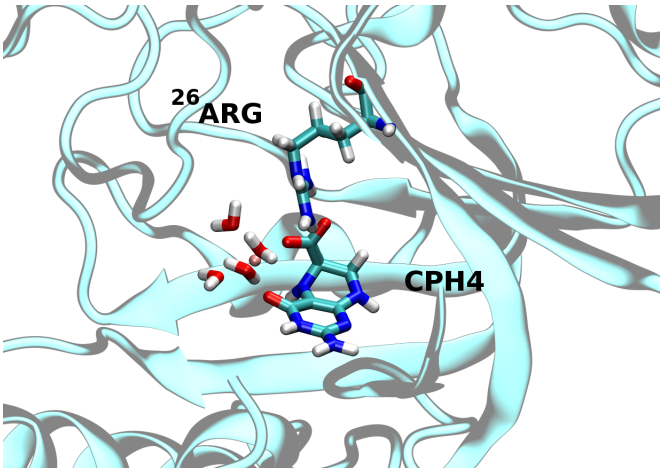


Figure 6.2: Radical SAM enzyme (CPH4) complexed with an Mg ion, arginine and surrounding waters.

Radical stabilization energies (RSEs)[203] are an example that offers an attractive possibility to assess the overall thermodynamics of radical reactions and rearrangements as central steps in radical SAM enzyme catalysis. Through the example of the bacterial 7-carboxy-7-deazaguanine (CDG) synthase (QueE, figure 6.1),[204] we will highlight key features and details of the biocatalytic radical rearrangement mechanism involved[205] and will discuss the potential of using radical stabilization energies for thermodynamic reaction profiling directly from enzyme substrate complexes (figure 6.3). Further, we will provide insights into other challenges of radical SAM enzymes addressed in the context of enzyme engineering.

6.2 An outlook to the future of simulations

The future of running simulations is changing due to the shift in infrastructure made feasible by Cloud computing. Already we are seeing conventional High Performance Computers (HPC) adapting to the changing computing climate by adopting cloud bursting abilities which is perhaps a small step towards fully adopting a purely Cloud HPC. In the US, high performance cloud computing facilities are now part

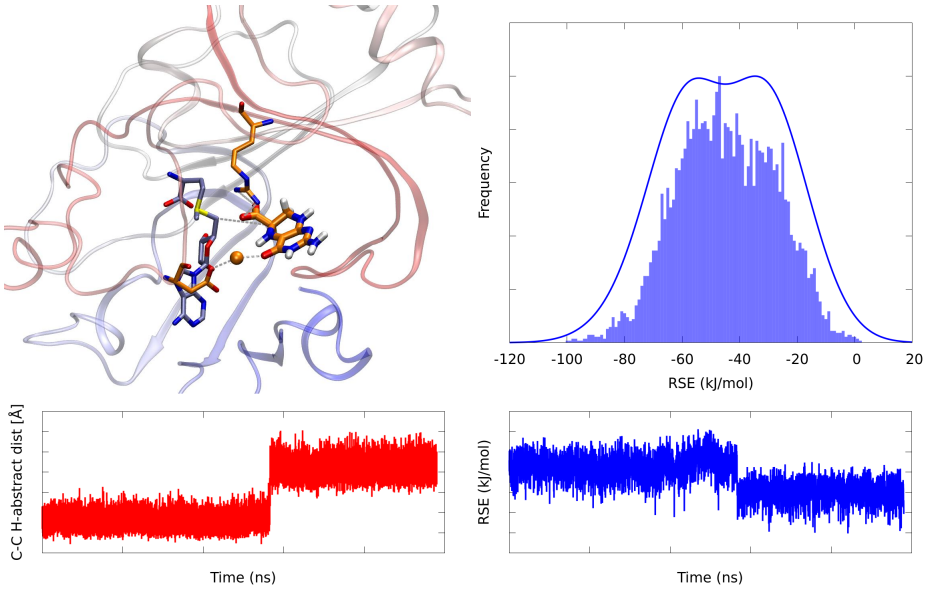


Figure 6.3: An overview of how RSE's are computed. Absolute energies of both substrate and its radical are calculated on MD snapshots. In the above figure the conformation of the substrates changes to a more stable configuration after around 60ns.

of the national academic sector provision and large pharmaceutical companies are beginning to embrace the cloud due to its elasticity. One of the main reason for lack of widespread adoption is that cloud based solutions require significant adjustments to established methods of working. Resource access, job submission, monitoring and data manipulation require rethinking. As the infrastructure changes so do the possibilities arise to make simulations 'better' and work is currently ongoing in order to fully take exploit all these opportunities.

The key feature of cloud computing is the flexibility it allows. For scientific research the solution of cloud computing is now more attractive than ever due to recent developments in computational power (CPU, GPU and FPGA) and competitive pricing between the key players in the industry. Amazon Web Services (AWS) have a market known as 'Spot Market' where unused computing resources are sold at a significantly reduced price. The cost-effectivness of cloud-based solutions for molecular dynamic (MD) simulations on B-cell lymphoma-extra large (BCL-XL) has been benchmarked in order to substantiate claims that cloud solutions can be

cost effective.

	ARCHER			EC2	
	24 Cores	**48 Cores**	**96 Cores**	**On Demand**	**Spot**
GROMACS	0.07 (71)	0.11 (84)	-	0.30 (35)	0.04 (35)
AMBER	-	0.33 (28)	0.41 (28)	0.20 (51)	0.02 (51)

Table 6.1: Price/Performance for MD using GROMACS and AMBER for BCL-XL. Costs are in £/ns (smaller is better); figures in brackets are ns/day (larger is better)

Table 6.1 reports the cost for both 'On Demand' and 'Spot' access available through AWS. The former is the cost of the computing resource that is available instantly and the latter is the cost for an unused resource being sold at reduced price. For the systems on ARCHER (national supercomputing resource) the two codes do not scale the same so a range of cores have been used. It is found that for AMBER the cloud option is always cheaper and always faster. This reflects the particularly efficient GPU implementation in the code. For GROMACS, ARCHER is always faster but the cloud option is considerably cheaper if a 'spot' resource is used.

Everyone uses a form of HPC on a daily basis with the majority of users being unaware of this interaction. The idea that HPC should be invisible to the user is becoming a more and more relevant particularly towards simulations run in a cloud environment. This form of invisible cloud based HPC will hopefully lower the adoption barrier for traditional bench based scientists who do not try theoretical techniques due to the steep learning curve towards running simulations. Of course this is a very 'black' box approach but for those who simply want to use computational codes it becomes very appetising.

Although the *status quo* of running computational tasks is well optimised it seems bizarre that a new user must learn outdated techniques. As the infrastructure changes it is important that methodologies adapt to keep up. Computational chemistry is evolving and it is exciting.

Bibliography

[1] Edwards, L.; Dolphin, D.; Gouterman, M.; Adler, A. D. *Mol. Spectrosc.* **1971**, *32*, 16–32.

[2] Harriman, A. *J. Chem. Soc., Faraday Trans. 1* **1980**, *76*, 1978–1985.

[3] Peeks, M. D.; Claridge, T. D. W.; Anderson, H. L. *Nature* **2007**, *541*, 200–203.

[4] Suess, C. J.; Hirst, J. D.; Besley, N. A. *J. Comput. Chem.* **2017**, *38*, 1495–1502.

[5] Rizzo, T. R.; Park, Y. D.; Peteanu, L. A.; Levy, D. H. *J. Chem. Phys.* **1986**, *84*, 2534–2541.

[6] Gouterman, M.; Wagnire, G. H.; Snyder, L. C. *J. Mol. Spectrosc.* **1963**, *11*, 108–127.

[7] Gentemann, S.; Medforth, C. J.; Forsyth, T. P.; Nurco, D. J.; Smith, K. M.; Fajer, J.; Holten, D. *J. Am. Chem. Soc.* **1994**, *116*, 7363–7368.

[8] Santoro, F.; Lami, A.; Improta, R.; Bloino, J.; Barone, V. *The J. Chem. Phys.* **2008**, *128*, 224311.

[9] Yao, P.; Han, S.; Zhang, Y.; Zhang, X.; Jiang, J. *Vib. Spectrosc.* **2009**, *50*, 169–177.

[10] Hutchison, J. A.; Sintic, P. J.; Brotherhood, P. R.; Scholes, C.; Blake, I. M.; Ghiggino, K. P.; Crossley, M. J. *J. Phys. Chem. C* **2009**, *113*, 11796–804.

[11] Pileni, M.-p.; Gratzel, M. *J. Chem. Phys.* **1980**, *2*, 1822–1825.

[12] Hückel, E. *Zeitschrift für Physik* **1931**, *70*, 204–286.

[13] Feixas, F.; Matito, E.; Solà, M.; Poater, J. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7126–7137.

[14] Hsia, C. C. *N. Engl. J. Med.* **1998**, *338*, 239–247.

[15] Giardina, B.; Messana, I.; Scatena, R.; Castagnola, M. *Crit. Rev. Biochem. Mol. Biol.* **1995**, *30*, 165–196.

[16] Moore, E.; Mander, A.; Ames, D.; Carne, R.; Sanders, K.; Watters, D. *Int. Psychogeriatrics* **2012**, *24*, 541–556.

[17] Son, H.-J.; Jin, S.; Patwardhan, S.; Wezenberg, S. J.; Jeong, N. C.; So, M.; Wilmer, C. E.; Sarjeant, A. a.; Schatz, G. C.; Snurr, R. Q.; Farha, O. K.; Wiederrecht, G. P.; Hupp, J. T. *J. Am. Chem. Soc.* **2013**, *135*, 862–9.

[18] Suppan, P. *Chemistry and light*; The Royal Society of Chemistry: Cambridge, 1994.

[19] Hartensteiner, S.; Kruetler, B. *Biochim. Biophys. Acta - Bioenerg.* **2011**, *1807*, 977–988.

[20] Santhanamoorthi, N.; Lo, C.-M.; Jiang, J.-C. *J. Phys. Chem. Lett.* **2013**, *4*, 524–530.

[21] Li, L.-L.; Diau, E. W.-G. *Chem. Soc. Rev.* **2013**,

[22] Rintoul, L.; Harper, S. R.; Arnold, D. P. *Phys. Chem. Chem. Phys.* **2013**, *15*, 18951–64.

[23] Pushpan, S. K.; Venkatraman, S.; Anand, V. G.; Sankar, J.; Parmeswaran, D.; Ganesan, S.; Chandrashekar, T. K. *Curr. Med. Chem. Anticancer. Agents* **2002**, *2*, 187–207.

[24] Wiehe, A.; Shaker, Y. M.; Brandt, J. C.; Mebs, S.; Senge, M. O. *Tetrahedron* **2005**, *61*, 5535–5564.

[25] Pandey, R. K.; Zheng, G. *Porphyrin Handbook* **2000**, *6*, 157–230.

[26] Silva, E. F. F.; Serpa, C.; Dabrowski, J. M.; Monteiro, C. J. P.; Formosinho, S. J.; Stochel, G.; Urbanska, K.; Simões, S.; Pereira, M. M.; Arnaut, L. G. *Chemistry* **2010**, *16*, 9273–9286.

[27] Govind, N.; Valiev, M.; Jensen, L.; Kowalski, K. *J. Phys. Chem. A* **2009**, *113*, 6041–6043.

[28] Słota, R.; Broda, M. a.; Dyrda, G.; Ejsmont, K.; Mele, G. *Molecules* **2011**, *16*, 9957–71.

[29] Zheng, W.; Shan, N.; Yu, L.; Wang, X. *Dye. Pigment.* **2008**, *77*, 153–157.

[30] Segev, M.; Silberberg, Y.; Christodoulides, D. N. *Nat. Photonics* **2013**, *7*, 197–204.

[31] Liang, X.; Li, X.; Yue, X.; Dai, Z. *Angew. Chemie - Int. Ed.* **2011**, *50*, 11622–11627.

[32] Kasha, M.; Rawls, H. R.; Ashraf El-Bayoumi, M. *Pure Appl. Chem.* **1965**, *11*, 371–392.

[33] Anderson, H. L. *Inorg. Chem.* **1994**, *33*, 972–981.

[34] Esdaile, L. J.; Jensen, P.; McMurtrie, J. C.; Arnold, D. P. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 2090–3.

[35] Sedghi, G.; García-Suárez, V. M.; Esdaile, L. J.; Anderson, H. L.; Lambert, C. J.; Martín, S.; Bethell, D.; Higgins, S. J.; Elliott, M.; Bennett, N.; Macdonald, J. E.; Nichols, R. J. *Nat. Nanotechnol.* **2011**, *6*, 517–523.

[36] Kocherzhenko, A. A.; Patwardhan, S.; Grozema, F. C.; Anderson, H. L.; Siebbeles, L. D. A. *J. Am. Chem. Soc.* **2009**, *131*, 5522–5529.

[37] Grozema, F. C.; Siebbeles, L. D. *Int. Rev. Phys. Chem.* **2008**, *27*, 87–138.

[38] Sprafke, J. K.; Stranks, S. D.; Warner, J. H.; Nicholas, R. J.; Anderson, H. L. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 2313–2316.

[39] Sessler, J. L.; Cho, D.-G.; Stepień, M.; Lynch, V.; Waluk, J.; Yoon, Z. S.; Kim, D. *J. Am. Chem. Soc.* **2006**, *128*, 12640–12641.

[40] Sessler, J. L.; Seidel, D. *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 5134–5175.

[41] Born, M.; Oppenheimer, R. *Annalen der Physik* **1927**, *389*, 457–484.

[42] Hartree, D. R. *Ann. der Phys* **1928**, *24*, 89–110.

[43] Fock, V. *Zeitschrift fur Physik* **1930**, *61*, 126–148.

[44] Slater, J. C. *Phys. Rev.* **1951**, *81*, 385–390.

[45] Roothaan, C. *Rev. Mod. Phys.* **1951**, *23*, 69–89.

[46] Hall, G. G. *Proc. R. Soc. Lond. A* **1951**, *205*, 541–552.

[47] Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939–947.

[48] Dunning Jr, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.

[49] Martin, J. M. L.; Parthiban, S. *Quantum Mechanical Prediction of Thermochemical Data* **2001**, *22*, 31–65.

[50] Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, 864–871.

[51] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, 1133–1138.

[52] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[53] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

[54] Lee, C.; Yang, W.; Parr, R. *Phys Rev B* **1988**, *37*, 785–789.

[55] Becke, A. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

[56] Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.

[57] Lange, A.; Herbert, J. M. *J. Chem. Theory Comput* **2007**, *3*, 1680–1690.

[58] Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51–57.

[59] Chai, J. D.; Head-Gordon, M. *Chem. Phys. Lett.* **2008**, *467*, 176–178.

[60] Schultz, N. E.; Zhao, Y.; Truhlar, D. G. *J Comput Chem* **2008**, *29*, 185–189.

[61] Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*.

[62] Mardirossian, N.; Head-Gordon, M. *J. Chem. Theory Comput* **2016**, *12*, 4303–4325.

[63] Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463–1473.

[64] Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

[65] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*.

[66] Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2005**, *123*.

[67] Johnson, E. R.; Becke, A. D. *J. Chem. Phys.* **2006**, *124*.

[68] Marques, M. a. L.; Gross, E. K. U. *Annu. Rev. Phys. Chem.* **2004**, *55*, 427–55.

[69] Runge, E.; Gross, E. K. U. *Phys. Rev. Letters* **1984**, *52*, 997–1000.

[70] Jamorski, C.; Casida, M. E.; Salahub, D. R. *J. Chem. Phys.* **1996**, *104*, 5134.

[71] Casida, M. E. *Recent advances in density functional methods.* **1995**, *1*, 1–34.

[72] Hirata, S.; Head-Gordon, M. *Chem. Phys. Lett.* **1999**, *314*, 291–299.

[73] Gilbert, A. T. B.; Besley, N. A.; Gill, P. M. W. *J. Phys. Chem. A* **2008**, *112*, 13164–71.

[74] Ziegler, T.; Rauk, A.; Baerends, E. J. *Theoretica Chimica Acta* **1977**, *43*, 261–271.

[75] Roos, B. *Advances in Chem. Phys.: Ab Initio Methods in Quantum Chemistry Part II* **1987**, *69*, 399–445.

[76] Andersson, K.; Roos, B. O. *Int. J. Quantum Chem.* **1993**, *45*, 591–607.

[77] Schlegel, H. B. *Bull. Korean Chem. Soc.* **2003**, *24*, 837–842.

[78] Herbert, J. M.; Head-Gordon, M. *PCCP* **2005**, *7*, 3269.

[79] Pulay, P.; Fogarasi, G. *Chem. Phys. Lett.* **2004**, *386*, 272–278.

[80] Woodcock, H. L.; Hodošček, M.; Gilbert, A. T. B.; Gill, P. M. W.; Schaefer, H. F.; Brooks, B. R. *J. Comput. Chem.* **2007**, *28*, 1485–1502.

[81] Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2006**, *117*, 185.

[82] Thiel, W.; Hummer, G. *Nature* **2013**, *504*, 96–97.

[83] Hsu, C. P.; You, Z. Q.; Chen, H. C. *J. Phys. Chem. C* **2008**, *112*, 1204–1212.

[84] Levy, S.; Speiser, S. *The J. Chem. Phys.* **1992**, *96*, 3585–3593.

[85] Pettersson, K.; Kyrychenko, A.; Ronnow, E.; Ljungdahl, T.; Mortensson, J.; Albinsson, B. *J. Phys. Chem. A* **2006**, *110*, 310–318.

[86] Dexter, D. L. *The J. Chem. Phys.* **1953**, *21*, 836–850.

[87] Consani, C.; Auböck, G.; van Mourik, F.; Chergui, M. *Science* **2013**, *339*, 1586–1589.

[88] Gray, H. B.; Winkler, J. R. *Q. Rev. Biophys.* **2003**, *36*, 341–372.

[89] McLendon, G.; Hake, R. *Chem. Rev.* **1992**, *92*, 481–490.

[90] Cogdell, R. J.; Howard, T. D.; Bittl, R.; Schlodder, E.; Geisenheimer, I.; Lubitz, W. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2000**, *355*, 1345–9.

[91] West, B. A.; Moran, A. M. *J. Phys. Chem. Lett.* **2012**, *3*, 2575–2581.

[92] Abramavicius, D.; Jiang, J.; Bulheller, B. M.; Hirst, J. D.; Mukamel, S. *J. Am. Chem. Soc.* **2010**, *132*, 7769–7775.

[93] Frauenfelder, H.; McMahon, B. H.; Fenimore, P. W. *Proc. Natl. Acad. Sci. USA.* **2003**, *100*, 8615–8617.

[94] Shih, C.; Museth, A. K.; Abrahamsson, M.; Blanco-Rodriguez, A. M.; Di Bilio, A. J.; Sudhamsu, J.; Crane, B. R.; Ronayne, K. L.; Towrie, M.; Vlček, A.; Richards, J. H.; Winkler, J. R.; Gray, H. B. *Science* **2008**, *320*, 1760–1762.

[95] Consani, C.; Auböck, G.; van Mourik, F.; Chergui, M. *Science* **2013**, *339*, 1586–9.

[96] Monni, R.; Al Haddad, A.; van Mourik, F.; Auböck, G.; Chergui, M. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, 1–5.

[97] Nölting, B.; Salimi, N.; Guth, U. *J. Theor. Biol.* **2008**, *251*, 331–347.

[98] Qiu, W.; Li, T.; Zhang, L.; Yang, Y.; Kao, Y.-T.; Wang, L.; Zhong, D. *Chem. Phys.* **2008**, *350*, 154–164.

[99] Marcus, R. a. *J. Chem. Phys.* **1956**, *24*, 966.

[100] Marcus, R. a. *J. Chem. Phys.* **1965**, *43*, 679.

[101] Marcus, R. A. *Angew. Chemie Int. Ed. English* **1993**, *32*, 1111–1121.

[102] Newton, M. D.; Sutin, N. *Annu. Rev. Phys. Chem.* **1984**, *35*, 437–480.

[103] Marcus, R. A.; Sutin, N. *Biochim. Biophys. Acta* **1985**, *811*, 265–322.

[104] de la Torre, G.; Giacalone, F.; Segura, J. L.; Martín, N.; Guldi, D. M. *Chemistry* **2005**, *11*, 1267–80.

[105] Onuchic, J. N.; Beratan, D. N.; Winkler, J. R.; Gray, H. B. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 349–377.

[106] Hsu, C. P. *Acc. Chem. Res.* **2009**, *42*, 509–518.

[107] Koslowski, T.; Burggraf, F.; Krapf, S.; Steinbrecher, T.; Wittekindt, C. *Biochim. Biophys. Acta* **2012**, *1817*, 1955–7.

[108] Blumberger, J. *Chem. Rev.* **2015**, *115*, 11191–11238.

[109] Beratan, D. N.; Betts, J. N.; Onuchic, J. N. *Science* **1991**, *252*, 1285–1288.

[110] Skourtis, S. S.; Beratan, D. N. *JBIC J. Biol. Inorg. Chem.* **1997**, *2*, 378–386.

[111] Balabin, I. A.; Hu, X.; Beratan, D. N. *J. Comput. Chem.* **2012**, *33*, 906–910.

[112] Beratan, D. N.; Onuchic, J. N.; Winkler, J. R.; Gray, H. B. *Science* **1992**, *258*, 1740–1741.

[113] Keinan, S.; Nocek, J. M.; Hoffman, B. M.; Beratan, D. N. *Phys. Chem. Chem. Phys.* **2012**, *14*, 13881–13889.

[114] Voityuk, A. A. *J. Phys. Chem. C* **2014**, *118*, 1478–1483.

[115] Voityuk, A. A. *Chem. Phys. Lett.* **2006**, *427*, 177–180.

[116] Cave, R. J.; Newton, M. D. *Chem. Phys. Lett.* **1996**, *249*, 15–19.

[117] Olaya-Castro, A.; Scholes, G. M. *Int. Rev. Phys. Chem.* **2011**, *30*, 49–77.

[118] Skourtis, S. S.; Liu, C.; Antoniou, P.; Virshup, A. M.; Beratan, D. N. *Proc. Natl. Acad. Sci. USA.* **2016**, *113*, 8115–8120.

[119] Chen, H.-C.; Hsu, C.-P. *J. Phys. Chem. A* **2005**, *109*, 11989–11995.

[120] Subotnik, J. E.; Vura-Weis, J.; Sodt, A. J.; Ratner, M. A. *J. Phys. Chem. A* **2010**, *114*, 8665–8675.

[121] Subotnik, J. E. *J. Chem. Phys.* **2011**, *135*, 71104–71120.

[122] Leng, C.; Qin, H.; Si, Y.; Zhao, Y. *J. Phys. Chem. C* **2014**, 1401–1423.

[123] Yang, X.; Bittner, E. R. *J. Phys. Chem. A* **2014**, *118*, 5196–5203.

[124] Veldkamp, B. S.; Liu, X.; Wasielewski, M. R.; Subotnik, J. E.; Ratner, M. A. *J. Phys. Chem. A* **2015**, *119*, 253–262.

[125] Briggs, E. A.; Besley, N. A. *J. Phys. Chem. A* **2015**, *119*, 2902–2907.

[126] Narth, C.; Gillet, N.; Cailliez, F.; Lévy, B.; de la Lande, A. *Acc. Chem. Res.* **2015**, *48*, 1090–1097.

[127] de la Lande, A.; Martí, S.; Parisel, O.; Moliner, V. *J. Am. Chem. Soc.* **2007**, *129*, 11700–11707.

[128] Hammi, E. E.; Houee-Levin, C.; Rezac, J.; Levy, B.; Demachy, I.; Baciou, L.; de la Lande, A. *Phys. Chem. Chem. Phys.* **2012**, *14*, 13872–13880.

[129] Melia, C.; Ferrer, S.; Rezac, J.; Parisel, O.; Reinaud, O.; Moliner, V.; de la Lande, A. *Chem. – Eur. J.* **2013**, *19*, 17328–17337.

[130] Cailliez, F.; Müller, P.; Gallois, M.; de la Lande, A. *J. Am. Chem. Soc.* **2014**, *136*, 12974–12986.

[131] Rogers, D. M.; Besley, N. A.; O'Shea, P.; Hirst, J. D. *J. Phys. Chem. B* **2005**, *109*, 23061–23069.

[132] Robinson, D.; Besley, N. A.; O'Shea, P.; Hirst, J. D. *J. Phys. Chem. B* **2009**, *113*, 14521–14528.

[133] Fujimoto, K. J.; Hayashi, S. *J. Am. Chem. Soc.* **2009**, *131*, 14152–14153.

[134] Ren, Y.; Chi, B.; Melhem, O.; Wei, K.; Feng, L.; Li, Y.; Han, X.; Li, D.; Zhang, Y.; Wan, J.; Xu, X.; Yang, M. *J. Comp. Chem.* **2013**, *34*, 1005–1012.

[135] Ren, Y.; Melhem, O.; Li, Y.; Chi, B.; Han, X.; Zhu, H.; Feng, L.; Wan, J.; Xu, X. *J. Comp. Chem.* **2015**, *36*, 137–145.

[136] Barends, T. R. M. et al. *Science* **2015**, *350*, 445–450.

[137] Choe, Y.-K.; Hashimoto, T.; Nakano, H.; Hirao, K. *Chem. Phys. Lett.* **1998**, *295*, 380 – 388.

[138] Choe, Y.-K.; Nakajima, T.; Hirao, K.; Lindh, R. *J. Chem. Phys.* **1999**, *111*, 3837–3845.

[139] Liao, M.-S.; Scheiner, S. *J. Chem. Phys.* **2002**, *117*, 205–219.

[140] Shao, Y. et al. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172–3191.

[141] Solntsev, P. V.; Sabin, J. R.; Dammer, S. J.; Gerasimchuk, N. N.; Nemykin, V. N. *Chem. Commun.* **2010**, *46*, 6581–3.

[142] Hanson-Heine, M. W. D.; George, M. W.; Besley, N. a. *J. Chem. Phys.* **2013**, *138*, 064101.

[143] Spiro, T. G.; Burke, J. M. *J. Am. Chem. Soc.* **1976**, *98*, 5482–9.

[144] Zhou, Z.; Shen, M.; Cao, C.; Liu, Q.; Yan, Z. *Chem. - A Eur. J.* **2012**, *18*, 7675–7679.

[145] Hopfield, J. J. *Proc. Nat. Acad. Sci. USA* **1974**, *71*, 3640–3644.

[146] Bortolotti, C. A.; Siwko, M. E.; Castellini, E.; Ranieri, A.; Sola, M.; Corni, S. *J. Phys. Chem. Lett.* **2011**, *2*, 1761–1765.

[147] Aprá, E.; Harrison, R. J.; Shelton, W. a.; Tipparaju, V.; Vázquez-Mayagoitia, a. *Journal of Physics: Conference Series* **2009**, *180*, 012027.

[148] Anisimov, V. M.; Bauer, G. H.; Chadalavada, K.; Olson, R. M.; Glenski, J. W.; Kramer, W. T. C.; Aprà, E.; Kowalski., K. *J. Chem. Theory Comput.* **2014**, *10*, 4307–4321.

[149] Bhaskaran-Nair, K.; Kowalski, K.; Moreno, J.; Jarrell, M.; Shelton, W. A. *J. Chem. Phys* **2014**, *141*, 074304.

[150] Fought, E. L.; Sundriyal, V.; Sosonkina, M.; Windus, T. L. *J. Comput. Chem.* **2017**, *38*, 830–841.

[151] Hiroto, S.; Miyake, Y.; Shinokubo, H. *Chem. Rev.* **2017**, *117*, 2910–3043.

[152] Chatterjee, T.; Shetti, V. S.; Sharma, R.; Ravikanth, M. *Chem. Rev.* **2017**, *117*, 3254–3328.

[153] Hosomizu, K.; Oodoi, M.; Umeyama, T.; Matano, Y.; Yoshida, K.; Isoda, S.; Isosomppi, M.; Tkachenko, N. V.; Lemmetyinen, H.; Imahori, H. *J. Phys. Chem. B* **2008**, *112*, 16517–16524.

[154] Patra, R.; Bhowmik, S.; Ghosh, S. K.; Rath, S. P. *Dalton Transactions* **2010**, *39*, 5795–5809.

[155] Weiss, C.; Kobayashi, H.; Gouterman, M. *J. Mol. Spectrosc.* **1965**, 415–450.

[156] Gouterman, M.; Wagnière, G. H.; Snyder, L. C. *J. Mol. Spectrosc.* **1963**, *11*, 108–127.

[157] Maiti, N. C.; Mazumdar, S.; Periasamy, N. *J. Phys. Chem.* **1995**, *99*, 10708–10715.

[158] Kotani, H.; Ohkubo, K.; Crossley, M. J.; Fukuzumi, S. *J. Am. Chem. Soc.* **2011**, *133*, 11092–11095.

[159] Bagchi, B.; Oxtoby, D. W.; Fleming, G. R. *Chem. Phys.* **1984**, *86*, 257–267.

[160] Mishra, B. K.; Karthikeyan, S.; Ramanathan, V. *J. Chem. Theory Comput* **2012**, *8*, 1935–1942.

[161] Karthikeyan, S.; Ramanathan, V.; Mishra, B. K. *J. Phys. Chem. A* **2013**, *117*, 6687–6694.

[162] Karthikeyan, S.; Lee, J. Y. *Chem. Phys. Lett.* **2014**, *602*, 16–21.

[163] Ethirajan, M.; Chen, Y.; Joshi, P.; Pandey, R. K. *Chem. Soc. Rev.* **2011**, *40*, 340–362.

[164] O'Connor, A. E.; Gallagher, W. M.; Byrne, A. T. *J. Photochem. Photobiol.* **2009**, *85*, 1053–1074.

[165] Anderson, H. L. *Chem. Commun.* **1999**, 2323–2330.

[166] Mozer, A. J.; Griffith, M. J.; Tsekouras, G.; Wagner, P.; Wallace, G. G.; Mori, S.; Sunahara, K.; Miyashita, M.; Earles, J. C.; Gordon, K. C.; Du, L.; Katoh, R.; Furube, A.; Officer, D. L. *J. Am. Chem. Soc.* **2009**, *131*, 15621–15623.

[167] Ryan, A.; Gehrold, A.; Perusitti, R.; Pintea, M.; Fazekas, M.; Locos, O. B.; Blaikie, F.; Senge, M. O. *Eur. J. Org. Chem.* **2011**, 5817–5844.

[168] Cheng, B.; Munro, O. Q.; Marques, H. M.; Scheidt, W. R. *J. Am. Chem. Soc.* **1997**, *119*, 10732–10742.

[169] Ohgo, Y.; Hoshino, A.; Okamura, T.; Uekusa, H.; Hashizume, D.; Ikezaki, A.; Nakamura, M. *Inorg. Chem.* **2007**, *46*, 8193–8207.

[170] Nakamura, M. *Coord. Chem. Rev.* **2006**, *250*, 2271–2294.

[171] Sazanovich, I. V.; Galievsky, V. A.; Van Hoek, A.; Schaafsma, T. J.; Malinovskii, V. L.; Holten, D.; Chirvony, V. S. *J. Phys. Chem. B* **2001**, *105*, 7818–7829.

[172] Cao, H. Y.; Si, D. H.; Tang, Q.; Zheng, X. F.; Hao, C. *J. Chem. Theory Comput.* **2016**, *1081*, 18–24.

[173] Marx, D.; Hutter, J. *Theor. Chem. Acc.* **2000**, *1*, 301–449.

[174] Blumberger, J.; Tateyama, Y.; Sprik, M. *Comput. Phys. Commun.* **2005**, *169*, 256–261.

[175] Zhang, Y.; Mao, J.; Godbout, N.; Oldfield, E. *J. Am. Chem. Soc.* **2002**, *124*, 13921–13930.

[176] Theisen, R. F.; Huang, L.; Fleetham, T.; Adams, J. B.; Li, J. *J. Chem. Phys.* **2015**, *142*.

[177] Van Cott, T. C.; Rose, J. L.; Misener, G. C.; Williamson, B. E.; Schrimpf, A. E.; Boyle, M. E.; Schatz, P. N. *J. Phys. Chem.* **1989**, *93*, 2999–3011.

[178] Nguyen, K. a.; Pachter, R. *J. Chem. Phys.* **2001**, *114*, 10757.

[179] Mack, J.; Stillman, M. J. *J. Phys. Chem.* **1995**, *99*, 7935–7945.

[180] Mack, J.; Stillman, M. J. *Inorg. Chem.* **2001**, *40*, 812–814.

[181] Condon, E. U. *Phys. Rev.* **1928**, *32*, 858–872.

[182] Coolidge, A. S.; James, H. M.; Present, R. D. *The J. Chem. Phys.* **1936**, *4*, 193–211.

[183] Fulton, R. L.; Gouterman, M. *The J. Chem. Phys.* **1961**, *35*, 1059–1071.

[184] Orlandi, G.; Siebrand, W. *The J. Chem. Phys.* **1973**, *58*, 4513–4523.

[185] Lin, S. H.; Eyring, H. *Proc. Natl. Acad. Sci. U. S. A.* **1974**, *71*, 3802–3804.

[186] Jarzqcki., A.; Kozlowski., P.; Pulay., P.; Ye., B.; Li, X. *Spectrochim. Acta: Part A* **1997**, *53*, 1195–1209.

[187] Cramariuc., O.; Hukka., T.; Rantala, T. *J. Phys. Chem. A* **2004**, *108*, 9435–9441.

[188] Miyahara., T.; Nakatsuji., H.; Hasegawa., J.; Osuka., A.; Aratani., N.; Tsuda, A. *J. Chem. Phys.* **2003**, *118*, 5802–5810.

[189] Nguyen., K.; Pachter, R. *J. Chem. Phys.* **2003**, *117*, 5802–5810.

[190] Liu., X.; Yeow., E.; Velate., S.; Steer, R. *Phys. Chem. Chem. Phys* **2006**, *8*, 1298–1309.

[191] Ricciardi., G.; Rosa., A.; Baerends., E.; Gisbergen, S. V. *J. Am. Chem. Soc* **2002**, *124*, 12319–12334.

[192] Baerends., E.; Ricciardi., G.; Rosa., A.; Gisbergen, S. V. *Coord. Chem. Rev.* **2002**, *230*, 5–27.

[193] Macak., P.; Luo., Y.; Agren, H. *Chem. Phys. Lett.* **2000**, *330*, 447–456.

[194] Duschinsky, F. *Acta. Physicochim* **1937**, *7*, 551–566.

[195] Cederbaum., L.; Domcke, W. *J. Chem. Phys.* **1976**, *64*, 603–611.

[196] Cederbaum., L.; Domcke, W. *Adv. Chem. Phys.* **1977**, *36*, 205–344.

[197] Gouterman, M. *The Porphyrins, vol. III* **1978**, *Academic Press, New York*.

[198] Peralta., G.; Seth., M.; Ziegler, T. *Inorg. Chem.* **2007**, *46*, 9111–9125.

[199] Peralta., G.; Seth., M.; H. Zhekova, T. Z. *Inorg. Chem* **2008**, *47*, 4185–4198.

[200] Yu., H.; Baskin., J.; Zewail, A. *J. Chem. Phys* **2002**, *106*, 9845–9854.

[201] He, R.; Li, H.; Shen, W.; Yang, Q.; Li, M. *J. Mol. Spectrosc.* **2012**, *275*, 61 – 70.

[202] Broderick, J. B.; Duffus, B. R.; Duschene, K. S.; Shepard, E. M. *Chem. Rev.* **2014**, *114*, 4229–4317.

[203] Hioe, J.; Zipse, H. *Faraday Discuss.* **2010**, *145*, 301–313.

[204] Dowling, D. P.; Bruender, N. A.; Young, A. P.; McCarty, R. M.; Bandarian, V.; Drennan, C. L. *Nat Chem Biol* **2014**, *10*, 106–112.

[205] Jäger, C. M.; Croft, A. K. *Chemistry* **2017**, *23*, 953–962.