PROBING THE STRUCTURE AND SOLVATION OF FERRIHAEM AND ITS CHLOROQUINE COMPLEX IN AQUEOUS SOLUTION: AN EXPERIMENTAL AND COMPUTATIONAL APPROACH

By

DAVID KUTER

MSc. (University of Cape Town)

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

In the Department of Chemistry

UNIVERSITY OF CAPE TOWN

February 2014

Declaration

PROBING THE STRUCTURE AND SOLVATION OF FERRIHAEM AND ITS CHLOROQUINE COMPLEX IN AQUEOUS SOLUTION: AN EXPERIMENTAL AND COMPUTATIONAL APPROACH

I, David Kuter, hereby declare the following:

1. The above-titled thesis is my own work in both concept and execution, apart from the normal guidance of my supervisors.
2. In cases where others work have been cited, this has been acknowledged and referenced.
3. No part of this work has been, is being or is to be submitted for another degree at this or any other university.
4. I grant the University of Cape Town free license to reproduce this work in whole or in part for the purposes of research.

Having declared these things, I hereby present this thesis for examination for the degree of Doctor of Philosophy (PhD) in Chemistry.

Signed at the UNIVERSITY OF CAPE TOWN this _____ day of February 2014.

David Kuter _____

Witness _____

Abstract

The complex formed between the antimalarial drug chloroquine (CQ) and iron(III)protoporphyrin IX (ferrihaem) in aqueous solution is of considerable interest, but the structure and stoichiometry of this complex are uncertain. The present study comprehensively investigated the interaction of CQ and ferrihaem in solution using multiple experimental and computational techniques with the aim of uncovering the solution-structure of the complex. In aqueous solution (pH 7.4), UV-visible spectrophotometric titration indicated a strong association between CQ and ferrihaem ($\log K_{obs}$ = 6.3 ± 0.2) and a Job plot showed a 1:2 CQ:ferrihaem binding ratio. The ferrihaem species in the complex was identified as the antiferromagnetically coupled μ-oxo dimer based on observed similarities with this species in magnetic moment (2.25 ± 0.02 vs. 2.0 ± 0.1 μ_B) and UV-visible and magnetic circular dichroism (MCD) spectroscopic features. These findings indicated that CQ induced μ-oxo dimer formation from the π-π dimeric form of ferrihaem in aqueous solution. The CQ-(μ-oxo ferrihaem) complex was found to persist in the solid state when precipitated from aqueous solution at pH 5. Solid-state magnetic susceptibility measurements made using an Evans balance produced a magnetic moment similar to that observed in solution (2.3 ± 0.1 μ_B). The 1:2 CQ:ferrihaem binding ratio was confirmed in the solid state based on the UV-visible spectrum of the redissolved precipitate. The presence of CQ in the solid was confirmed using mass spectrometry.

To conduct molecular modelling of the complex, reliable structural and electronic data for ferrihaem species were computed using density functional theory (DFT) calculations (OPBE/LANL2DZ). DFT models compared well with crystal structures of related iron(III)porphyrins and produced satisfactory agreement with experimental IR spectra. UV-visible spectra of H_2O- and μ-oxo ferrihaem calculated using time-dependent DFT (PBE0/LANL2DZ:LANL2DZdp) reproduced the major spectroscopic features observed from experiment for both species. Excitations giving rise to the Q-band in H_2O-ferrihaem mainly involved the Gouterman $a_{1u}/_{2u}(\pi)\rightarrow e_g(\pi^*)$ transitions. The same transitions were implicated in the Soret band but substantial contributions were also observed from porphyrin $\pi\rightarrow e_g(\pi^*)$ transitions originating in lower energy molecular orbitals. The spectroscopic feature previously assigned solely as a $\pi\rightarrow d_\pi$ charge transfer peak at lower energy than the Q-band was, in fact, found to consist of heavily admixed $\pi\rightarrow\pi^*$ and $\pi\rightarrow d_\pi$ character. In the case of μ-oxo ferrihaem, the Soret band as well as low-energy spectroscopic envelope comprised many overlapping $\pi\rightarrow\pi^*$ and $\pi\rightarrow d_\pi$ excitations with the

low-energy peak mixed to such an extent that it could not be thought of as solely a Q- or charge transfer band, but rather as a superposition of both. Characteristic spectroscopic features of the μ-oxo dimer were attributed to α- and β-spin transitions allowed by the antiferromagnetic coupling of the iron centres; the decreased ligand-field strength of the oxide ligand compared to the water ligand; doming of the iron porphyrin; and increased delocalisation of electrons over the two porphyrin rings.

Using geometries, water interaction energies and electrostatic surface potentials derived from DFT calculations, a force field compatible with the CHARMM software package was established for the simulation of non-protein bound five-coordinate ferrihaem species in aqueous solution. Molecular dynamics (MD) simulations of monomeric, π-π dimeric and μ-oxo dimeric ferrihaem species were conducted and solvation of the porphyrin system determined using spatial distribution functions (SDFs). Hydration predominated around the axial ligands in all species, but in the case of some monomeric species and μ-oxo ferrihaem, the unligated face was also strongly solvated. Based on the excellent fit to the experimental extended X-ray absorption fine-structure (EXAFS) spectrum of μ-oxo ferrihaem using SDF-derived structural data, an equilibrium between five- and six-coordinate species with water acting as the sixth ligand was proposed. MD simulations of the CQ-ferrihaem complex was conducted in two conformations – one where CQ π-stacked on the outer face of μ-oxo ferrihaem and the other where CQ docked between the porphyrin rings. Correlation of computed and experimental ^{1}H NMR T_1 relaxation times as well as fits to the EXAFS spectrum of the complex provided convincing evidence that the latter model was most likely closest to the solution structure. Three prominent interactions identified in the docked structure provided the first rationalisation of previously reported structure-activity relationships for CQ-ferrihaem association. These included an intermolecular hydrogen bond between the protonated quinoline nitrogen and μ-oxo bridge; π-stacking between the quinoline and porphyrin ring system; and close contact interactions between methyl hydrogen atoms of ferrihaem and the chlorine substituent of CQ. Correlation of computed and experimental diffusion coefficients suggested a tetrameric, [CQ-(μ-oxo ferrihaem)]$_2$, complex formed in aqueous solution which was supported by an empirical relationship between diffusion coefficient and molecular volume. This methodical approach and combination of multiple experimental and computational methods has provided the most accurate structural model of the CQ-ferrihaem complex to date.

Publications and Conference Proceedings

Publications

1. D. Kuter, G. A. Venter, K. J. Naidoo and T. J. Egan, Experimental and time-dependent density function theory characterisation of the UV-visible spectra of monomeric and μ-oxo dimeric ferriprotoporphyrin IX, *Inorg. Chem.*, 2012, **51**, 10233-10250.

2. D. Kuter, S. J. Benjamin and T. J. Egan, Multiple spectroscopic and magnetic techniques show that chloroquine induces formation of the μ-oxo dimer of ferriprotoporphyrin IX, *J. Inorg. Biochem.*, 2014, **113**, 40-49.

Conference Proceedings

1. South African Chemical Institute National Conference, **2011**, Johannesburg, RSA.
   Poster: D. Kuter, K. J. Naidoo and T. J. Egan, Understanding the mechanism of action of the antimalarial chloroquine.

2. 15th International Conference on Biological Inorganic Chemistry, **2011**, Vancouver, Canada.
   Poster: D. Kuter, K. J. Naidoo and T. J. Egan, A density functional theory investigation of ferriprotoporphyrin IX species.

3. Frontiers in Scientific Computing Symposium, **2012**, Cape Town, RSA.
   Poster: D. Kuter, G. A. Venter, K. J. Naidoo and T. J. Egan, Predicting UV-visible spectra of ferriprotoporphyrin IX species using time-dependent density functional theory.

4. 16th International Conference on Biological Inorganic Chemistry, **2013**, Grenoble, France.
   Poster: D. Kuter, K. J. Naidoo and T. J. Egan, Investigating aqueous solvation of ferriheme species using molecular dynamics.

Dedication

To Craig and Sandy Kuter

Acknowledgments

While many people have made varying degrees of contributions to my work, I would like to express my deepest and sincere gratitude to a few below:

Professor Timothy Egan: For being undoubtedly the best supervisor a student could hope for. I will be forever grateful for the mentorship you have given me. If I get to be even half the supervisor to others as you have been to me, I will count myself incredibly fortunate.

Professor Kevin Naidoo: For taking a chance on an experimentalist and allowing me to dabble in the realm of Computational Chemistry.

To my "Pseudo-Supervisors": Dr Gerhard Venter and **Dr Chris Barnett** – For the many discussions both related and completely unrelated to my project. Without these, I wouldn't know half as much as I do now about Scientific Computing. **Pete Roberts, Professor Ravenscroft** and **Professor Jackson** – For their extensive time spent showing me the ropes of the NMR spectrometers. **Dr Victor Streltsov** – For the many long-distance EXAFS lessons as well as for hosting me during my time at the Australian Synchrotron.

The "Haem Team" members: Particularly **Dr Kanyile Ncokazi**, **Dr Katherine de Villiers**, **Aneesa Omar**, **John Woodland** and **Roxanne Mohunlal**, for the friendship, putting up with my neat-freak ways and making the overall time spent in the lab most enjoyable. A person could not ask for better lab mates. **SCRU members:** Importantly, **Krishna Govender** and **Riedaa Gamieldien**, for the coffee breaks and continuous computing-related questions. Your patience and willingness to help is much appreciated.

To present and past friends: Specifically **Siri Lie Olsen**, **Craig and Caylee Rossouw**, **Dale Taylor**, **Roxanne Mohunlal** and **Nikki Dare**, for keeping me sane and helping me to realise that sometimes it's OK not to work.

Staff at the CHPC: For all the help getting my jobs to run, especially with the months of problems faced by TD-DFT calculations.

Finally, and most importantly, to my family, especially my **mother** and **father** for their never-ending support, love and care as well as for showing so much interest in my project despite never understanding a word of it.

Abbreviation List

CD	Circular dichroism
CGenFF	CHARMM general force field
CHES	*N*-Cyclohexyl-2-aminoethanesulfonic acid
CHPC	Centre for High Performance Computing
CI-S	Configuration Interaction – Singles
COM	Centre of mass
CQ	Chloroquine
CQR	Chloroquine-resistant
CQS	Chloroquine-sensitive
CSD	Cambridge Structural Database
CT	Charge transfer
DMSO	Dimethyl sulfoxide
DV	Digestive vacuole
DFT	Density functional theory
ECP	Effective core potential
ESP	Electrostatic surface potential
EXAFS	Extended X-ray absorption fine-structure
Ferrihaem	Iron(III)protoporphyrin IX
FID	Free induction decay
GGA	Generalised gradient approximation
GIXD	Grazing incidence X-ray diffraction
GTO	Gaussian-type orbitals
Haem	Iron(II)protoporphyrin IX
Haematin	HO-iron(III)protoporphyrin IX
Haemin	Cl-iron(III)protoporphyrin IX
Hb	Haemoglobin
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HF	Hartree-Fock
HOMO	Highest occupied molecular orbital
HPC	High performance computing
IC_{50}	50% inhibitory concentration
IR	Infrared
ITC	Isothermal titration calorimetry
K76T	Lysine 76 mutation to threonine
KS	Kohn-Sham
LCAO	Linear combination of atomic orbitals
lcp	Left circularly polarised light
LDA	Local density approximation
LSDA	Local spin density approximation
LSP	Least-squares plane
LUMO	Lowest unoccupied molecular orbital
MCD	Magnetic circular dichroism
MD	Molecular dynamics
metMb	Metmyoglobin
M_L	Angular momentum quantum number
MM	Molecular mechanics
MMG	1-Myristoyl-glycerol
MO	Molecular orbital
MSD	Mean square displacement
MUE	Mean unsigned error
N-AcMP8	*N*-Acetylmicroperoxidase-8
NMR	Nuclear magnetic resonance
NPT	Isobaric-isothermal ensemble
NVT	Canonical ensemble
OEP	Octaethylporphyrin
OOP	Out of plane
PDB	Protein Data Bank
PEG400	Polyethyleneglycol 400
PfCRT	*Plasmodium falciparum* chloroquine resistance transporter
PfMDR1	*Plasmodium falciparum* multi-drug resistance transporter
QM	Quantum mechanical
RBC	Red blood cell
rcp	Right circularly polarised light
RDF	Radial distribution function
RR	Resonance Raman
SAR	Structure-activity relationship
SCF	Self-consistent field
SDF	Spatial distribution function
SQUID	Superconducting quantum interference device
STO	Slater-type orbitals
TD-DFT	Time-dependent DFT
TPP	Tetraphenylporphyrin
WIE	Water interaction energy
XAFS	X-ray absorption fine-structure
XANES	X-ray absorption near-edge spectroscopy

Table of Contents

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 The Role of Haem in Biology

1.2.1 Haemoproteins.

Iron(II)protoporphyrin IX or haem is one of the most ubiquitous prosthetic groups in biological systems and is essential to the survival of almost all higher forms of life.[1] Haem consists of an iron atom incorporated into a protoporphyrin IX ring consisting of a macrocycle of four pyrrole rings with four methyl, two vinyl and two propionic acid substituents (see Fig. 1.1a). This relatively simple molecule plays a crucial role in a variety of biological processes, often through incorporation into protein structures to form haemoproteins. To date, over 4000 crystal structures of haemoproteins are listed in the Protein Data Bank (PDB).[2] Functions of haemoproteins cover a broad range in biological systems, although the most well-known is arguably that of oxygen storage and transport by haemoglobin (Hb, Fig. 1.1b) in red blood cells (RBCs). In humans, RBCs produce and contain approx. 80% of haem in the form of haemoglobin and other haemoproteins.[1] Roughly 15% of haem is made and present in the liver where processes including metabolism of compounds by haem-containing enzymes such as cytochrome P450 occurs (Fig. 1.1c).[1, 4] Other important functions of haemoproteins include electron transport by cytochrome c; the biosynthesis of endogenous molecules such as lipids and steroids; synthesis of compounds crucial to the immune response by cytochrome peroxidases; and synthesis of neuromodulators such as nitric oxide by nitric oxide synthase and carbon monoxide by haem oxygenase enzymes.[1, 5-8]

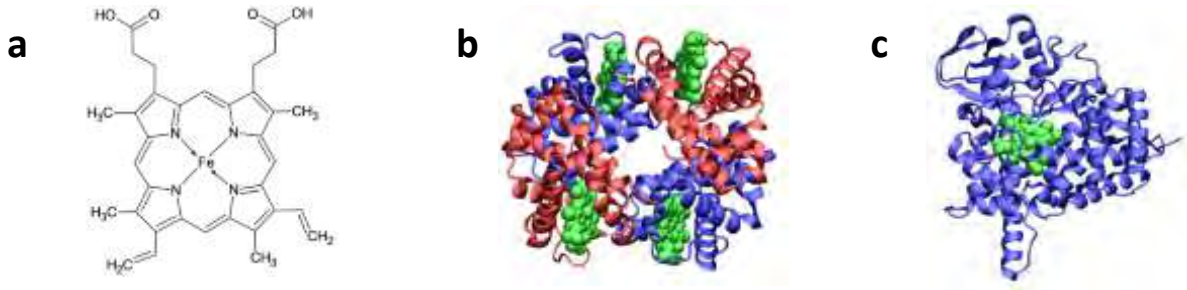


Figure 1.1. (**a**) The structure of iron(II)protoporphyrin IX or haem. In the oxidised iron(III) form, it is known as ferrihaem. Possible axial ligands not shown for clarity. Examples of haemoproteins, (**b**) haemoglobin (Hb) and (**c**) cytochrome P450. Haem moiety shown in green for clarity. Protein structures were created using crystal structure data obtained from the PDB,[2] PBD ID: (b) 2W72,[9] (c) 2Q9F.[10]

The versatility of biological functions of haemoproteins is mainly attributed to the haem prosthetic group. Firstly, the iron centre is able to accommodate one or two axial ligands which allows it to associate with proteins and small molecules such as oxygen and carbon monoxide.[1] Common protein ligands include side chains of histidine, methionine and cysteine as well as, more rarely, tyrosine residues. The binding of small molecules by five-coordinate haem moieties facilitates their transport, sensing and use by haemoproteins.[1] Secondly, the iron centre of haem can interconvert between a variety of oxidation states, enabling electron transport and oxidation/reduction processes. The more stable oxidation states of the iron centre are +2 and +3.[11] Finally, the porphyrin ring system is able provide both non-covalent and covalent linkages with the protein. The hydrophobic methyl and vinyl substituents interact with hydrophobic regions of proteins.[1, 8] In addition, these substituents are able to form covalent bonds with proteins as in the case of c-type cytochromes in which two cysteine amino acids are attached to the two vinyl moieties. Similar attachments have also been observed at the propionic acid substituents.[8, 12] Furthermore, propionic acid groups are able to form hydrogen bonds with amino acid residues or solvent molecules within the protein.[1] A summary of the features of the haem prosthetic group is shown in Fig. 1.2.

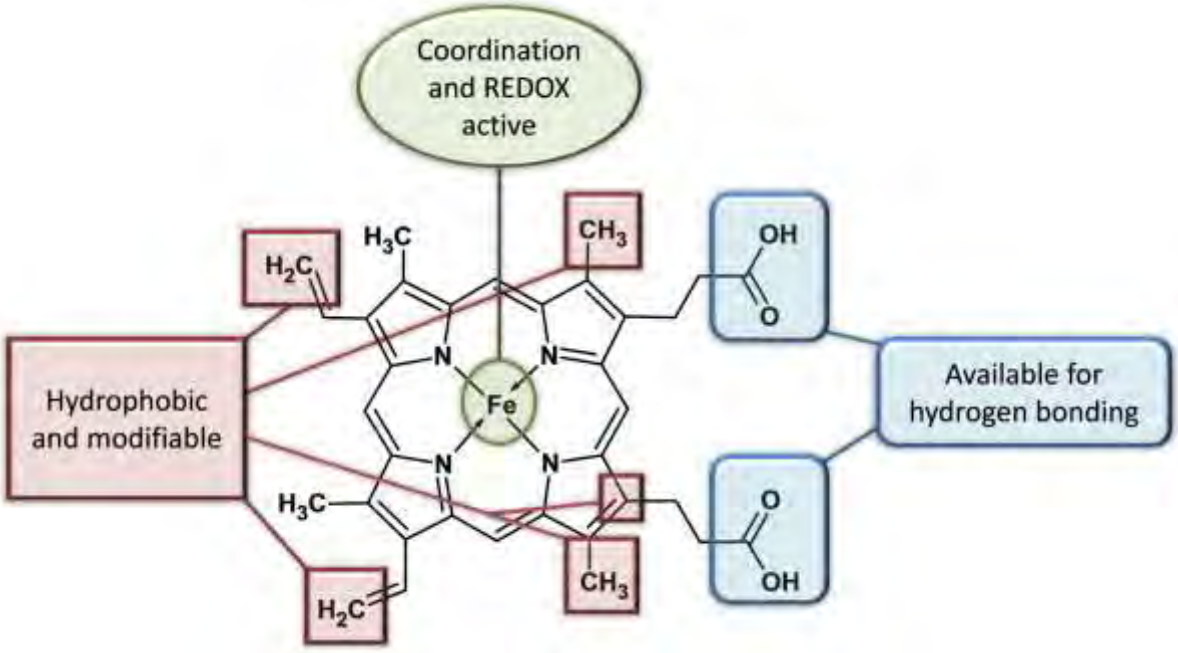


Figure 1.2. Diagrammatic representation of the functionality of the haem prosthetic group which facilitates its wide biological applications.

1.1.2 Free haem.

While extensive literature exists regarding the role haem plays in haemoprotein function, the free (unbound) form of this molecule has received much less attention. This has changed in recent years owing to its implication in a wide range of biological functions such as signalling, gene transcription and protein folding.[7, 12-14] Studies have shown haem to be an inducer of differentiation in a variety of cell types including mouse 3T3,[15] neuroblastoma[16] and leukemic cells[17] as well as in human hematopoietic[18] and leukemic cells.[13, 17, 19] On the other hand, the release of free haem from haemoproteins has been associated with a variety of conditions and diseases, for example inflammation, atherogenesis, cancer and Alzheimer's disease.[20-24] The negative influence of free haem in these illnesses is brought about by the inherent toxicity of this molecule. In high concentrations, haem catalyses the production of reactive oxygen species which damage cellular components such as DNA, proteins and lipids.[25-28] Furthermore, it has a tendency promote cell lysis and death by accumulating in membranes, introducing disorder and thus increasing permeability.[27, 29, 30]

The human body employs haem-degrading enzymes in the form of haem oxygenases to manage haem concentrations and thus prevent the toxic effects associated with this molecule from overwhelming cells.[6, 31] Haem oxygenase acts by binding free haem within a cell and degrading it into biliverdin, carbon monoxide and free ferrous iron (see Fig. 1.3).[31] For every two haem molecules degraded, six oxygen molecules, seven NADPH molecules and fourteen electrons are consumed.[31]

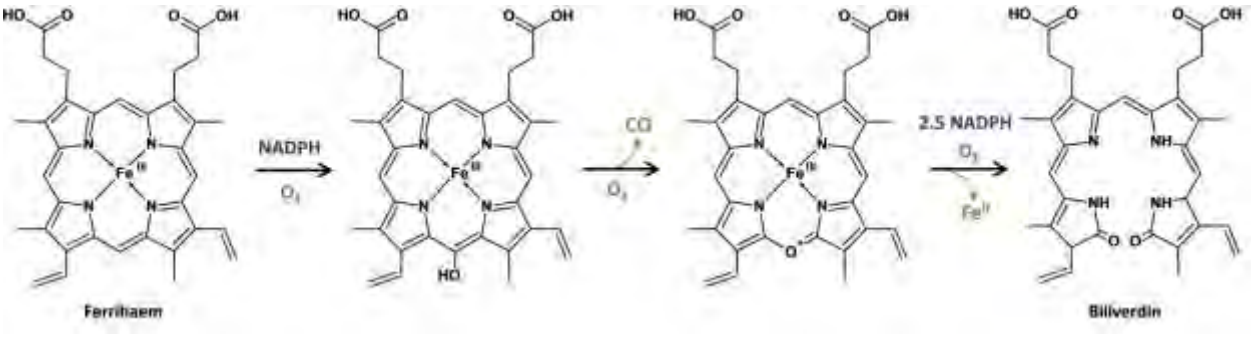


Figure 1.3. Mechanism of haem degradation.

However, during severe haemolytic events, such as after a haemorrhagic stroke, the concentrations of haem released from haemoproteins can be as high as 10 mM and simply overwhelm haem oxygenase enzymes.[32] In these circumstances, catabolism of haem quickly depletes NADPH stores, slowing the recycling of glutathione disulphide by enzymes which are the primary defence against reactive oxygen species and thus allowing the toxic effects of haem to persist.[32] Severe haemolysis also occurs during diseases; for example in sickle cell anaemia or malaria.[33, 34] On the other hand, ferrihaem has also been shown to play a crucial role in treatment of the latter disease which contributes significantly to the global health burden with more than 219 million cases being reported in 2010 and resulting in approximately 660 000 deaths.

1.2 Haem in Malaria

1.2.1 The malaria parasite.

Malaria is caused through infection by the *Plasmodium* parasite which is transmitted to the human host through the saliva of a female *Anopheles* mosquito vector.[36, 37] Five species of the parasite have been associated with malaria; namely *P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale* and *P. knowlesi*.[37, 38] *P. falciparum* is considered the mostly deadly and, while *P. vivax* is also considered dangerous, it tends to cause fewer deaths than *P. falciparum*. *P. malariae* and *P. ovale* are mild by comparison and are not known to cause severe disease in humans.[39] Finally, *P. knowlesi* has only recently been discovered in the human population in South East Asia, but this species is well known for causing malaria in rhesus monkeys.[38, 40]

The malaria parasite has a complicated life cycle consisting of several stages in which it undergoes many morphological changes (see Fig. 1.4). It first enters the human blood stream in the form of sporozoites which migrate to the liver, invade hepatic cells and undergo asexual fission forming liver schizonts.[36] During this liver stage, lasting between five and twenty-five days, the human host remains free of the symptoms associated with malaria, which include fever and chills.[37, 41] Rupture of the liver schizont releases parasites into the blood stream in the form of merozoites which invade RBCs and begin

what is termed the blood stage.[37] Whilst in the RBC, the parasite utilises Hb as a food source and develops into a form known as a trophozoite. Subsequent formation of a blood schizont then occurs and, when ruptured, releases between six and twenty-four parasites into the blood stream.[37] The rupture of the blood schizont also releases toxins and RBC debris into the blood stream which are responsible for the observed symptoms of malaria. The parasites released from schizont-rupture then re-infect RBCs and the cycle continues. During these cycles, the parasite load is increased between eight and ten-fold in *P. falciparum* and occurs at intervals between 48 and 72 hours depending on the parasite species.[36, 37] Certain parasites within the RBCs develop into male and female gametocytes which undergo sexual reproduction in the mosquito vector when it feeds on an infected human. These produce oocytes that rupture to release sporozoites. Finally, the sporozoites migrate to the salivary glands of the mosquito, completing the life cycle of the parasite.[36, 37]

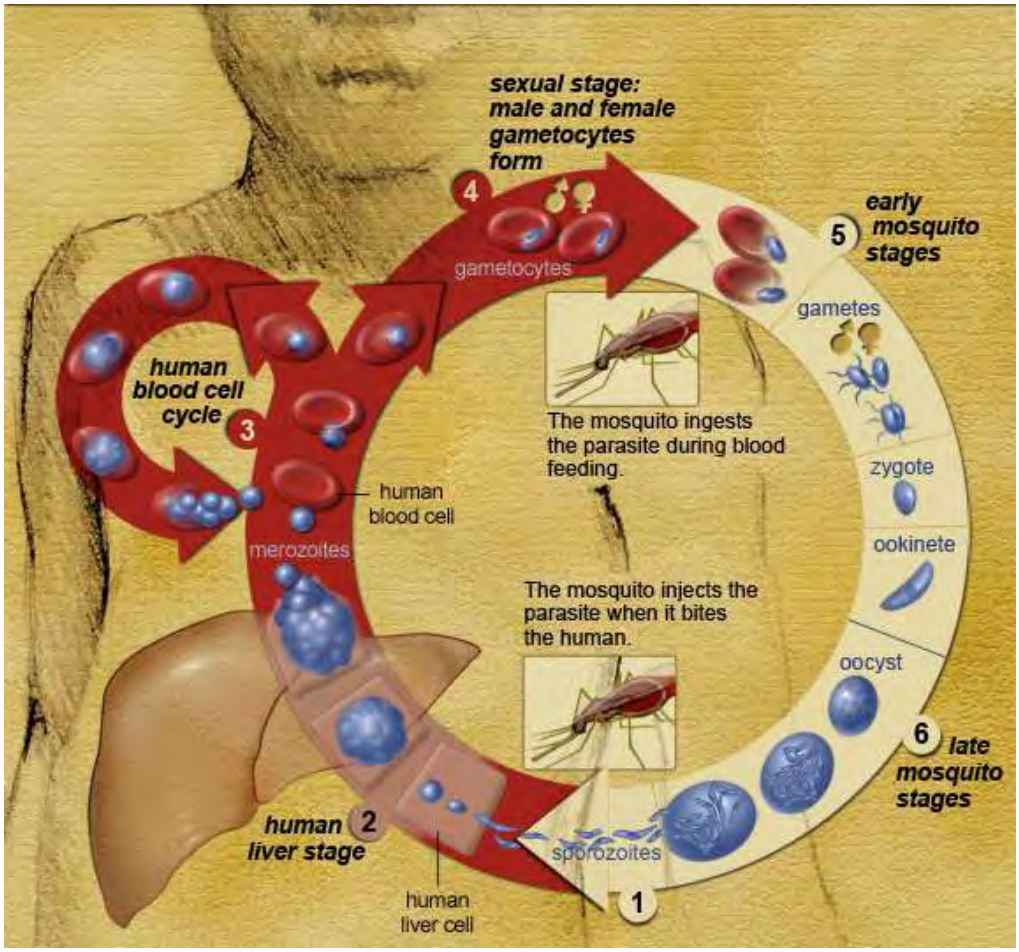


Figure 1.4. The life cycle of the malaria parasite.[42]

1.2.2 Haemoglobin degradation and haem detoxification.

The most pertinent phase relating to haem in the malaria parasite life cycle is in the blood stage, where the parasites invade human RBCs to evade the host immune system. Whilst inside the RBC, the malaria parasite ingests between 60 and 80% of the Hb present to use as a source of amino acids for its continued proliferation as well as to maintain osmotic balance and provide space for growth.[43, 44] Endocytosis of Hb occurs via vesicles created from parasitic plasma membranes which transport the protein to a specialised acidic compartment (pH 4.8 – 5.2) known as the digestive vacuole (DV).[45, 46] Inside the DV, the globin portion of Hb is degraded into peptides by proteolytic enzymes. These include cysteine proteases falcipain 1 and 2; zinc metalloprotease facilysin and aspartic proteases plasmepsin I, II and IV as well as histo-aspartic protease.[47-49] Degradation of peptides to amino acids occurs through enzyme-catalysed hydrolysis in the parasitic cytosol.[50] While the globin portion of Hb provides the amino acids essential to parasite survival, the haem portion is released as a by-product which is rapidly oxidised to ferrihaem.[51] The presence of this molecule is harmful to the parasite given the toxic effects to biological systems described in Section 1.1.2. Unlike humans, the malaria parasite does not possess haem oxygenase enzymes to degrade ferrihaem and circumvent its toxicity. Consequently, the parasite has evolved an alternate mechanism of detoxification. Instead, free ferrihaem is converted into a less toxic crystalline form, known as haemozoin, which is made up of ferrihaem dimer units known as μ-propionato ferrihaem.[52] These involve reciprocal coordination of the propionate side chains of one monomer to the iron centre of the other. Hydrogen bonding between the remaining propionic acid side chains and its neighbours facilitates the biomineralisation of the haemozoin crystal (see Fig. 1.5 inset).[53] Hb degradation and ferrihaem detoxification processes are summarised in Fig. 1.5.

The subject of haemozoin formation has been one of great interest and not only to those in the field of malaria. Other blood-feeding organisms such as *Schistosoma* trematodes and *Rhodnius* insects, which cause bilharzia and act as vectors in Chagas disease respectively, share the same detoxification process.[54, 55] Understanding the formation of this molecule has been approached from two aspects. The first relates to investigations of the solution behaviour of ferrihaem as well as its speciation, while the second has focused on the mechanism of haemozoin crystallisation which includes what components are necessary for its formation. Both are discussed in Sections 1.2.3 and 1.2.4 respectively.

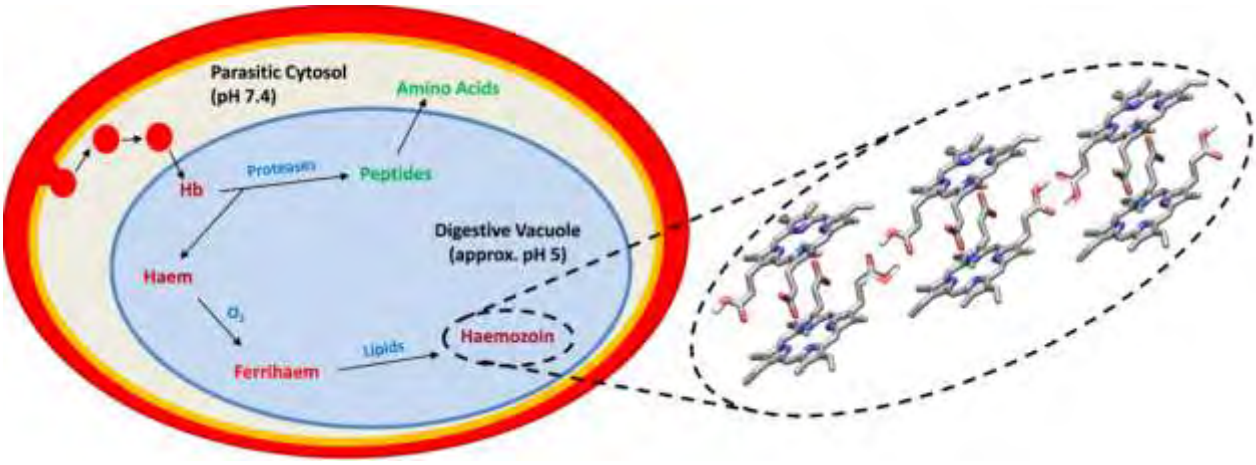


Figure 1.5. Hb degradation and ferrihaem detoxification by the malaria parasite (grey) inside the human RBC (red). The structure of haemozoin is displayed in the dashed inset.

1.2.3 Ferrihaem in solution.

Several early studies of ferrihaem in aqueous solution using diffusion measurements suggested it formed large aggregates, although consensus on the size of these aggregates was not reached.[56-60] In one study, particles with weights of 50 000 to 145 000,[56] and even 4 000 000 over time,[57] were proposed to exist in solution, while another suggested aggregates with average weight of 31 000 were instead present.[59] Subsequent kinetic experiments were seemingly supportive of aggregate formation in solution, where changes to UV-visible spectra and differences in kinetics were observed when using aged solutions as opposed to using fresh ones.[61, 62] In a seminal paper by Brown *et al.*[63], however, a thorough investigation of ferrihaem aggregation in solution was undertaken using UV-visible spectroscopy. Based on observed spectroscopic changes in the UV-visible spectrum of ferrihaem with increasing concentration which adhered to a dimerisation process, they concluded that ferrihaem existed in solution as a dimeric species and not as higher order aggregates.

The form of the dimeric ferrihaem species was proposed by the same authors in a separate paper as a μ-oxo dimer which contained an oxide ligand bridging two ferrihaem molecules through coordination to the iron centres (see Fig. 1.6).[64] This was based on infrared (IR) evidence obtained from the solid precipitated when NaOH was added to haemin in DMSO in which a characteristic Fe-O-Fe stretching frequency was observed at 900 cm^{-1}. Their model of the μ-oxo ferrihaem dimer is depicted in Fig 1.6a which consisted of antiferromagnetically coupled iron centres deviating 0.5 Å out of the porphyrin plane towards the axial oxide ligand, a Fe-O bond length of 1.8 Å, Fe-O-Fe bond angle of between 165 and 180°, and a water molecule present in the sixth coordination site. With the exception of the last point, the model proposed by Brown *et al*. is consistent with crystal structures reported for a range of μ-oxo dimeric metalloporphyrins,[65-71] including that of the μ-oxo ferrihaem dimethyl ester (see Fig. 1.6b).[72]

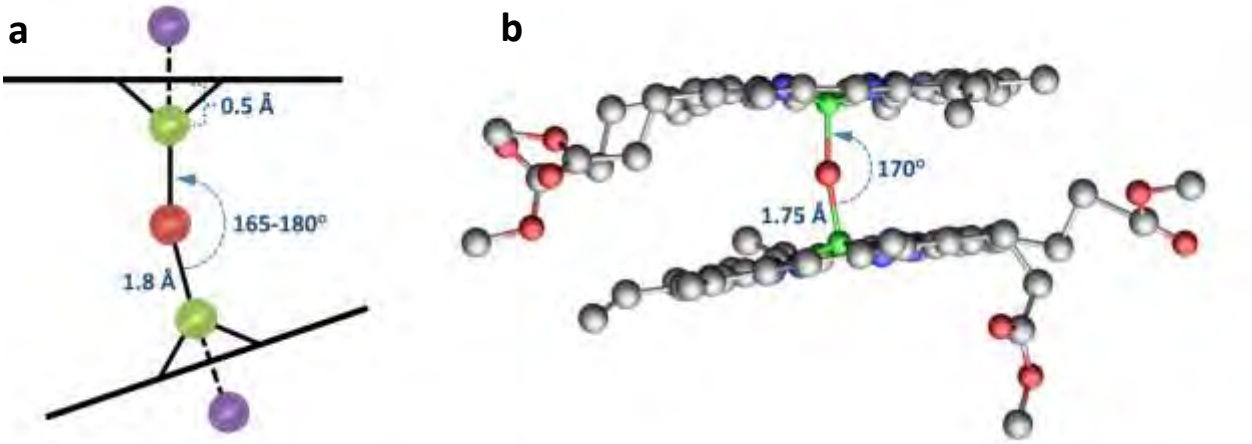


Figure 1.6. (**a**) Model of the μ-oxo ferrihaem dimer as proposed by Brown *et al*.[64] Green and red spheres represent the iron atoms and oxide ligand respectively, while the purple sphere represents the sixth coordinated water or hydroxide ligand. (**b**) The crystal structure of the μ-oxo dimer of ferrihaem dimethyl ester reported by Chen *et al*.[72] bears close similarity to that proposed in (a). Structure reproduced using the crystal structure obtained from the Cambridge Structural Database (CSD).

The existence of μ-oxo dimeric ferrihaem as the dominant form in aqueous solution was later questioned based on the work of O'Keeffe *et al*.[74] These authors performed extensive studies on 2,4-disubstituted iron(III)porphyrins, characterising them using magnetic susceptibility measurements, ^{1}H NMR, IR and UV-visible spectroscopy. They found that μ-oxo dimers did not spontaneously form in aqueous solution but rather had to be induced by addition of a small amount of a weak organic base (e.g. 10% pyridine) to an alkaline aqueous solution. Despite these findings, the misconception that μ-oxo ferrihaem was the major species in aqueous solution persisted until recently when it was convincingly shown by de Villiers *et al*.[75] to exist as a π-π dimer instead. This form consists of two five-coordinate ferrihaem molecules associated through π-stacking interactions between the unligated faces and is readily distinguishable from the μ-oxo species by magnetic susceptibility measurements and UV-visible spectroscopy. The former experimental method produced magnetic moments of approx. 4.6 for the π-π dimer and about 2.0 for the antiferromagnetically coupled μ-oxo species, while the latter technique produced characteristic spectroscopic changes, particularly in the long wavelength region, that were consistent with those observed by O'Keefe *et al*. for related μ-oxo dimers (see Fig. 1.7).

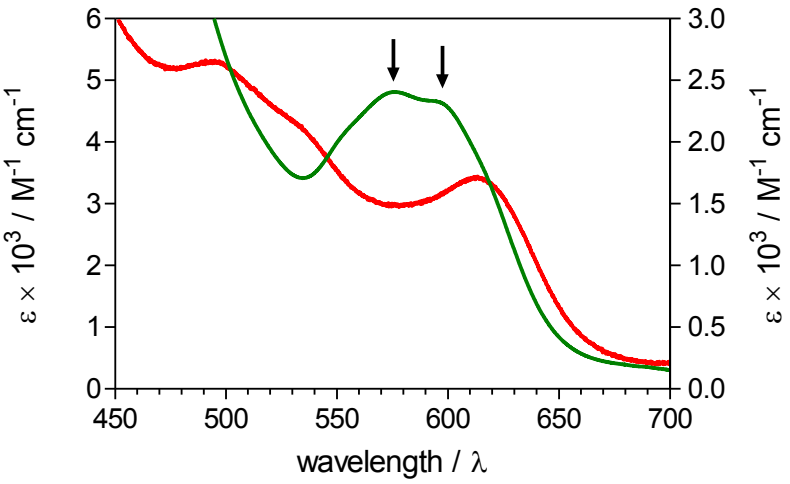


Figure 1.7. UV-visible spectra of π-π dimeric (red, left axis) and μ-oxo dimeric ferrihaem (green, right axis) species. Charge transfer bands proposed by O'Keeffe *et al*.[74] are labelled with arrows.

The existence of the π-π dimer in aqueous solution was further supported by a later investigation of the effect of alkaline aqueous mixtures of water-miscible organic solvents on ferrihaem speciation. This study found that inclusion of aprotic solvents, such as DMSO, induced μ-oxo dimer formation, while protic solvents such as methanol favoured π-π dimer formation.[76] Furthermore, high salt concentrations in aqueous solution were also shown to induce μ-oxo dimer formation. The induction of the μ-oxo dimer by aprotic solvents was attributed to the entropically-favourable release of water solvating the axial hydroxide ligand which was concentrated at this position relative to the bulk. This phenomenon was thought to only occur with aprotic solvents because they are unable act as hydrogen bond donors to the hydroxide ligand. On the other hand, increased localisation of water relative to the bulk does not occur with protic solvents as they are able act in a similar manner to water having both hydrogen bond donor and acceptor groups. Unfortunately, the misconception that the μ-oxo dimer is the major species in aqueous solution still permeates the literature with even a recent review[77] stating this while referencing the paper that disproves it.

The structure of the π-π dimer in implicit water solvent was investigated by de Villiers *et al*.[78] using molecular dynamics (MD) simulations. The study employed the force field developed by Marques and co-workers[79, 80] for metalloporphyrins. These parameters did not take into account atomic charges and force constant restraints were imposed on the structure to maintain π-stacking interactions between the ferrihaem molecules using dummy atoms. The findings indicated that orientations in which the propionate side chains were 180° (anti) or ± 135° with respect to each other were favoured (see Fig. 1.8a) and that the majority of geometries had lateral Fe-Fe shifts of between 2 and 4 Å (see Fig. 1.8b).

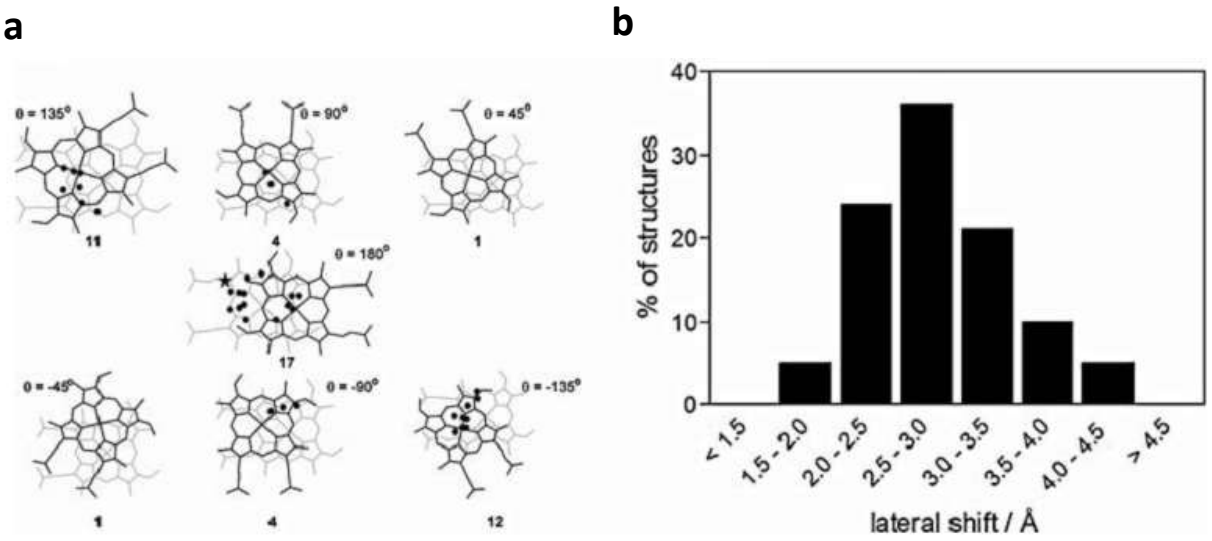


Figure 1.8. (**a**) Representative clusters of the π-π dimer structure computed using MD simulation. (**b**) Histogram of lateral shifts over the course of the simulation. Figure reproduced with permission from de Villiers *et al*.[75] © 2006, Springer.

1.2.4 Mechanism of haemozoin formation.

The mechanism of haemozoin formation has been the topic of much debate in the literature. There are two main schools of thought that have been proposed in this regard. The first is that proteins such as histidine-rich protein II or haem detoxification protein facilitate haemozoin formation.[81-83] The second proposed hypothesis is that neutral lipids within the parasite DV or even phospholipids provide a nucleation point and mediate its formation.[84, 85] While some support for the first mechanism of haemozoin formation has been recently reported,[82] the large majority of literature reports tend to support the second mechanism. Early reports by Pisciotta *et al*.[86] showed haemozoin crystals closely associated with neutral lipid droplets within the DV of *P. falciparum* parasites (see Fig. 1.9a and b) and it was suggested that haemozoin formation occurs inside these lipid nanospheres. Indeed, formation of the synthetic equivalent of haemozoin, β-haematin, was shown to spontaneously and rapidly occur *in vitro* at neutral lipid-water interfaces under biological conditions (pH 4.8, 37 °C).[87, 88] Alcohols such as pentanol and octanol which have a hydrophobic chain and hydrophilic head group were also able to mediate its formation.[89, 90] A grazing incidence X-ray diffraction (GIXD) study determined the packing structure of a neutral lipid, 1-myristoyl-glycerol (MMG), monolayer which shared similar head group

spacing to the propionic acid side chains of the (100) face of β-haematin crystals. [91] These authors suggest this partial alignment would be able to promote β-haematin formation through epitaxy (see Fig. 1.10). Recent support for this was reported by Ambele *et al*.[92] where the (100) face of β-haematin was observed *in vitro* to be in contact with the surface of pre-formed neutral lipid. No evidence of formation within the lipid droplets was observed. Furthermore, these authors and Kapishnikov *et al*.[93, 94] could not find evidence of neutral lipid droplets associated with haemozoin inside the malaria parasite DV. Rather, the latter authors provided evidence that haemozoin aligned with the membrane of the DV (see Fig. 1.9c and d). Recent *in vitro* evidence in support of membrane-mediate haemozoin formation has been reported by Huy *et al*.[85] where β-haematin was efficiently synthesised at phospholipid-water interfaces under biological conditions.

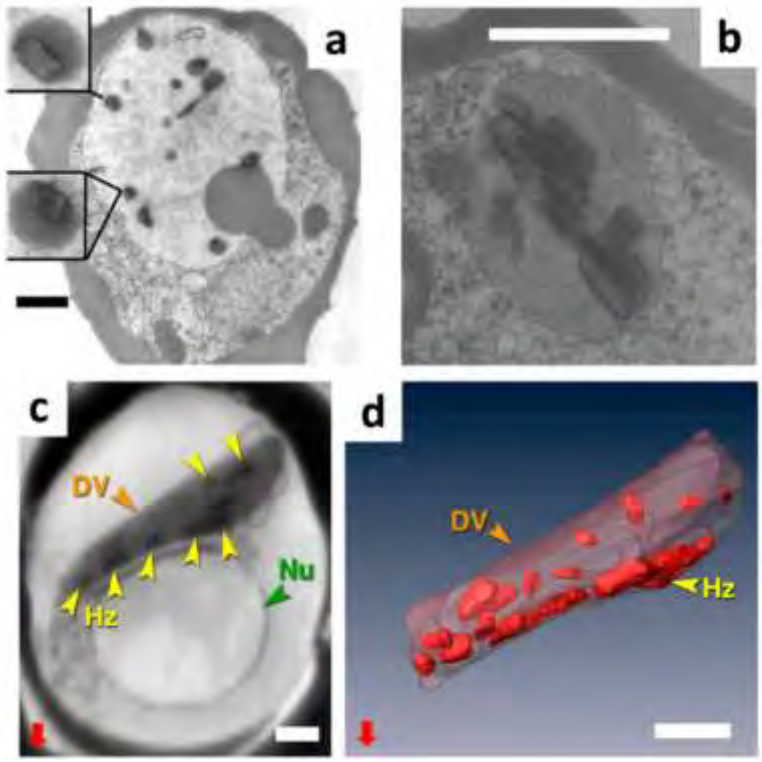


Figure 1.9. Transmission electron micrograph images of (**a**) an early stage and (**b**) a mature *P. falciparum* trophozoite showing haemozoin crystals associated with lipid nanospheres in the parasite DV. Scale bar = 1 μm. Figure adapted with permission from Pisciotta *et al*.[86] © 2007, Biochemical Society. (**c**) Serial surface view electron micrograph of *P. falciparum* trophozoite showing alignment of haemozoin (Hz) with the DV membrane. Nu = nucleus. (**d**) Surface rendering of Hz crystals on the DV membrane. Scale bar = 0.5 μm. Figure adapted with permission from Kapishnikov *et al*.[93] © 2012, National Academy of Sciences.

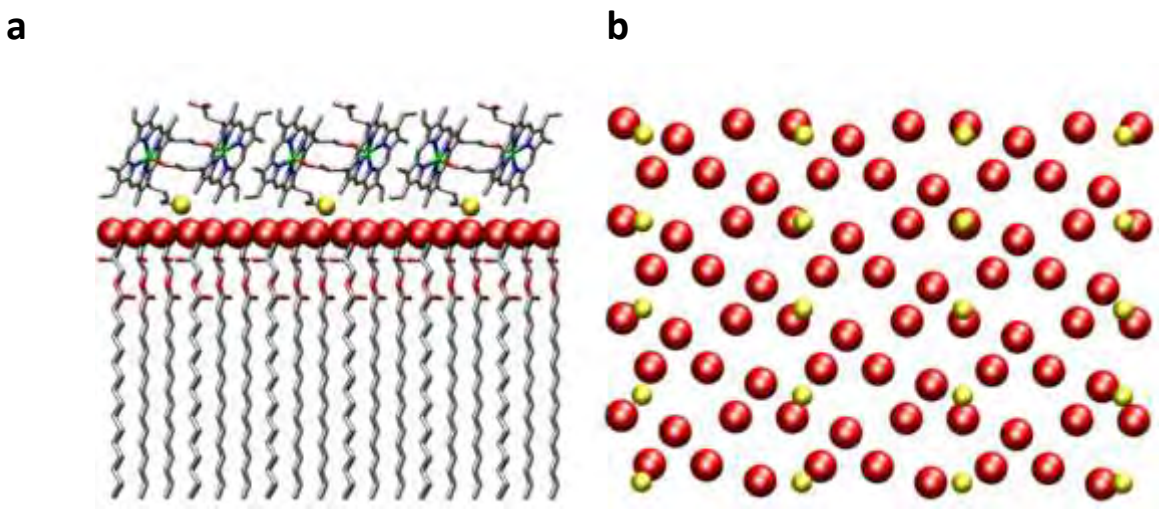


Figure 1.10. Packing structure of the neutral lipid MMG as determined by GIXD, overlaid with the crystal structure of haemozoin from the (**a**) side and (**b**) top views. Lipid head groups are represented by red spheres and propionic acid side chains of haemozoin by yellow spheres. Figure adapted with permission from Ambele *et al*.[92] © 2013, American Chemical Society.

MD simulations by Egan *et al*.[89] of the π-π ferrihaem dimer in vacuum and aqueous solution provided some insight into the mechanism of haemozoin formation. In vacuum, these authors showed that the π-π dimer spontaneously and rapidly adopted a conformation that is a precursor to the μ-propionato ferrihaem species and thus also a precursor to haemozoin (see Fig. 1.11a, structure ii). Conversion from the precursor to μ-propionato species simply requires coordination of propionic acid groups to the iron centre of ferrihaem and subsequent release of axially ligated water molecules (see Fig. 1.11b). This precursor species was shown not to survive in aqueous solution, however, owing to competing hydrogen bonding interactions between water molecules and the propionate side chains which caused the latter to move away from the iron centres (see Fig. 1.11a, structure iii). It was proposed that the precursor dimer prevails at lipid-water interfaces.

The formation of haemozoin presents an attractive target for antimalarial drugs as this detoxification mechanism is not shared by the human host and is not under the direct genetic control of the parasite since haemoglobin is sourced from its host. Indeed, the activity of a number of antimalarial drugs has been proposed to arise from the inhibition of haemozoin formation.

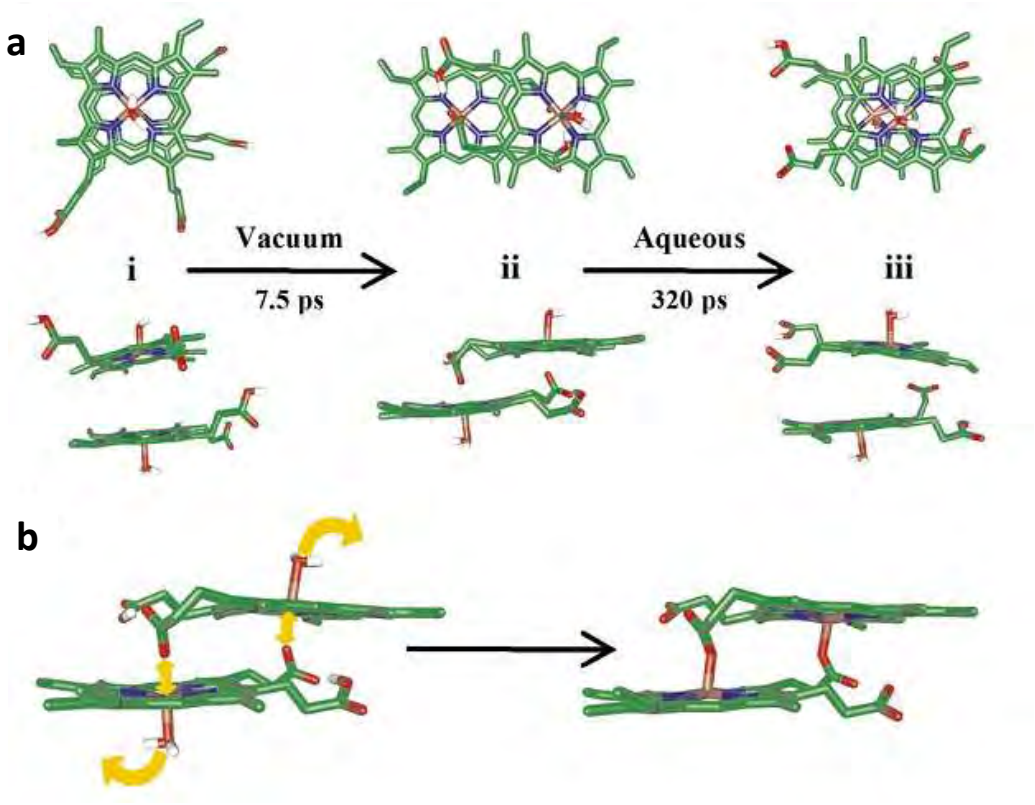


Figure 1.11. MD simulations of π-π dimeric ferrihaem in vacuum and aqueous solution. (**a**) The initial starting structure of the π-π dimer in vacuum (i), rapidly forms the haemozoin precursor (ii) which is not maintained in aqueous solution (iii). (**b**) Proposed mechanism of haemozoin formation from the precursor. Figure reproduced with permission Egan *et al*.[89] © 2006, Elsevier.

1.3 Antimalarial Drugs and Resistance

1.3.1 Antimalarial drug overview.

Antimalarial drugs have been shown to act on different stages of the malaria parasite life cycle but the majority are active in the blood stage. These drugs can be split into five main classes namely (i) folate antagonists such as sulfadoxine and pyrimethamine; (ii) naphthoquinones such as atovaquone; (iii) endoperoxides such as artemisinin; (iv) 4-aminoquinolines such as chloroquine (CQ) and

amodiaquine; and (v) aryl-methanols including compounds such as quinine and mefloquine.[95, 96] Structures of the antimalarial drugs listed in (i) – (v) are shown in Fig. 1.12. The first class represents nucleic acid synthesis inhibitors which target pathways involved in the pyrimidine biosynthetic pathway.[95-97] The second class targets electron transport in the mitochondrion. This class causes a collapse in the organellar membrane potential, preventing the regeneration of ubiquinone, a necessary electron acceptor for dihydroorotate dehydrogenase which is the enzyme crucial for transferring electrons to coenzyme Q.[96, 97] The mechanism of action of the third drug class is less certain. Hypotheses proposed regarding this include that they are calcium-ATPase inhibitors; that they induce oxidative stress through formation of radicals upon cleavage of their peroxide bridges; or possibly act in both manners.[95, 98] The final two classes of compounds target the haem detoxification process (see Section 1.2.2).[99-101] The latter are of particular interest since some of the most commonly used antimalarial drugs belong to this class.

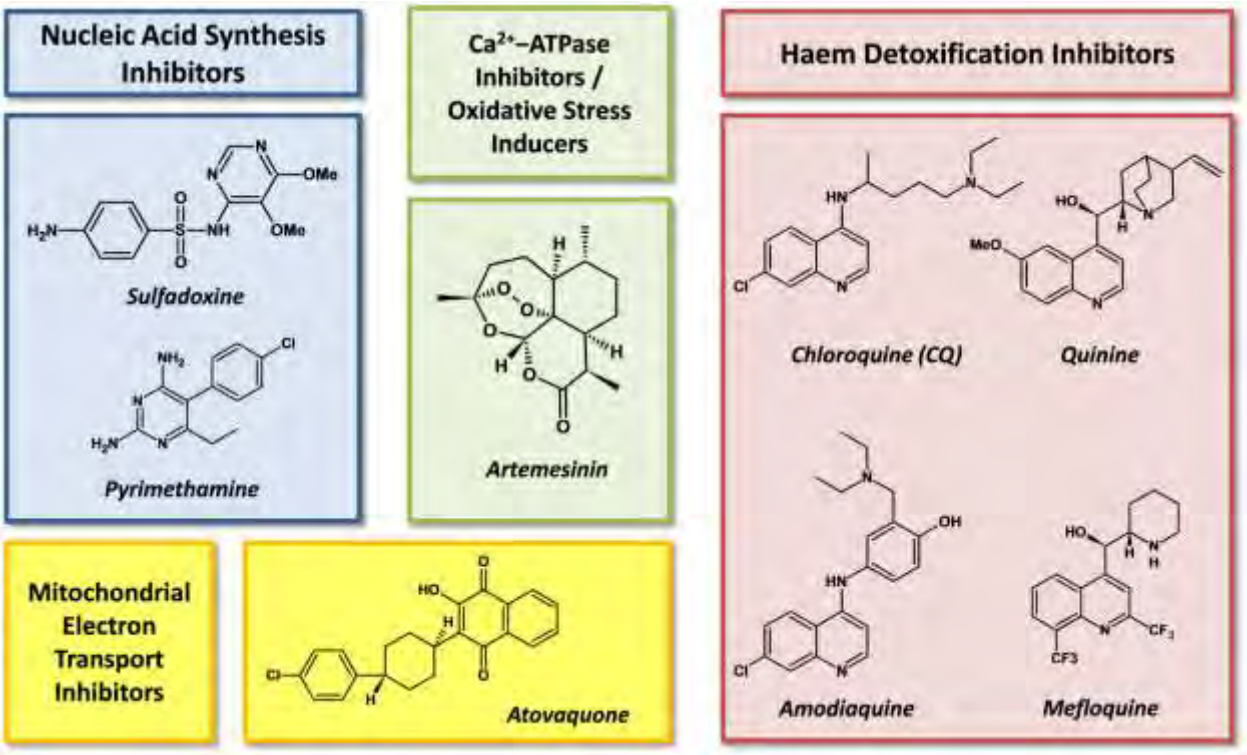


Figure 1.12. Structures of common antimalarial drugs targeting the blood stage of the parasite life cycle.

1.3.2 Quinoline antimalarial drugs.

As early as the 1600s, quinoline antimalarials were used to treat malaria. The first treatment was administered in the form of dried and powdered bark from the cinchona tree, the active component of which is the cinchona alkaloid quinine (see Fig. 1.12 above).[102] Treatment of malaria continued in this fashion until 1820, after which quinine was extracted and isolated from the bark.[103] It was subsequently administered in the purified state and historically remains the first successful treatment of an infectious disease using a chemical compound.[102] While this compound as well as the related cinchona alkaloids quinidine, cinchonine and cinchonidine had cure rates of >98%, they were limited by supply because large-scale production was only possible through extraction from cinchona tree bark.[102] Nevertheless, these compounds, particularly quinine, were the mainstay of malaria treatment until the mid-1920s. Subsequently, synthetic molecules were developed which could provide large-scale production of antimalarials not limited by supply or location of natural resources.[104] The first synthetic compound successfully used to treat malaria was the parasite staining reagent, methylene blue, in 1891 although it was not effective enough to supplant quinine. A number of synthetic modifications were subsequently made to methylene blue in an attempt to increase its activity which resulted in the development of quinacrine in 1931 and CQ in 1935 (see Fig. 1.13 for structures).[104]

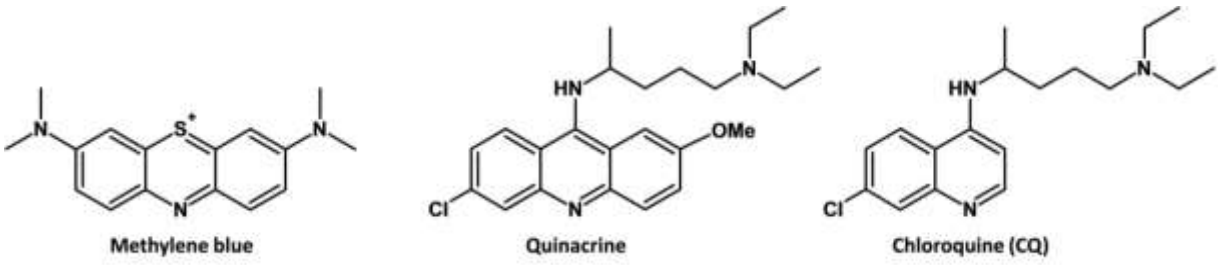


Figure 1.13. Precursors to the antimalarial drug CQ.

Following the discovery of CQ, it became the first-line of antimalarial therapy and is still regarded as the archetypal antimalarial drug owing to its effectiveness in curing malaria, its low cost, ease of synthesis, relatively few side effects and safety for use in pregnant women.[105] While the target of this compound is widely considered to be the inhibition of the haem detoxification process, several others have proposed alternate modes of action. These include (i) intercalation into GC-rich DNA;[106-108]

(ii) inhibition of polyamine metabolism by binding ornithine decarboxylase;[109, 110] (iii) inhibition of haem-dependent protein synthesis;[111] (iv) inhibition of vacuolar phospholipase;[112, 113] (v) inhibition of haemoglobin proteases;[114] (vi) inhibition of haem degradation by hydrogen peroxide;[43] and (vii) inhibition of haem degradation by glutathione in the cytosol.[115-117] Several of these mechanisms, (i) – (iii) are unlikely, since the CQ concentrations in the parasitic cytosol are in the nanomolar range and are thus too low to exert the proposed activity.[118] On the other hand, accumulation of CQ occurs in the DV and is predicted to reach millimolar concentrations, making this the likely site of action.[119-121] Furthermore, numerous studies have provided *in vitro* evidence that CQ is able to inhibit the formation of the synthetic form of haemozoin, β-haematin.[81, 121-127] Strong support of CQ targeting this pathway *in vivo* has also very recently been reported, where Combrinck *et al*.[128] were able to show a dose-dependent increase in free ferrihaem concentrations (see Fig. 1.14a) which correlated closely with decreased parasite growth crossing over close to the concentration required to inhibit 50% of parasite growth (IC_{50}, see Fig. 1.14c). A corresponding decrease in haemozoin concentrations was also observed (see Fig. 1.14b).

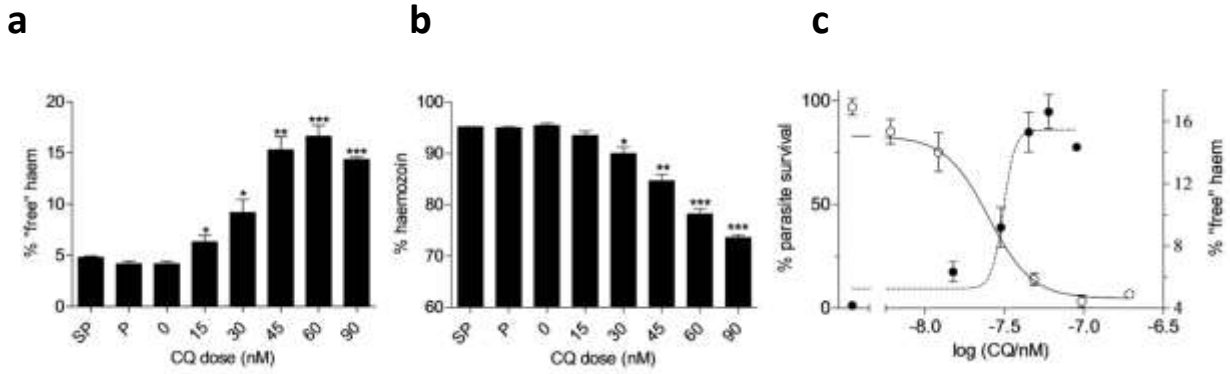


Figure 1.14. (**a**) Increase in free haem concentrations in *P. falciparum* as a function of CQ concentration. (**b**) Decrease in haemozoin concentration in *P. falciparum* as a function of CQ concentration. (**c**) Parasite survival (plotted on left axis) and free ferrihaem concentration (plotted on right axis) as a function of CQ concentration. Curves intersect near the measured IC_{50} of CQ providing *in vivo* evidence that ferrihaem detoxification is the target of this drug. Figure adapted with permission from Combrinck *et al*.[128] © 2012, American Chemical Society.

While the application of CQ from the 1940s resulted in the near eradication of malaria, its heavy use resulted in the emergence of CQ-resistant (CQR) parasites in the 1950s. By the 1980s, CQ was all but ineffective for antimalarial therapy owing to widespread parasite resistance and malaria was resurgent.[102, 129] Combination therapy is currently employed to treat malaria, which involves the use of two antimalarial drugs with different modes of action. These include artemisinin or one of its derivatives and usually a quinoline compound such as mefloquine, amodiaquine, piperaquine (Euartesim®, MMV, Switzerland) or lumefantrine (Coartem®, Novartis, Switzerland).[129] The combination approach is used in order to limit the development of parasite resistance because for this to occur, the parasite must produce advantageous mutations for two distinct and separate pathways. While this method has proved successful, there have been recent reports of the first signs of parasite resistance to combination therapy, including prolonged clearance times resulting in artemisinin treatment failure.[130-133] Thus there is a need to develop new and improved antimalarial drugs to combat malaria. Consequently, much research is still invested into understanding the mechanism of action of CQ and producing CQ-analogues, despite the extensive resistance to this drug. Indeed, two related 4-aminoquinoline compounds, piperaquine and ferroquine, have recently entered the clinic or are currently undergoing clinical trials.[134, 135] Investigation and development of CQ-like compounds as antimalarial drugs continues because the mechanism of CQ resistance has been shown to be independent of its target ferrihaem detoxification process.[129]

1.3.2 Mechanism of CQ resistance.

Resistance to CQ is not caused by any known change to the haem detoxification process. Rather, it arises from the reduction of CQ accumulation in the DV of CQR parasites compared to CQ-sensitive (CQS) ones.[136, 137] Initially, it was proposed that the reduced concentration of CQ in the DV resulted from increased vacuolar pH.[138] This was suggested based on the pH trapping effect of CQ which is known to cause accumulation in the DV and occurs when neutral CQ present in the parasitic cytosol (approximate pH of 7.4) becomes diprotic after crossing the membrane into the DV (approximate pH of 5). Owing to the double positive charge, CQ is not able to diffuse back across the membrane into the cytosol and thus accumulates in the DV. The proposal of increased vacuolar pH would result in a larger percentage of neutral CQ existing in the DV, thus allowing more to diffuse into back across the membrane with the

result of reduced accumulation in the DV. This hypothesis, however, was later discounted by a study which showed the vacuolar pH of CQR and CQS parasites to be similar.[46]

Subsequently, several other mechanisms of CQ resistance have been proposed but the principle determinant of CQ resistance is widely accepted to involve mutations in the *P. falciparum* CQ resistance transporter (PfCRT) protein.[129, 139] PfCRT is a transmembrane protein which localises in the DV membrane and is a member of the drug metabolite transporter family.[140] The predicted arrangement of the protein is depicted in Fig. 1.15. A second protein, *P. falciparum* multi-drug resistance 1 transporter (PfMDR1) has also been implicated in CQ resistance, although is thought to modulate sensitivity to CQ in the presence of PfCRT mutants rather than being directly responsible for it.[141] A recent study has also suggested that CQ inhibits the transport of natural substrates of PfCRT and PfMDR1 and CQ binding to PfCRT has been observed using a perfluorophenyl azido CQ analogue.[142, 143] Strong evidence to support the proposed mechanism of PfCRT in facilitating CQ resistance was reported by Martin *et al*.[144], where PfCRT expressed in the membrane of *Xenopus laevis* oocytes was shown to transport CQ across the membrane. Indeed, this finding supports the idea that decreased CQ activity stems from the efflux of CQ from the DV, effectively lowering the concentrations and allowing the ferrihaem detoxification process to proceed unhindered. Further support for transport of CQ by PfCRT was shown when the known membrane transport inhibitor, verapamil, was included in this system which resulted in the transport of CQ being abated. This finding mirrors those observed *in vivo* where CQR parasites were found to become CQS once more when incubated with CQ and verapamil.[145] Mutation of lysine at position 76 to threonine (K76T) in PfCRT is crucial to confer CQ resistance, although several other mutations are also present which vary from strain to strain.[139, 146] These are believed to be secondary mutations that counteract the loss of function of PfCRT as a result of the K76T mutation.[139] A summary of reported mutations associated with CQ resistance is shown in Fig. 1.15. In wild-type PfCRT, it is thought that the presence of the positively-charged lysine at position 76 near the entrance to the transmembrane channel on the DV side is responsible for cationic repulsion of diprotic CQ which prevents its transport by the protein.[147] The crucial K76T mutation results in the loss of the positive charge and thus there are no longer unfavourable interactions preventing CQ from being transported across the membrane. This is represented schematically in Fig. 1.16.

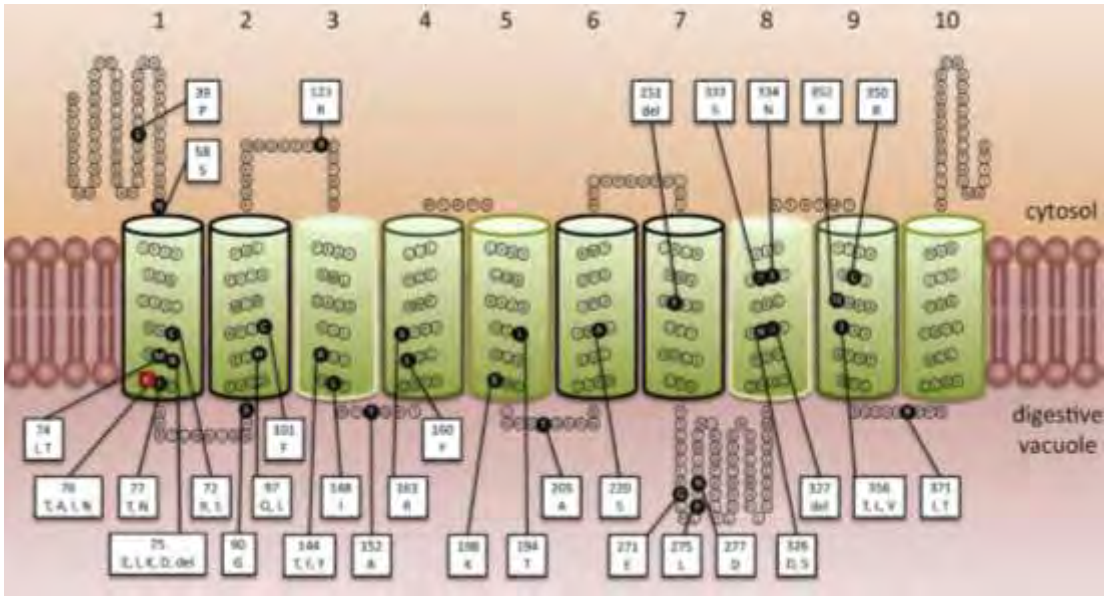


Figure 1.15. Predicted arrangement of wild-type PfCRT in the parasite DV membrane, consisting of ten α-helical transmembrane domains. Reported mutations in CQR parasites are indicated in white boxes. The crucial K76T mutation is highlighted in red. Figure reproduced with permission from Summers *et al.*[139] © 2012, Springer Basel AG.

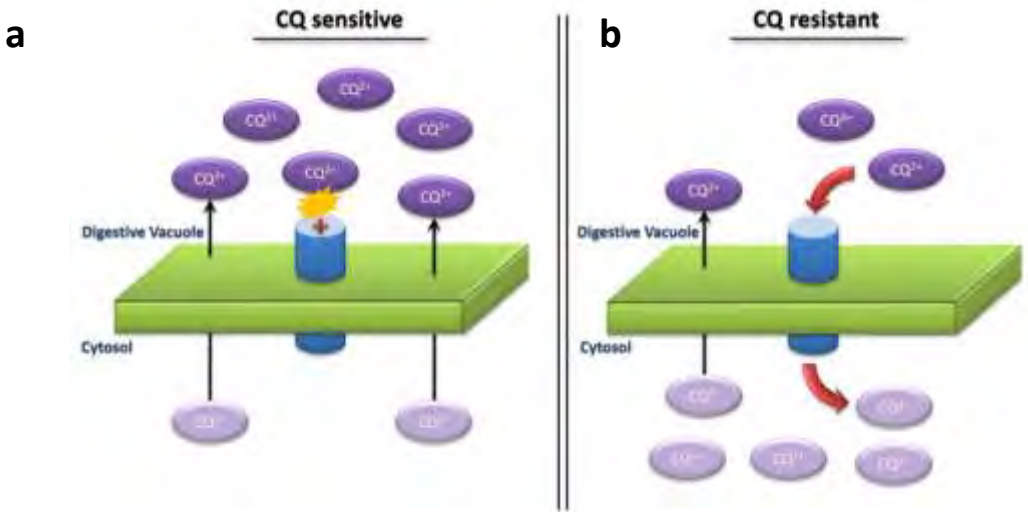


Figure 1.16. Proposed interaction of CQ with PfCRT in (**a**) wild-type CQS and (**b**) CQR parasites. In (a), pH trapping of CQ across the membrane (green) into the DV causes accumulation of CQ. The positive charge on lysine at position 76 of PfCRT (blue cylinder) prevents transport out of the DV. In (b), the K76T mutation removes the postive charge at the entrance of PfCRT and facilitates transport of CQ across the membrane into the parasitic cytsol. This efflux reduces the concentration of CQ in the DV allowing haemozoin formation to continue unabated.

1.4 Mechanism of Haemozoin Inhibition by Quinoline Antimalarial Drugs

While the quinoline class of antimalarial drugs have been one of the most successful and widely-used in antimalarial therapy, the mechanism of interference with haemozoin formation is not completely understood. There are two prevailing theories, however, which have enjoyed much attention in the literature. These include the inhibition of the fastest growing face of the haemozoin crystal[53, 148] and complex formation between antimalarial drugs and free ferrihaem.[99, 122, 125] These are discussed in Sections 1.4.1 and 1.4.2 respectively.

1.4.1 Interference with haemozoin crystal formation.

Pagola *et al*.[53] were the first propose that CQ adsorbs onto the fastest growing (001) and $(00\bar{1})$ faces of haemozoin crystals which are orientated in such a manner as to form grooves that expose the unligated porphyrin face and propionic acid side chains of ferrihaem molecules (see Fig. 1.17a). This hypothesis was made based on the surface area of haemozoin crystals and the amount of CQ predicted to be in the parasite DV. The structure of the adsorbed CQ-haemozoin complex was subsequently proposed by Buller *et al*.[148] in which four specific intermolecular interactions between monoprotic CQ and ferrihaem molecules were suggested to stabilise the complex. These included (i) a salt bridge between propionate COO^- of ferrihaem and NH^+ of the CQ side chain; (ii) a contact between methyl protons of ferrihaem and the 7-chloro substituent of CQ; (iii) an interaction between the vinyl proton on the ferrihaem porphyrin and quinoline nitrogen of CQ; and (iv) an interaction between the proton of the 4-amino substituent of CQ and the π-cloud of a neighbouring vinyl group of ferrihaem. Proposed interactions are depicted in Fig. 1.17b.

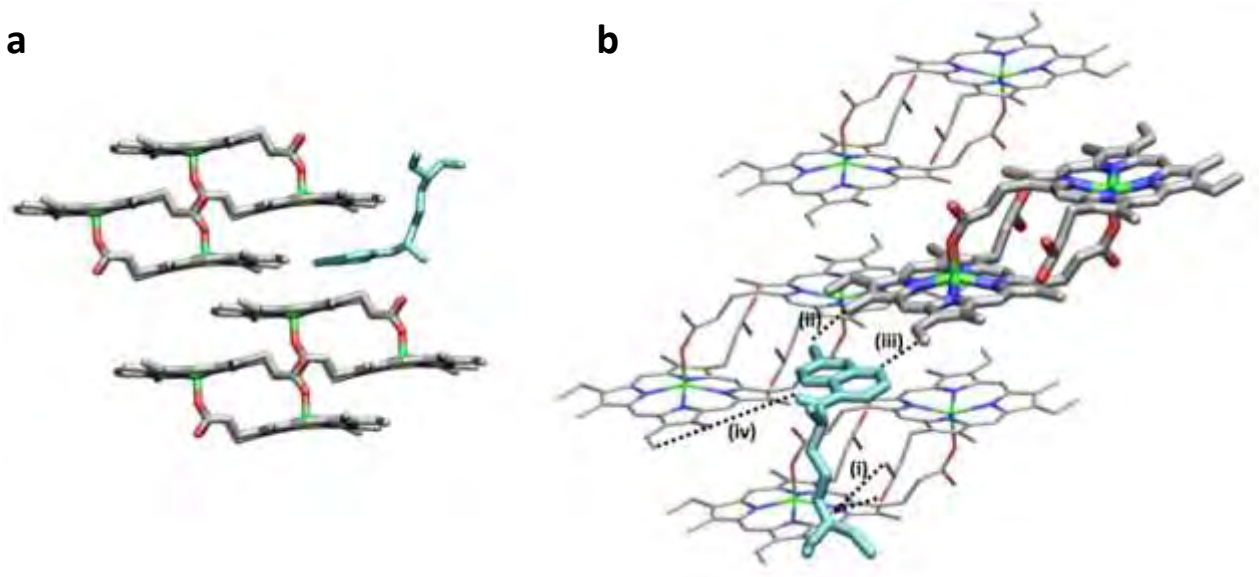


Figure 1.17. Proposed structure of CQ docked with the fastest growing face of haemozoin viewed (**a**) along the a-axis. (**b**) Intermolecular contacts stabilising the CQ-haemozoin docked complex. Details of interactions are discussed in the text. Porphyrin hydrogen atoms and selected hydrogen atoms of CQ not shown for clarity. Monoprotic CQ has been coloured in cyan to aid visualisation. Figures were created using the crystal structures of β-haematin and CQ reported by Pagola *et al*.[53] and Karle[149], respectively, and were obtained from the CSD.[150]

A criticism of the structural model employed by Buller *et al*. was the use of the monoprotic form of CQ which is likely to only be present in small quantities given the pH of the DV. On the other hand, a recent study by Nsumiwa *et al*.[151] tends to support the implication of the monoprotic species being involved in haemozoin inhibition. In this study, a significant correlation between synthetic haemozoin inhibition and QM-derived atomic charges for a range of 4-aminoquinoline derivatives was only obtained when compounds were computed in the free base form. Furthermore, correlations with atomic charges were observed for atoms on the quinoline ring implicated or near those implicated in the Buller *et al*. structural model (see Fig. 1.18a). Extended X-ray absorption fine-structure (EXAFS) and UV-visible spectroscopic studies by Bohle and co-workers have also provided some support for the interaction of CQ and μ-propionato ferrihaem, although these authors used the more soluble mesoporphyrin analogue to conduct investigations.[152, 153] UV-visible spectrophotometric titrations in organic media indicated CQ bound strongly to the μ-propionato mesoporphyrin species ($\log K_{obs} \approx 5.6$).[153] EXAFS spectra of the CQ-(μ-propionato mesoporphyrin) complex in DMSO suggested that CQ atoms are positioned between 2.48 and 3.77 Å from an iron centre (see Fig. 1.18b).[152] Recent kinetics evidence provided by Gildenhuys *et al*.[154] supports the adsorption of CQ as well as the cinchona alkaloid quinidine

onto the surface of haemozoin crystals. These authors found that the decreased rate of formation of synthetic haemozoin in the presence of low drug concentrations could only be modelled using a drug adsorption process. At higher drug concentrations, precipitation of drug-ferrihaem complexes caused a reduction in the overall synthetic haemozoin yield. This kinetic model (see Fig. 1.18c) enabled calculation of equilibrium adsorption constants (K_{ads}) for CQ and quinidine (log K_{ads} = 5.55 ± 0.03 and 4.92 ± 0.01, respectively).

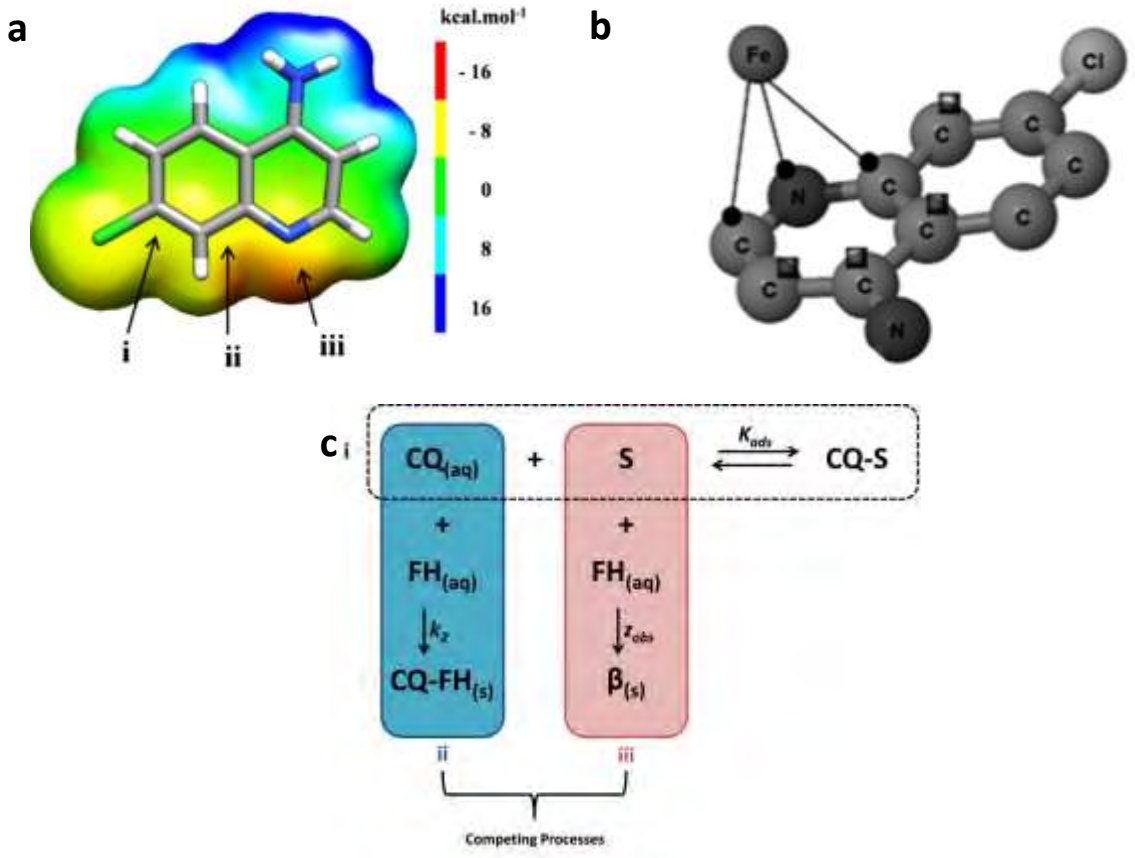


Figure 1.18. (**a**) Aromatic moiety of CQ showing partial charges correlating with synthetic haemozoin inhibition reported by Nsumiwa *et al*.[151] Decreased positive potential at C7 (i) and decreased negative potential at C8 (ii) and N1 (iii) favour inhibition of haemozoin formation. (**b**) Model of CQ quinoline atoms in relation to the iron centre of μ-propionato mesoporphyrin as reported by Walczak *et al*.[152] Figure adapted with permission from Walczak *et al*.[152] © 2011, American Chemical Society. (**c**) Kinetic model used to describe CQ adsorption to the surface of synthetic haemozoin crystals. (i) S, surface binding site of haemozoin; CQ-S, CQ adsorbed to haemozoin surface. (ii) Precipitation of the drug-ferrihaem complex, CQ-$FH_{(s)}$; FH, ferrihaem. (iii) Formation of synthetic haemozoin, $\beta_{(s)}$. K_{ads}, k_2 and z_{obs} are the equilibrium adsorption and rate constants for drug-ferrihaem complex and synthetic haemozoin formation, respectively.

1.4.2 Ferrihaem association.

The second proposed mechanism of quinoline antimalarial activity is that of drug-ferrihaem complex formation in aqueous solution. These complexes have been suggested to bind ferrihaem in a form which no longer allows its incorporation into haemozoin. Moreover, the complex has been suggested to be more soluble than ferrihaem itself thus further discouraging its biomineralisation into haemozoin.[155] Strong support for this mode of action has been reported for the cinchona alkaloids, quinine and quinidine, as well as for the related aryl-methanol antimalarial, halofantrine, where crystal structures of the antimalarial drug-ferrihaem complexes have been elucidated (see Fig. 1.19).[78, 156] Three specific interactions were identified, including (i) coordination of the drug alkoxide moiety to iron centre of ferrihaem; (ii) π-stacking interaction between aromatic ring systems of the drug and porphyrin of ferrihaem; and (iii) hydrogen bonding to propionate side chains on the same porphyrin in the case of quinine and quinidine, or on a neighbouring complex in the case of halofantrine. Furthermore, molecular mechanics (MM) modelling showed the energy required to maintain the hydrogen bonded conformation of quinine and quinidine as well as their inactive epimers, epiquinine and epiquinidine, was directly correlated with their corresponding parasite IC_{50} values.[78] This suggested that the comparative ease of forming hydrogen bonds with the propionate side chain of ferrihaem was important for antimalarial activity.

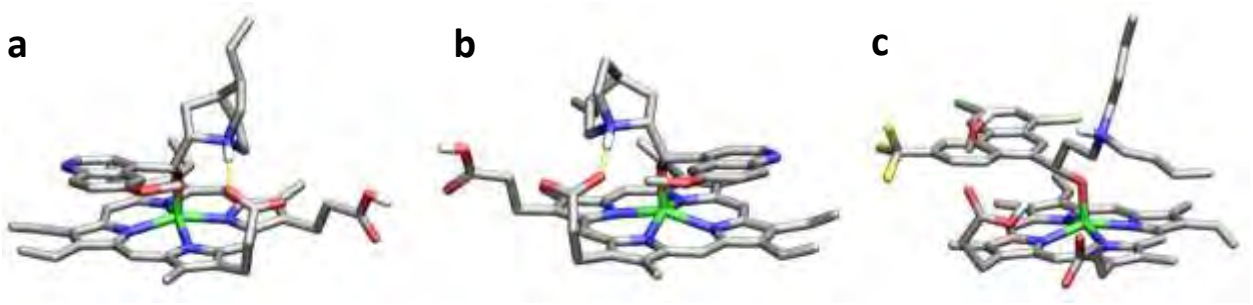


Figure 1.19. Crystal structures of (**a**) quinine-, (**b**) quinidine- and (**c**) halofantrine-ferrihaem complexes reported by de Villiers and co-workers.[78, 156] Intramolecular hydrogen bond between propionate side chain and quinuclidine NH in (**a**) and (**b**) is shown in yellow. Structures reproduced using crystal structure data obtained from the CSD.

While the structural evidence for these antimalarial drug-ferrihaem complexes is undeniable, there is debate as to whether they exist in the same form in aqueous solution given that the all three were crystallised from predominately organic solvent systems. Alumasa *et al.*[157] have suggested an alternate solution-phase structure for cinchona alkaloid-ferrihaem complexes in which association is driven by π-stacking interactions between the quinoline ring and porphyrin system, which is supplemented by an electrostatic interaction between the aryl-methanol moiety and iron centre of ferrihaem. Furthermore, this solution-structure was proposed to have an intramolecular hydrogen bond forming a five-atom ring between the aryl-methanol and quinuclidine nitrogen (see Fig, 1.20a). This model is not without criticism either, however, as it requires the cinchona alkaloid to be in the neutral state which is unlikely given the pK_a of the quinuclidine nitrogen atom (approx. 8.6) compared to pH of the DV (approx. 5). Furthermore, the aryl-methanol antimalarial-drug complexes have been proposed to form in the organic-like environment of the lipid nanospheres within the DV where haemozoin formation has been suggested to occur.[78] This is depicted schematically in Fig. 1.20b.

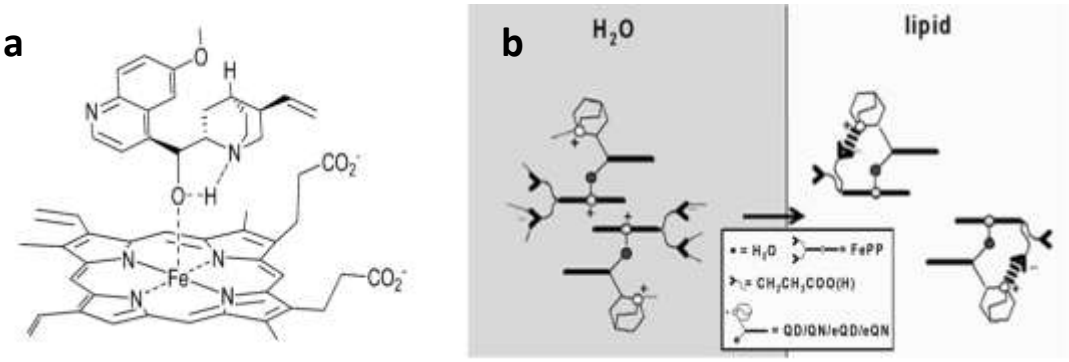


Figure 1.20. (**a**) Solution structure of the quinine-ferrihaem complex proposed by Alumasa *et al.*[157] Figure adapted with permission from Gorka *et al.*[77] © 2013, American Chemical Society. (**b**) Structure of the cinchona alkaloid-ferrihaem complex in solution and lipid media suggested by de Villiers *et al.*[78] Figure adapted with permission from de Villiers *et al.*[78]© 2008, Elsevier.

Regardless of whether the mechanism of action of these antimalarial drugs involves complex formation in solution or inhibition of the fastest growing face of haemozoin crystals, investigation of the solution interaction between antimalarials and ferrihaem is still of interest. Increased concentrations of ferrihaem present in the parasite after inhibition of haemozoin very likely associate with antimalarial drugs which are present in the DV. Thus the subsequent toxicity of ferrihaem may well be a result of the ferrihaem-antimalarial complex.

In contrast to the complexes of ferrihaem and aryl-methanol antimalarial drugs, there is much confusion regarding the structures of CQ-ferrihaem and other 4-aminoquinoline-ferrihaem complexes which are less well understood. This stems from previous investigations being conducted without fully appreciating the importance of experimental conditions on ferrihaem speciation as well as with the frequent reliance on a single technique to obtain experimental information regarding the CQ-ferrihaem complex. A brief discussion of these investigations as well as the conflicting findings is presented Section 1.5.

1.5 Investigations of the CQ-Ferrihaem Complex

The first reported solution interaction between CQ and ferrihaem was described by Cohen *et al*.[158] in 1964, where changes to the UV-visible spectrum of CQ in aqueous solution were observed upon addition of ferrihaem. Subsequent work by Chou *et al*.[99] in 1980 produced the first equilibrium association constant between CQ and ferrihaem in aqueous solution ($\log K_{obs}$ = 8.5) measured using equilibrium dialysis. Furthermore, these authors reported a binding stoichiometry of 1:2 CQ:ferrihaem. Following these two seminal studies, numerous investigations of the CQ-ferrihaem interaction have been reported employing a variety of experimental as well as computational techniques.

Dorn *et al*.[125] were one of the first to use isothermal titration calorimetry (ITC) to determine binding stoichiometries and association constants as well as to probe the thermodynamics of complex formation in the form of entropic and enthalpic contributions. While this study, performed in aqueous solution (pH 6.5) at 37°C, indicated complex formation was driven by enthalpy and resulted in an unfavourable decrease in entropy, binding stoichiometries obtained (1:4 CQ:ferrihaem) contradicted that reported by Chou *et al*.[99] On the other hand, the same findings reported by Dorn *et al*. were obtained by Cheruku *et al*.[159] using ITC measurements for CQ and its isostere in which the 4-amino substituent of the quinoline ring was replaced with a CH group. Both authors rationalised the 1:4 binding ratio as one with CQ sandwiched between two μ-oxo dimer units. A thorough study of the effects of temperature, salt concentration and pH by Bacchawat *et al*.[160], however, produced yet more contrasting findings. Using a model involving two sets of interacting sites, these authors showed that at pH 5.6 over

the temperature range 280.5 to 300.5 K, complex formation almost always had a positive entropy component. By contrast, titrations performed at pH 7.4 had negative entropy and were more in line with the previous two studies. Interestingly, addition of salt to the system abolished the first binding site at pH 5.6 but not at pH 7.4. The findings reported by Bacchawat *et al*. have provided some insight into the thermodynamics governing CQ-ferrihaem complex formation, although should be treated with caution since titrations were conducted using a solvent system containing 0.04% of the detergent, Triton X. This substance was used to apparently render ferrihaem monomeric in solution and, although they mentioned that monomeric ferrihaem was qualitatively determined, no direct evidence was provided to support this. Furthermore, reported binding stoichiometries in some cases were up to six ferrihaem molecules per CQ which is difficult to rationalise. In a similar study by Vippagunta *et al*.[161], the effects of pH and temperature on CQ-ferrihaem association were determined using ITC measurements. Little difference in association strength was observed over the pH range 5.6 – 9.0 where the average $\log K_{obs}$ was 5.6 ± 0.1. Association at all pH values was enthalpy-driven and had negative entropy components, with complex formation at pH 9 having a negative value significantly greater than that observed at all other pH values. Association strengths at pH 6.5 were found to be independent of ionic strength up to 0.5 M salt concentration, after which somewhat larger constants were obtained. The negative entropy component reduced in magnitude with increasing ionic strength becoming positive at 0.75 M salt concentration. This study, however, also reported varying binding stoichiometries from 1:3 to 1:7 CQ:ferrihaem.

UV-visible spectroscopy has likewise been extensively used to probe interactions of CQ with ferrihaem. Egan and co-workers have measured association constants of a number of antimalarials, including CQ, with ferrihaem in 40% (v/v) aq. DMSO, pH 7.4.[124, 162] This solvent system was shown to form monomeric ferrihaem and was used to the circumvent issues of ferrihaem aggregation and changes in speciation. These authors also showed that increases in solvent polarity resulted in increased CQ-ferrihaem association strength.[162] In contrast to the findings obtained from ITC measurements at the same pH, thermodynamics determined from van't Hoff plots indicated formation of the CQ-ferrihaem complex produced a positive entropy component. Interestingly, despite the differences observed in the thermodynamic properties, association constants determined using the two techniques were similar ($\log K_{obs} \approx 5.5$), although aren't strictly comparable given the different binding stoichiometries used to fit the titration data. In the case of Egan *et al*., a 1:1 binding model was employed. More recently, Crespo *et*

al.[163] reported a detailed study on the interaction of CQ with ferrihaem in 40% aq. methanol over a wide pH range. This solvent system was chosen to ensure the biologically relevant π-π ferrihaem dimer species was produced but also to facilitate UV-visible titrations at low pH which weren't possible in pure aqueous solution owing to ferrihaem insolubility. These authors observed the same 1:2 CQ:ferrihaem binding ratio as reported by Chou *et al*.[99] and found a bell-shaped pH dependence of CQ-ferrihaem association constant which peaked around pH 6 – 7 ($\log K_{obs} \approx 6$). Association constants at lower and higher pH, however, still produced moderately strong values ($\log K_{obs} \approx 3$). Earlier work by Constantinidis and Satterlee[164] as well as by Marques *et al*.[165] utilised urohaemin I (Fig. 1.21a) and the haem octapeptide *N*-acetylmicroperoxidase-8 (N-AcMP8, Fig. 1.21b), respectively, as model systems for ferrihaem in order to avoid complicated solution speciation and aggregation. While both ferrihaem models were able to dimerise, experimental conditions under which UV-visible titrations were performed ensured monomeric species were produced. Findings obtained using N-AcMP8 were inconsistent with the data determined using ferrihaem itself, where a binding stoichiometry of 2:1 CQ:N-AcMP8 and comparatively weak association constants were observed ($\log K_{obs,1} \approx 3$ and $\log K_{obs,2} \approx 1.5$). On the other hand, the same binding stoichiometry was obtained from CQ-urohaemin investigations as that reported by Chou *et al*. and Crespo *et al*., but a much larger association constant was observed ($\log K_{obs} \approx 9$).

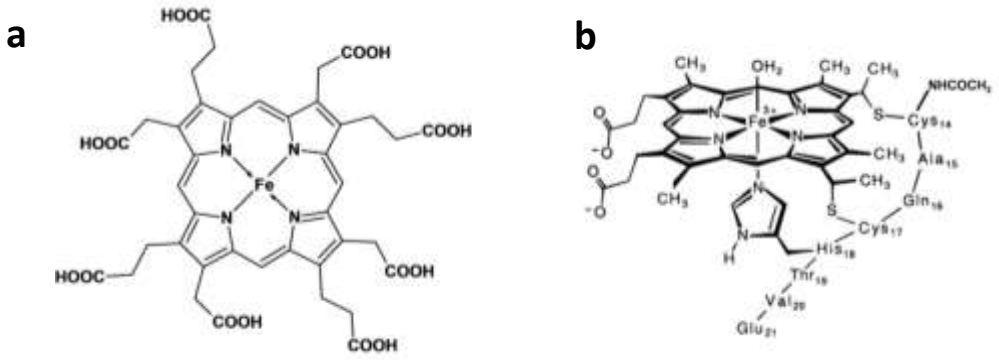


Figure 1.21. Metalloporphyrins used as ferrihaem models. (**a**) (Urohaemin I)Fe(III) chloride used by Constantinidis and Satterlee[164], axial chloride not shown for clarity. (**b**) Ferrihaem *N*-acetylmicroperoxidase-8, N-AcMP8, used by Marques *et al*.[165] Figure adapted with permission from Munro *et al*.[166] © 1996, American Chemical Society.

While ITC and UV-visible methods have provided insight into the thermodynamics of complex formation, other techniques such as NMR spectrometry, Mössbauer spectroscopy and computational methods have been employed to investigate the ferrihaem species present in the complex as well as to attempt to determine the structure of the complex. The earliest proposed structures of the CQ-ferrihaem complex in solution were reported by Moreau *et al*.[60, 167] Initially, a qualitative model was proposed based on observed shifts in the ^{1}H NMR spectrum of CQ as a result of increasing concentration of ferrihaem.[60] Later, the ^{1}H NMR shifts of CQ in the presence of increasing concentrations of uroporphyrin were used in conjunction with a mathematical ring current model to proposed a structure which consisted of co-facial π-stacking interaction between the aromatic moiety of CQ and porphyrin ring system of ferrihaem (see Fig. 1.22a).[167] A similar structure was later proposed by Constantinidis and Satterlee[164] based on ^{13}C NMR shifts using urohaemin, although these authors suggested CQ as being situated between two urohaemin molecules while Moreau *et al*. assumed μ-oxo ferrihaem was the dominant species in aqueous solution and placed CQ between two of these dimers. Interestingly, however, despite later evidence showing the π-π dimer being the predominant species in aqueous solution, Casabianca *et al*.[168, 169] reported clear evidence that ferrihaem existed in the μ-oxo dimeric form in the complex in aq. DMSO solution (pH 7.4) as well as in detergent micelles. In contrast to the μ-oxo dimeric form of ferrihaem, Schwedhelm and co-workers[170] suggested a ferrihaem tetramer adduct formed in aqueous solution based on mass spectrometry studies of the CQ-ferrihaem complex. This structure involved two μ-oxo ferrihaem species reciprocally coordinated by propionate side chains to iron centres in a similar fashion to haemozoin (see Fig. 1.22b). However, the structure was proposed in order to account for the low magnetic moment observed which they erroneously attributed to a low-spin S = ½ species. Rather, the low magnetic moment is indicative of the μ-oxo dimeric species and thus the formation of this ferrihaem species is not needed to explain the magnetic properties of the complex. In fact, this complex is highly unlikely because the characteristic C=O and C-O stretching bands of coordinated propionates are not seen in the IR spectrum of the complex. Resonance Raman (RR) studies by Webster *et al*.[171] suggested that CQ interacts with a variety of ferrihaem species, including the μ-oxo dimer, although these authors suggested interaction occurs in a 1:1 ratio (CQ:ferrihaem, see Fig. 1.22c).

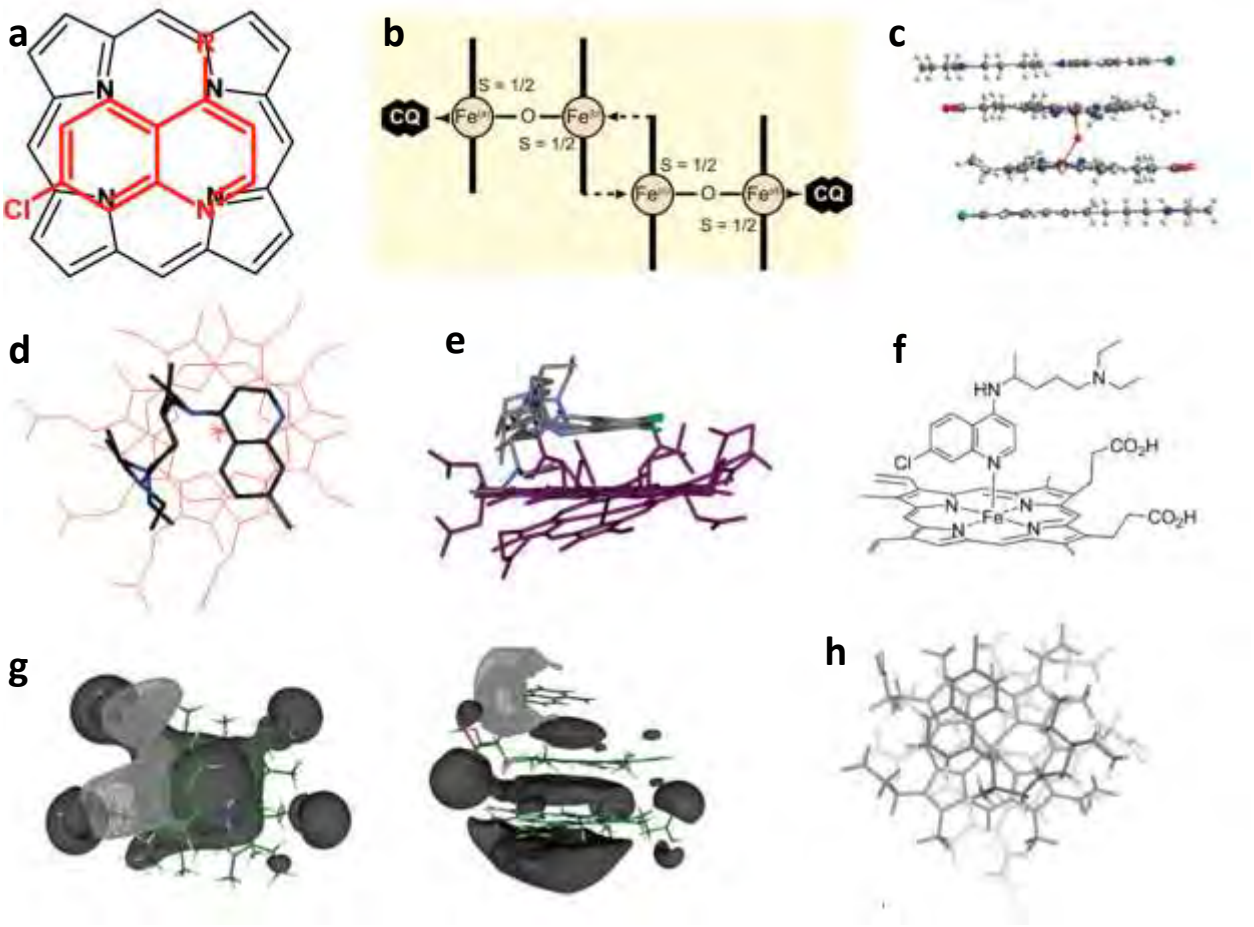


Figure 1.22. (**a**) CQ-urohaemin structure proposed by Moreau *et al*.[167] and Constantinidis and Satterlee[164]. Porphyrin substituents and iron centre as well as CQ (red) side chain not shown for clarity. (**b**) Schematic representation of the tetramer adduct proposed by Schwedhelm *et al*.[170] Figure adapted with permission from Schwedhelm *et al*.[170] © 2008, Bentham Science Publishers. (**c**) CQ-(μ-oxo ferrihaem) complex proposed by Webster *et al*.[171] on the basis of resonance Raman data. Figure adapted with permission from Webster *et al*.[171] © 2009, American Chemical Society. (**d**) Lowest-energy structure of the diprotic CQ-(μ-oxo ferrihaem) complex reported by Leed *et al*.[172] For clarity, only the upper porphyrin of the μ-oxo dimer is shown. Figure adapted with permission from Leed *et al*.[172] © 2002, American Chemical Society. (**e**) Lowest-energy structures of CQ-ferrihaem complex reported by Natarajan *et al*.[173] Figure adapted with permission from Natarajan *et al*.[173] © 2008, American Chemical Society. (**f**) Dative monomeric CQ-ferrihaem complex in equilibrium with the CQ-μ-oxo ferrihaem complex proposed by de Dios *et al*.[174] Figure adapted with permission from Gorka *et al*.[77] © 2013, American Chemical Society. (**g**) Structures proposed by Portela *et al*.[175] based on complimentary ESPs of μ-oxo ferrihaem and the free-base quinoline moiety of CQ. Figure adapted with permission from Portela *et al*.[175] © 2004, Elsevier. (**h**) Lowest-energy docked structure of the CQ-ferrihaem complex reported by Portela *et al*.[176] Figure adapted with permission from Portela *et al*.[176] © 2003, Elsevier.

Computational investigation of the CQ-ferrihaem complex was undertaken by Leed *et al*.[172] using MD modelling in which CQ was π-stacked with the unligated face of μ-oxo ferrihaem. Distance constraints obtained from ^{1}H NMR T_1 relaxation data were also applied in order to provide representative solution structures. The lowest-energy structure of the complex obtained for the diprotic form of CQ is shown in Fig. 1.22d, but is tentative at best given three major shortfalls of the study. The first criticism is that the incorrect magnetic moment corresponding to a low-spin state (S = ½) was used; the second is the neglect of solvent in the simulation; and the third is the statistical error associated with the calculation of the distance constraints. These authors made the assumption that the average T_1 relaxation time could be used to calculate average Fe-H distance, *r*, using the Solomon-Bloembergen equation which relates T_1 relaxation times to r^6. However this approach was incorrect since $\langle r \rangle \neq \sqrt[6]{\langle r^6 \rangle}$.[177] A similar study was later undertaken by Natarajan *et al*.[173] in which CQ and a number of its analogues were modelled using the same approach. The lowest-energy structures obtained for the CQ-ferrihaem complex are shown in Fig. 1.22e. Portela *et al.*[175] suggested a structure for the complex based on complimentary electrostatic surface potentials (ESPs) of CQ and μ-oxo ferrihaem computed using density functional theory (DFT) calculations. These were aligned in the π-stacked orientation similar to that observed in the MD simulations by Leed *et al*. and Natarajan *et al*. (see Fig. 1.22g), although the quinoline was orientated towards the propionate groups in the DFT-structure while they were orientated away from them in the MD models. One possible criticism of this study is that only the free-base quinoline form of CQ was used in DFT calculations and ESPs were not generated for the protonated form which likely interacts with ferrihaem. Docking studies were also performed by the same authors with CQ and μ-oxo ferrihaem (see Fig. 1.22h).[176] In the case of docking and MM/MD calculations, a further possible criticism relates to force fields employed to describe ferrihaem. These are generally developed for ferrihaem species in proteins which are most often six-coordinate. Consequently, the force fields used may not correctly describe ferrihaem in aqueous solution since all known crystal structures involving ferrihaem indicate this molecule is, in fact, five-coordinate outside of a protein environment.[53, 72, 78, 154, 156, 178] Five-coordinate ferrihaem in solution has been further supported experimentally by de Villiers *et al*.[75] and Asher *et al*.[76]

Conflicting results have also been reported in the case of the solid-state CQ-ferrihaem complex precipitated from more biologically relevant conditions (pH 5). A monomeric CQ-ferrihaem complex in which CQ coordinates to the iron centre of ferrihaem (see Fig. 1.22f) was proposed by de Dios *et al*.[174]

on the basis of solid-state NMR investigations. These authors proposed this structure existed in equilibrium with the μ-oxo dimer complex (Fig. 1.22d and e). While Dascombe *et al*.[179] showed using MM modelling that the necessary porphyrin distortion of the conformation may be less than expected, this solid-state complex is inconsistent with Mössbauer data reported by Adams *et al*.[180] for the precipitated complex which instead indicated ferrihaem was μ-oxo dimeric. On the other hand, a recent study by Asghari-Khiavi *et al*.[181] suggested the ferrihaem species in the complex was not the μ-oxo dimer in the solid state owing to the lack of Fe-O-Fe stretching frequencies in IR and RR spectra. Webster *et al*.[182] have also failed to observe the antisymmetric Fe-O-Fe stretching frequency using RR in the parasite itself and thus have suggested the CQ-ferrihaem complex does not involve the μ-oxo dimer *in vivo* either.

In light of the absence of a definitive CQ-ferrihaem structure in aqueous solution, investigations have focused on developing structure activity relationships (SARs) of CQ analogues to provide information that can be used for rational drug design. Early work by Vippagunta *et al*.[161] and Egan and co-workers[120, 127] showed that changes to the CQ side chain modulated the association strength with ferrihaem as well as affected β-haematin inhibition and *in vivo P. falciparum* activities. Modifications to the position and type of aromatic ring substituent were found to play a more important role in ferrihaem β-haematin inhibition and *in vivo P. falciparum* activities, while still modulating ferrihaem association strengths. Later Cheruku *et al*.[159] demonstrated that CQ isosteres in which the quinoline nitrogen atom or 4-amino substituent were replaced by carbon atoms had a marked effect on ferrihaem association and β-haematin inhibition. More recently, Nsumiwa *et al*.[151] reported SARs of a wide range of CQ-analogues and provided the first convincing empirical correlation between ferrihaem association and β-haematin inhibition. In essence, all authors came to the same conclusions and proposed that the quinoline moiety was responsible for association with ferrihaem. Indeed, a weak, but measurable association constant between ferrihaem and quinoline has been previously reported by Kuter *et al*.[183] The presence of the 4-amino substituent as well as electronegative groups in the 7-position was shown to be essential for β-haematin inhibition and antiplasmodial activity, but also modulated ferrihaem association. Lastly, side chains were found to modulate *in vivo* activity presumably through pH trapping into the acidic DV, although replacement of the basic tertiary amino group on the side chain with a carbon atom still maintains *in vivo* activity, and does not drastically affect β-haematin inhibition or ferrihaem association strength. These findings are summarised in Fig. 1.23.

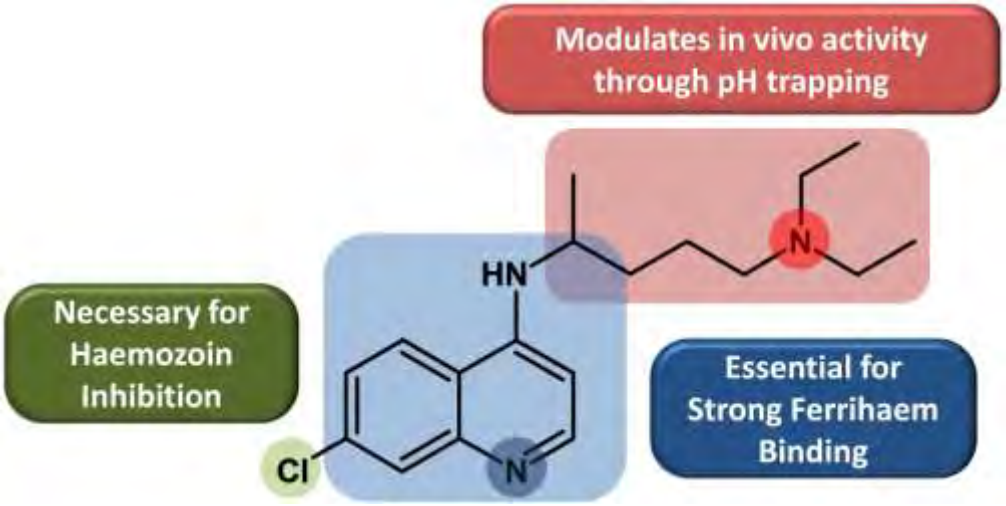


Figure 1.23. Summary of CQ SARs reported by Egan and co-workers[120, 127], Vippagunta *et al*.[161] and Cheruku *et al*.[159] The 4-aminoquinoline moiety (blue) is the smallest substructure that retains strong binding to ferrihaem. Replacement of the 4-amino substituent with a CH group results in weak binding, while replacement of the quinoline nitrogen by a carbon atom abolishes binding as well as β-haematin inhibition and *in vivo* activity. Electron-withdrawing, hydrophobic substituents in the 7-position (green) are crucial for haemozoin inhibition. The side chain (red) can be variable and modulates *in vivo* activity through pH trapping by the basic tertiary amine.

Several other studies have employed DFT calculations to predict properties such as chemical shift tensors and atomic charges of the CQ analogues described above in an attempt to correlate these with the corresponding experimental data and provide a means for rational drug design.[151, 184-188] Varying degrees of success have been observed in this regard. However, none of the proposed structures detailed above have been able to compellingly rationalise the reported SARs and as a result, the fundamental reasons behind the observed correlations are not understood. While these SARs have been able to produce compounds with respectable antiplasmodial activities, invariably the molecules closely resemble CQ. Consequently, drugs with novel scaffolds are unlikely be developed using this rational drug design approach without further structural knowledge of the CQ-ferrihaem complex in solution.

It should be evident from the discussion that there is no consensus regarding the nature of the CQ-ferrihaem complex in solution or solid state. In the absence of a crystal structure, there is no definitive technique for determining the structure of the complex, although, as highlighted by the criticisms of the aryl-methanol-ferrihaem complexes, even a crystal structure does not necessarily provide unequivocal support for the structure of the complex in solution either. This is because a crystal structure is obtained in the solid state and is strongly dependent on the solvent system from which it was crystallised. On the other hand, given the incremental advances in both experimental and computational techniques, an approach combining these methods to investigate the CQ-ferrihaem complex can assist in determining the solution structure. To date no such investigation has been convincingly reported. A number of previously proposed structural models of the complex have been suggested based on inference from experimental data only,[60, 164, 167, 170, 171] while others have been developed using DFT and docking methods but have no direct correlation with experiment.[175, 176] Only one study approached the problem of the CQ-ferrihaem complex using a combination of computational and NMR experimental techniques,[172] however, the methods were not independent from each other, the proposed structure of the complex was not tested using additional experimental data and the study had a number of shortcomings (see above for details).

Finally, it must be emphasised that while more recent evidence tends to indicate CQ exerts is antimalarial activity through adsorption onto the crystal face of haemozoin, understanding the interaction of CQ and ferrihaem in solution is still of relevance given that these two molecules will almost certainly form a complex in the DV of the parasite once haemozoin formation is inhibited. The complex probably plays a central role in the consequent toxicity of ferrihaem towards the parasite. Thus the proposed comprehensive computational and multi-technique experimental investigation of ferrihaem and its complex with CQ is overdue and would help considerably to reduce the confusion surrounding this topic. In this thesis, access to the fastest high performance computing (HPC) cluster on the African continent provided a means to perform more sophisticated calculations than have previously been performed in CQ-ferrihaem complex structural investigations.

1.6 Aims and Objectives

1.6.1 Aims.

The specific aims of this project were:

1. To use multiple experimental techniques to determine stoichiometry and speciation of the CQ-ferrihaem complex in solution and solid state.
2. To use the capabilities of HPC to produce a more realistic model of ferrihaem employing multiple spectroscopic and crystallographic data as a direct test of the quality of the model.
3. To employ the findings of aims 1 and 2, together with HPC, to perform MD simulations in aqueous solution in order to produce structural models of the CQ-ferrihaem complex and to test their validity using independent experimental measurements.

1.6.2 Objectives.

The specific objectives required to achieve the aims of the project were:

1. To comprehensively investigate the CQ-ferrihaem complex in aqueous solution and solid state using a variety of experimental techniques, including UV-visible, magnetic circular dichroism (MCD), IR, EXAFS and NMR spectroscopy as well as magnetic susceptibility and diffusion measurements.
2. To obtain realistic structural and electronic descriptions of monomeric and μ-oxo dimeric ferrihaem species using quantum mechanical (QM) calculations and employing corresponding experimental IR and UV-visible spectra as well as reported crystallographic structures of related iron(III)porphyrins as validation criteria.
3. To develop accurate MM force fields describing five-coordinate ferrihaem species using QM-derived structures, water interaction energies and ESP data.

4. To perform MD simulations of monomeric, π-π dimeric and μ-oxo dimeric ferrihaem species in aqueous solution using the newly established ferrihaem force fields and to extract information regarding solvation, diffusion and structure of these molecules over the course of the simulation.
5. To optimise force field parameters for CQ in order to conduct MD simulations of the CQ-ferrihaem complex in aqueous solution according to findings gained from Objective 1 and to determine its structure, solvation and aggregation state.
6. To use EXAFS and NMR data as independent tests of the validity of the structural models proposed from MD simulations.

CHAPTER 2

BACKGROUND TO TECHNIQUES EMPLOYED

2.1 Introduction

A number of experimental and computational techniques were used throughout this thesis to probe ferrihaem and its interaction with CQ. In order to enable more meaningful interpretation of the reported data, these techniques have been separately reviewed in this chapter.

Experimental techniques described include: (2.2) UV-visible spectroscopy; (2.3) MCD spectroscopy; (2.4) X-ray absorption fine-structure (XAFS) spectroscopy; and (2.5) magnetic susceptibility. UV-visible spectroscopy has been extensively applied to investigate ferrihaem in solution owing to the characteristic spectra ferrihaem species produce. The origins of these spectroscopic features are discussed based on porphyrin models proposed in the literature. The use of MCD as a complimentary technique to UV-visible spectroscopy is described and the fundamentals of the unique MCD spectroscopic terms explained. A brief overview of this technique and its application to metalloporphyrins is also provided. The theory of XAFS is discussed and its application to ferrihaem speciation briefly described. Magnetic susceptibility measurements have been particularly useful in determining ferrihaem speciation. The theory of determining this property in solution and solid state using two methods is detailed.

Computational techniques detailed in this chapter include: (2.6) DFT and (2.7) MM modelling. The theory governing time-independent as well as time-dependent DFT (TD-DFT) calculations is presented. Explanation of DFT functionals and basis sets is also included. The fundamentals of MM modelling are described with an emphasis on the parameters required to be optimised to produce reliable force fields. The theory of MD simulations is also briefly discussed.

2.2 UV-visible Spectroscopy

2.2.1 The Gouterman four-orbital model.

The UV-visible spectra of porphyrins have long been interpreted employing the Gouterman four-orbital model which has been used to account for the two prominent spectroscopic features observed in a variety of porphyrin molecules.[189-193] These features include the intense peak at high energy known as the B- or Soret Band as well as the less intense feature at lower energy termed the Q-band. The simple model proposed by Gouterman considers only $\pi \rightarrow \pi^*$ transitions between the two highest energy occupied π and two lowest energy unoccupied π^* molecular orbitals (MOs) of porphinate (see Fig. 2.1a).[189] Since porphinate has 4-fold D_{4h} symmetry, the unfilled MOs are degenerate (labelled $e_{g,x}$ and $e_{g,y}$ in Fig. 2.1b) and the filled MOs, while close in energy, are non-degenerate with the highest energy occupied MO having the symmetry label a_{2u} and the lower energy MO, a_{1u}.

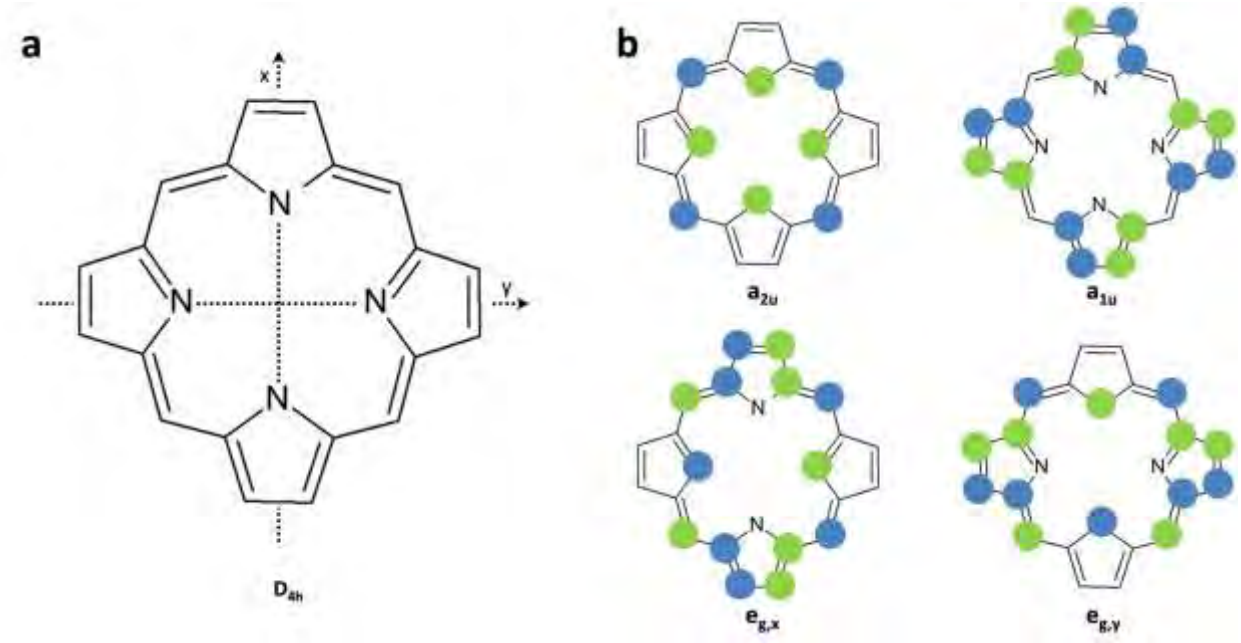


Figure 2.1. (**a**) The porphine anion displaying D_{4h} symmetry. Dashed lines indicate polarisation direction. (**b**) Schematic of the four MOs implicated in the Gouterman model. Filled MOs are a_{2u} and a_{1u} and unfilled MOs are $e_{g,x}$ and $e_{g,y}$. The latter are degenerate, while the former are nearly degenerate.

The two occupied MOs have angular momentum quantum numbers (M_L) of ± 4 while the two unoccupied MOs have M_L = ± 5 (see Fig. 2.2a). Configurational mixing of transitions from the filled to empty MOs give rise to two possible excitations, both with E_u symmetry ($a_{1u}/a_{2u} \times e_g$, see Fig. 2.2b).[189] The first excitation involves transitions in which the direction of angular momentum is conserved with ΔM_L = ± 1 ([–5 – (–4) = –1] or [+5 –(+4) = +1]), while the second arises from transitions where direction has inverted with ΔM_L = ± 9 ([+5 –(–4) = +9] or [–5 –(+4) = –9]). According to Hund's rule, the former has higher energy owing to greater electron-electron repulsion since electrons circulate in opposite directions and thus encounter each other more frequently.[194] This leads to a high-energy excited state. Conversely, the ΔM_L = ± 9 transition gives an excited state that is lower in energy because electrons circulate in the same direction, encountering each other less often and thus producing a lower electron-electron repulsion. Based on the electronic selection rule (ΔM_L = ± 1), transitions comprising the higher-energy excited state are allowed and give rise to an intense band, experimentally termed the B-band. On the other hand, transitions comprising the lower energy excited state are forbidden since ΔM_L = ± 9 and thus produce a band with weaker intensity, known as the Q-band.[195] The energy difference between these two bands, ΔE, has been shown to be approximately the repulsion between the electrons separated by the radius of the porphyrin ring, r_p, and can be described by equation 2.2.1.[196]

$$\Delta E \simeq \frac{e^2}{2r_p} = 7259 \text{ cm}^{-1} \qquad (2.2.1)$$

In addition to the B- and Q-bands, two related vibronic bands, B_v and Q_v, are also often observed and occur when electrons in the $e_g(\pi^*)$ MOs are promoted to a vibrationally excited state. In the case of porphyrins, the Q_v-band is approx. 1200 cm^{-1} higher in energy than the corresponding Q-band and borrows intensity from the B-band through a vibronic coupling mechanism.[197]

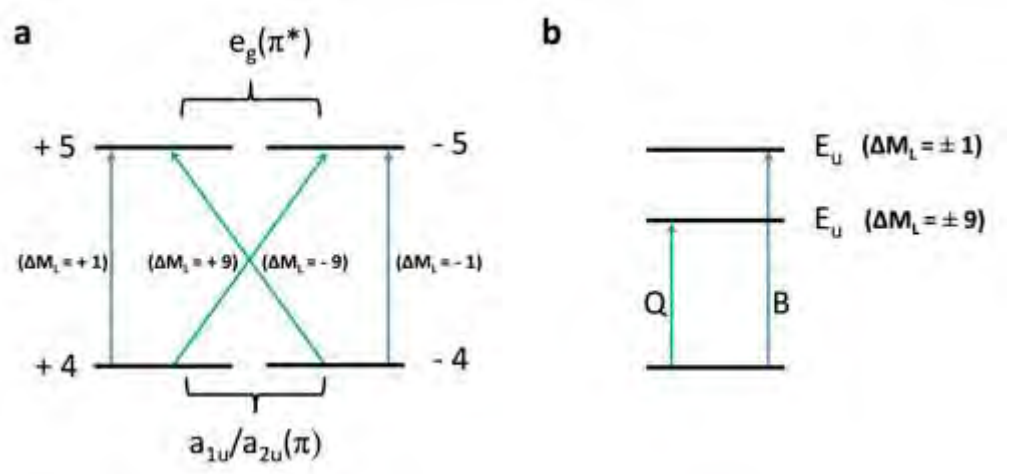


Figure 2.2. (**a**) Possible $\pi \rightarrow \pi^*$ transitions in porphinate implicated in the Gouterman four-orbital model. Numbers to the left and right of MOs represent M_L. (**b**) Excitations giving rise to the B- (Soret) and Q-bands. Transitions giving rise to the Q-band are forbidden and thus produce peaks with weak intensity while those comprising the B-band produce an intense peak.

2.2.2 Interpretation of the UV-visible spectrum of ferrihaem.

As described above, the Gouterman four-orbital model has been successfully applied to account for the $\pi \rightarrow \pi^*$ transitions of various metalloporphyrins. However, for those having unfilled metal d-orbitals, additional ligand to metal charge transfer transitions are possible which are not accounted for by the Gouterman model. Such a situation arises for monomeric H_2O-ferrihaem which is paramagnetic and high-spin with five unpaired electrons occupying metal d-orbitals ($S = {}^5/_2$). This molecule is structurally and spectroscopically similar to H_2O-metmyoglobin (metMb).[196] The UV-visible spectrum of H_2O-metMb was assigned by Makinen and Churg using the Hückel linear combination of atomic orbitals (LCAO) model (see MO diagram in Fig. 2.3a).[196] The same $\pi \rightarrow \pi^*$ transitions implicated in the Gouterman four-orbital model that comprise the B- and Q-bands were identified by the LCAO method.[189] Peaks at lower energy than the Q-band were assigned to charge transfer transitions from ligand a_{1u} and a_{2u} MOs to unfilled metal d_{xz}- and d_{yz}-orbitals (collectively termed d_π-orbitals owing to their near degeneracy). Given the spectroscopic similarity between H_2O-metMb and H_2O-ferrihaem, UV-visible band assignments of the former have been used to describe analogous peaks in the spectrum of the latter.[165] Prominent spectroscopic features of H_2O-ferrihaem are depicted in Fig. 2.3b.

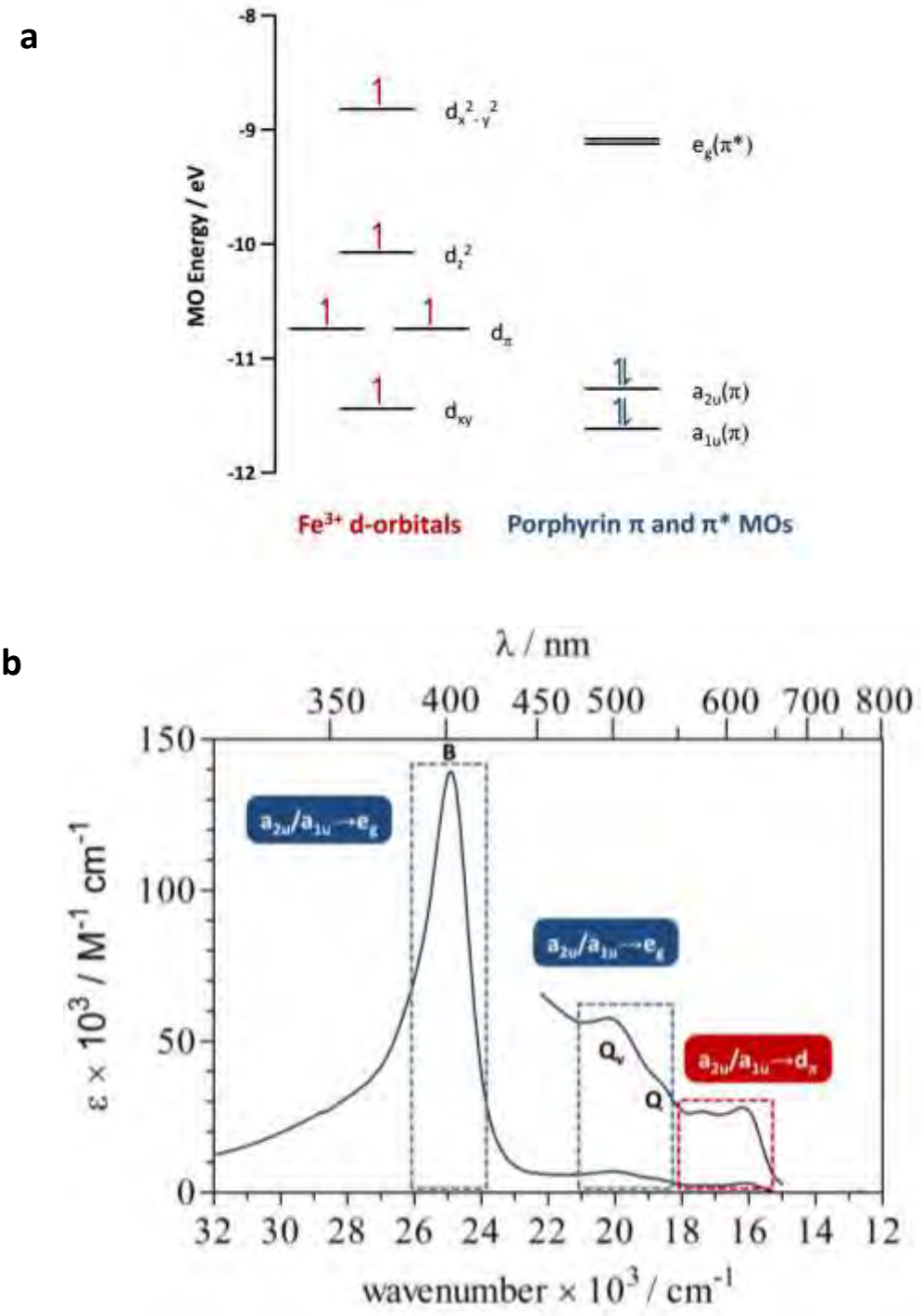


Figure 2.3. (**a**) MO diagram of the ferrihaem chromophore calculated using Hückel LCAO theory for H_2O-metMb. (**b**) UV-visible spectrum of H_2O-ferrihaem in 40% (v/v) aq. DMSO, pH 7.4. Putative assignments of prominent features have been made based on analogous bands in the spectrum of H_2O-metMb identified using the Hückel LCAO model.[196] Porphyrin $\pi \rightarrow \pi^*$ transitions are shown in blue while charge transfer transitions are displayed in red.

2.3 Magnetic Circular Dichroism Spectroscopy

2.3.1 Circularly polarised light.

Circular dichroism (CD) and MCD spectroscopy are based on the wavelength-dependent absorption of circularly polarised light to produce excited states. In the circularly polarised light, the electric field of light waves is comprised of x- and y-components which are orthogonal to each other and are shifted by one quarter of a wavelength ($^{\pi}/_{2}$). A resultant net electronic vector that circulates in a left or right-handed manner is thus produced (see Fig. 2.4).[198]

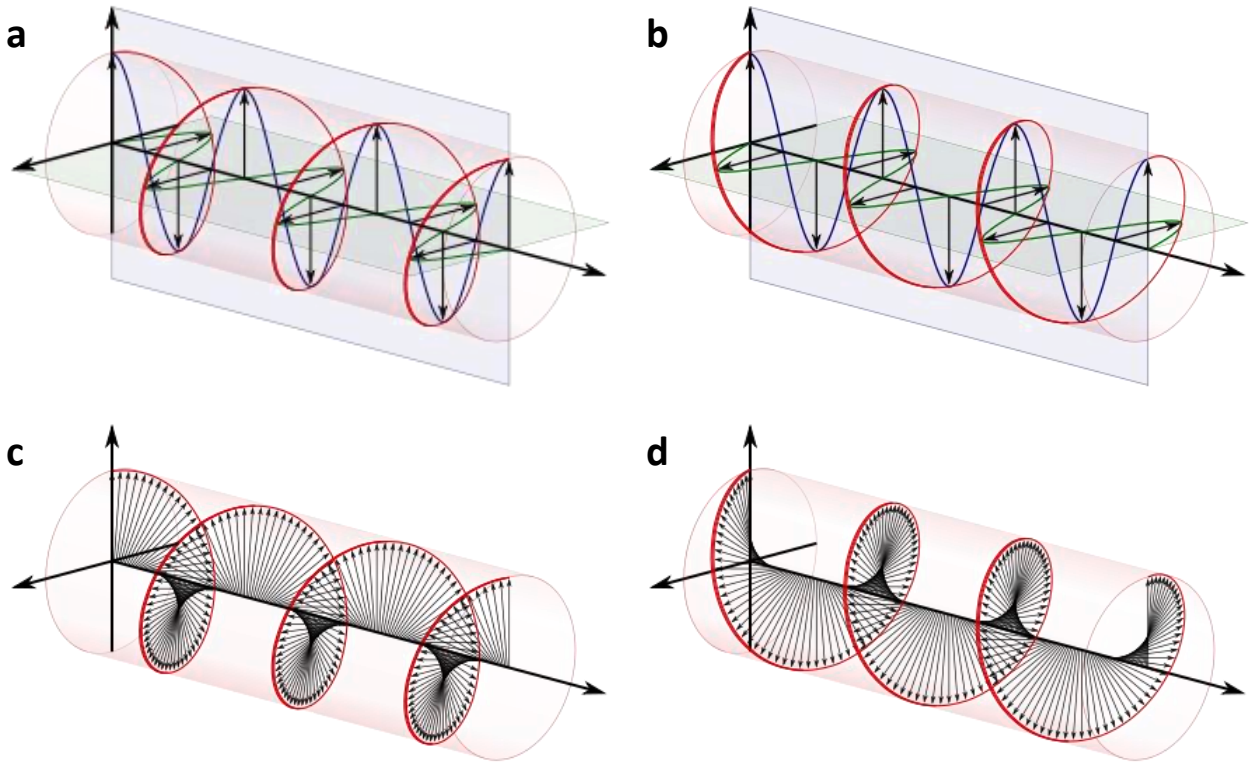


Figure 2.4. (**a** and **c**) Left- and (**b** and **d**) right-circularly polarised light. Blue and green waves represent the x- and y-electric field components of light, respectively. Black arrows in (c) and (d) are the resultant vector direction of electric field components.

2.3.2 MCD theory.

Excitations produced in MCD/CD spectroscopy are the same as those in UV-visible spectroscopy but band morphologies differ between the two techniques because the former measures the difference (ΔA) between left- and right-circularly polarised light (lcp and rcp) absorption, A_L and A_R, respectively.[199, 200] Since the difference in absorbance of lcp and rcp is measured, a handedness to the molecule under investigation is required. In the case of CD, this handedness is only achieved in chiral molecules. By contrast, any molecule can be investigated using MCD. This is because a strong magnetic field is applied along the direction of light propagation which lifts degeneracy of MOs and thus makes transitions involving these MOs non-equivalent in terms of energy. Splitting of degenerate MOs by a magnetic field is known as the Zeeman effect and occurs as a result of alignment of the M_L of MOs with or against the field (see Fig. 2.5).[199]

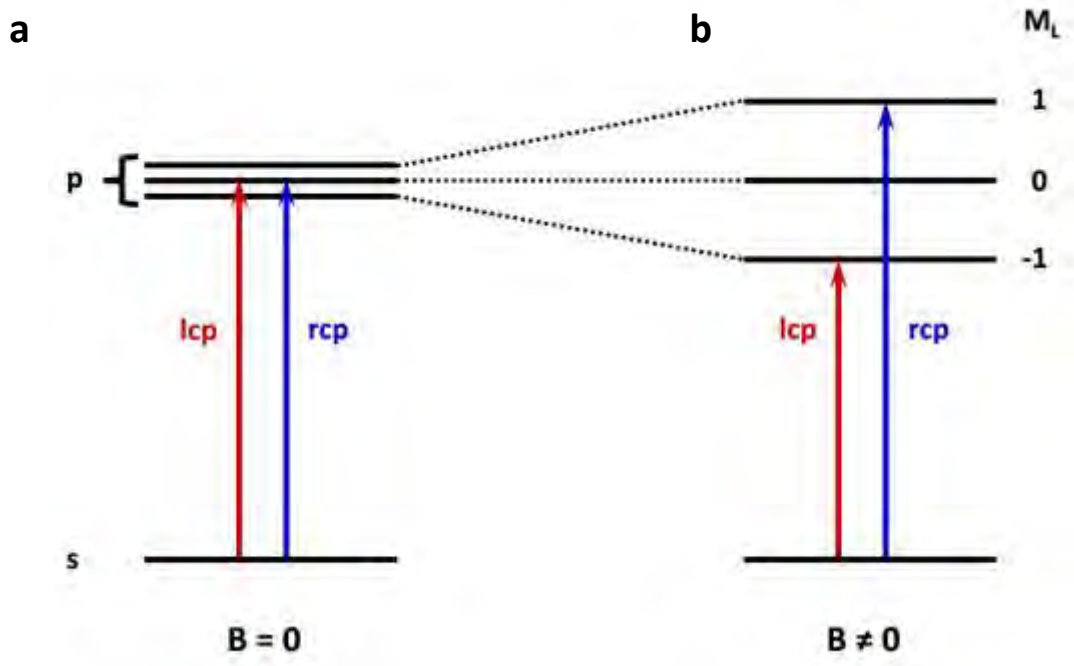


Figure 2.5. Effect on hypothetical degenerate p-orbitals in the absence (**a**) and presence (**b**) of an applied magnetic field, *B*. The latter case results in a difference in energy of left- and right-circularly polarised light absorption.

In both CD and MCD, extinction coefficients ($\Delta\varepsilon$, ε_L and ε_R) are related to absorbance via Beer's law as shown in equation 2.3.1, where *l* is the pathlength (in cm) and *c* the concentration of sample (in M). However, in the case of MCD extinction coefficients are also dependent on the magnitude of magnetic field, *B* (in T).[200] Experimentally, spectra are often recorded in millidegrees, θ, which is related to absorbance by equation 2.3.2.[198]

$$\Delta A = A_L - A_R = (\varepsilon_L - \varepsilon_R) \times l \times c \times B = \Delta\varepsilon \times l \times c \times B \quad (2.3.1)$$

$$\Delta A = \frac{\theta}{32982} \quad (2.3.2)$$

The MCD intensity of an electronic transition is described by equation 2.3.3 where *E* is energy and *f* a normalised band-shape function (usually a Gaussian).[201-204] The three Faraday terms A_1, B_0 and C_0 can be separated into characteristic MCD A-, B- and C-terms which form part of observed spectra. These are described in Section 2.3.3.

$$\frac{\Delta A}{E} = 152.5\, Bcl\left[A_1\left(-\frac{df}{dE}\right) + \left(B_0 + \frac{C_0}{kT}\right)\right] \quad (2.3.3)$$

It is the presence of the A-, B- and C-features and the fact that they can be either positive or negative owing to the difference nature of the MCD technique that gives the spectra a greater resolution than compared to UV-visible spectroscopy in which only positive Gaussian-type bands are observed.

2.3.3 A-, B- and C-terms.

In MCD spectra, A-terms are observed as highly distinctive Gaussian derivative-shaped bands. This feature arises from excitations originating in non-degenerate filled MOs and terminating in degenerate unfilled MOs. In the absence of a magnetic field, absorbance of lcp and rcp light, A_- (from lcp) and A_+ (from rcp), would be equal and result in no MCD spectrum (see Fig. 2.6a). In the presence of a magnetic field, Zeeman splitting occurs which causes a difference in energy between lcp and rcp absorption leading to the characteristic Gaussian derivative-shaped feature being observed (Fig. 2.6b).[200] Absorption of lcp can be higher or lower in energy thus both positive and negative features are possible.

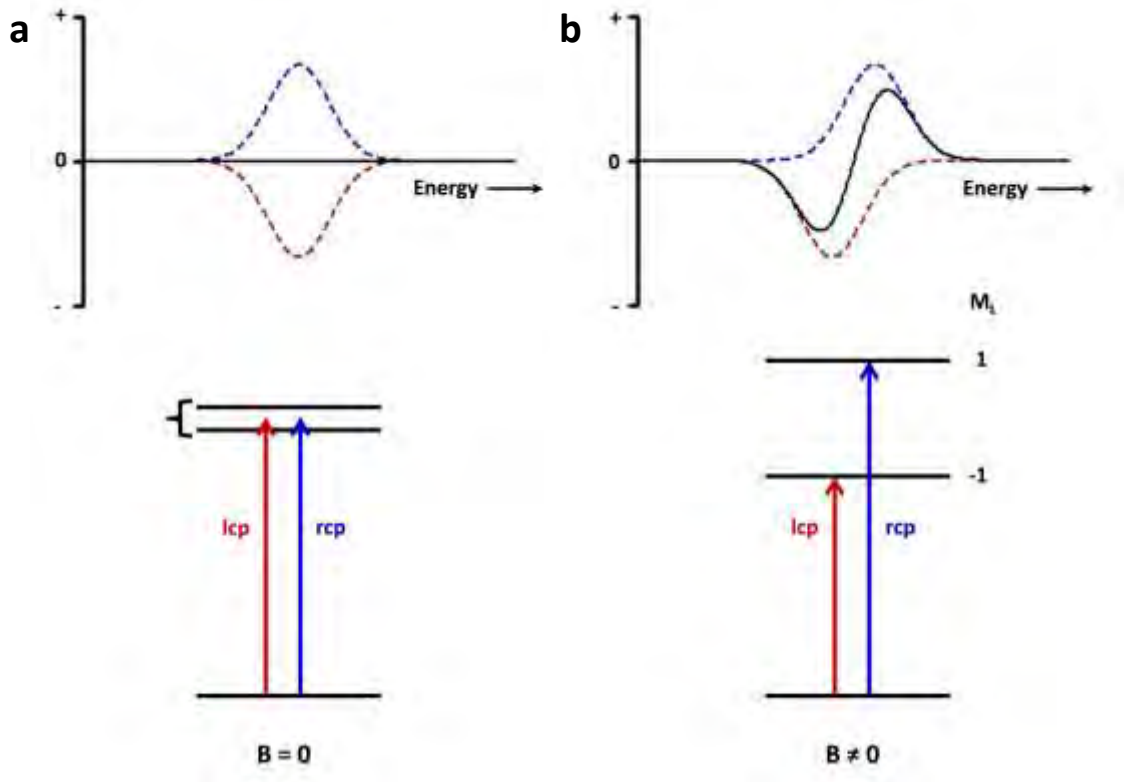


Figure 2.6. Origin of MCD A-term. (**a**) In the absence of magnetic field, B, transitions terminating in degenerate MOs have no energy difference in lcp and rcp (red and blue dashed lines, respectively), thus produce no MCD spectrum. (**b**) In the presence of a magnetic field, Zeeman splitting occurs resulting in the observed Gaussian derivative-shaped band (black solid line) known as an A-term.

The mechanism leading to Gaussian C-terms is similar to that described for A-terms with the exception that MO splitting occurs in the filled MOs. As in the case described above, when no magnetic field is present, no MCD spectrum is produced (see Fig. 2.7a). This is because excitations arising from transitions originating in degenerate MOs have the same energy and thus cancel. The presence of a magnetic field, however, can produce one of two features in the MCD spectrum. The first arises at higher temperatures where *kT* is of the order of the Zeeman splitting. In this case, a Gaussian derivative-shaped feature similar to that described for an A-term is produced (see Fig. 2.7b). The second feature is observed at low temperature where the Boltzmann population of the lower energy MO will be greater than that of the higher energy MO (see Fig. 2.7c).[200] Under these conditions, intense Gaussian-shaped bands are observed. Given the temperature dependence of this feature, variable temperature MCD is often used to distinguish C-term bands. In the case of metalloporphyrins, C-terms are usually associated with paramagnetic metal ions.[199]

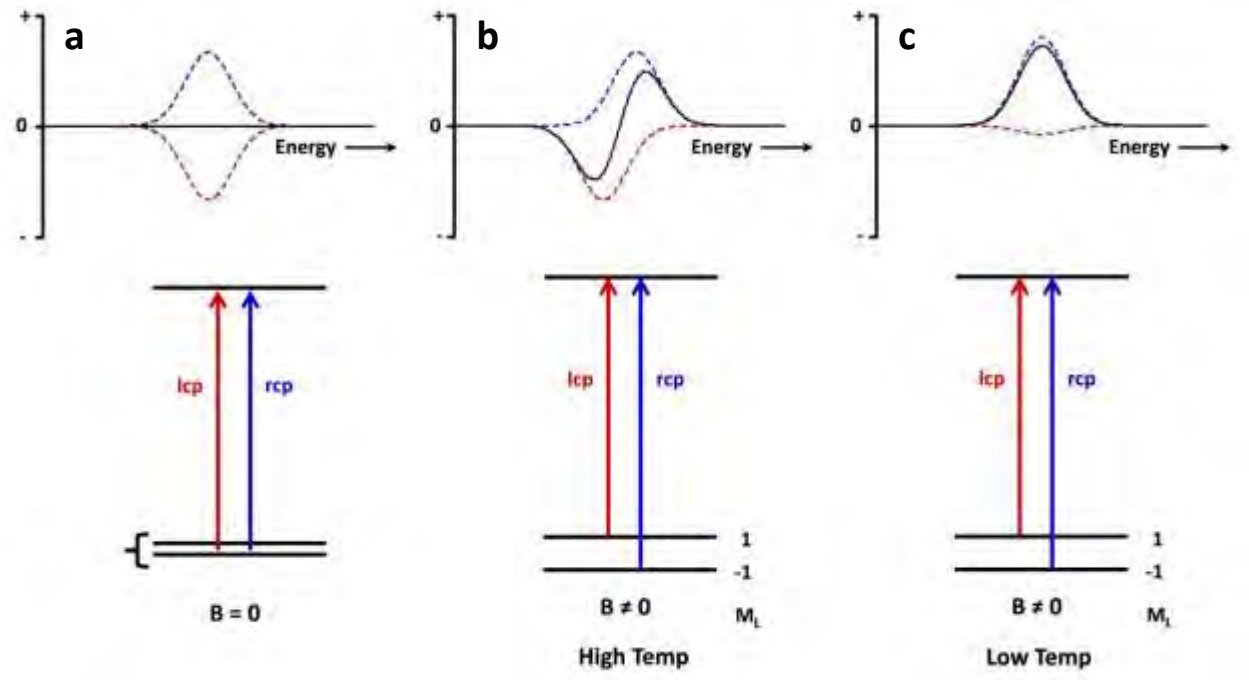


Figure 2.7. Origin of MCD C-term. (**a**) In the absence of magnetic field, *B*, no MCD spectrum is produced. (**b**) In the presence of a magnetic field and at higher temperatures, Zeeman splitting of degenerate ground-state MOs is observed which leads to a Gaussian derivative-shaped feature. (**c**) In the presence of a magnetic field and at low temperature, the lower energy ground-state MO will have a greater Boltzmann distribution population than the higher energy MO and thus produces a Gaussian-shaped band.

B-terms arise from field-induced mixing between excited states via magnetic dipole moments, μ_z.[199] Two such situations can occur, mixing between closely spaced unfilled MOs (see Fig. 2.8a) or between closely spaced occupied MOs (see Fig 2.8b). In the case of the former, a weak Gaussian-shaped feature is produced, while in the case of the latter, an intense Gaussian-shaped feature is observed and is qualitatively the same as a C-term. For molecules where there is no three-fold or higher rotation axis, B-terms dominate MCD spectra because there are no degenerate MOs to undergo Zeeman splitting.[199, 200] In special cases, where two excited states that undergo mixing are similar in energy but are oppositely signed, a Gaussian derivative-shaped feature known as a pseudo A-term can arise (see Fig. 2.8c).[198] These features often occur when small structural changes cause MOs to become non-degenerate such as in the case of ferrihaem where porphyrin substitution lifts the degeneracy of the $e_g(\pi^*)$ MOs, for example, which are equivalent in porphinate (see Section 2.2).

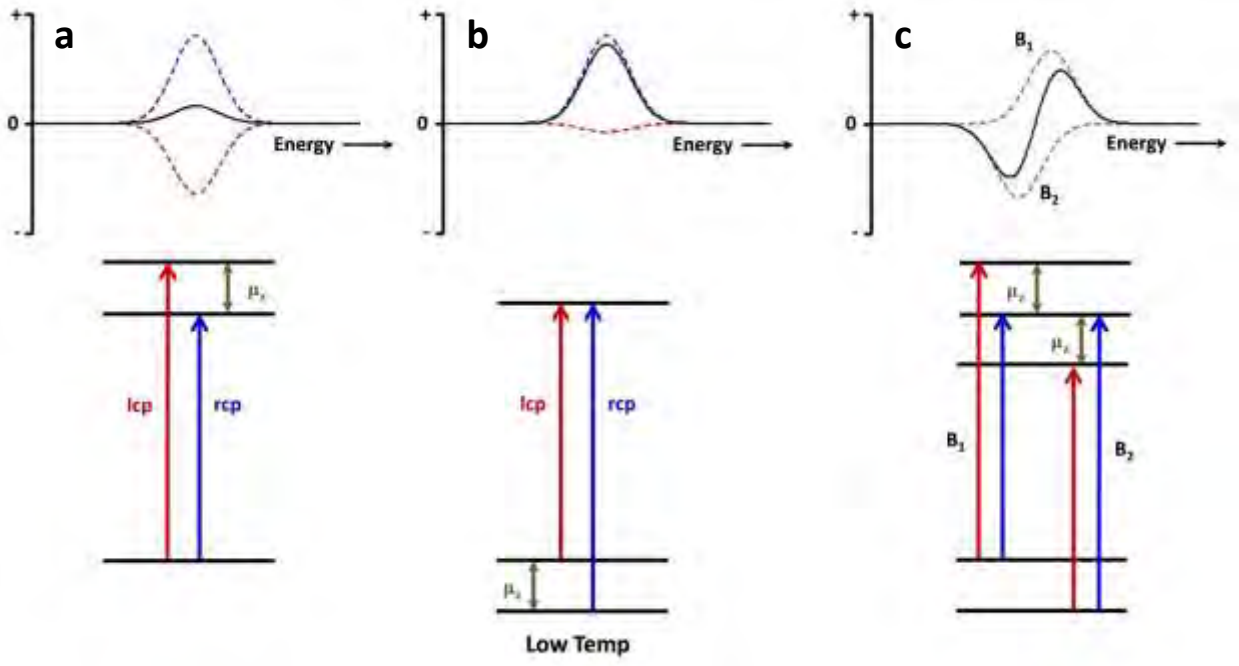


Figure 2.8. Origin of MCD B-terms. (**a**) When an intermediate state is close in energy to an excited state, mixing via magnetic dipole moments, μ_Z, occurs which gives rise to a weak Gaussian-shaped band. (**b**) When mixing occurs in the ground state, intense Gaussian shaped-features are observed which are effectively the same as C-terms (see Fig. 2.6c). (**c**) Pseudo A-terms are produced when two closely-spaced and oppositely signed B-terms (grey dotted lines) overlap.

2.3.4 Application of MCD to porphyrins.

MCD has been extensively applied to a variety of porphyrins and metalloporphyrins in the literature.[193, 199, 205-208] As in the case of UV-visible spectroscopy, the Gouterman four-orbital model can be applied to describe $\pi \rightarrow \pi^*$ transitions in MCD spectra (see Section 2.2.1 for details on the Gouterman model). This occurs because the electronic transitions which give rise to features in MCD spectra are the same as those producing features in UV-visible spectra. Thus MCD spectra of symmetric porphyrins have prominent Q- and B-band features. The excitations which give rise to these spectroscopic envelopes involve $\pi \rightarrow \pi^*$ transitions between the two highest occupied and lowest unoccupied MOs (see Section 2.2). Consequently, owing to the degeneracy of these MOs, intense A-term bands are expected (see Section 2.3.3 for explanation of A-term origin). Indeed these features are observed in the spectrum of zinc(II)octaethylporphyrin, Zn(II)OEP (see Fig. 2.9).[205] The intensity of the Q-band in MCD is enhanced relative to that observed in the UV-visible spectrum because the mechanism of intensity in MCD involves both the magnetic and transition dipole moments, whereas UV-visible intensities are governed only by the latter.[199]

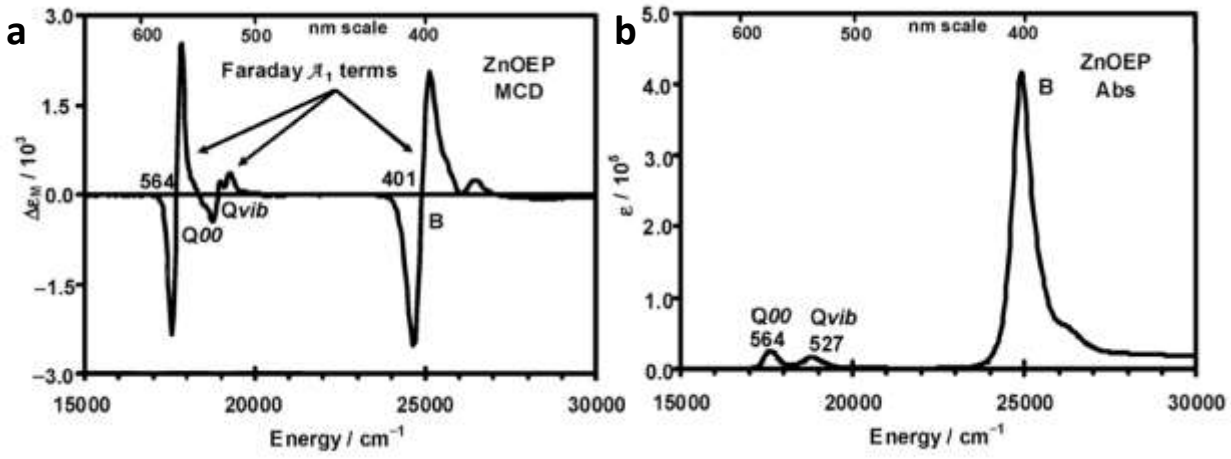


Figure 2.9. MCD (**a**) and UV-visible (**b**) spectra of Zn(II)OEP reported by Mack *et al*.[205] Figure adapted with permission from Mack *et al*.[205] © 2007, Elsevier.

As in the case of UV-visible spectra, the introduction of transition-metal ions with unfilled d-orbitals (both paramagnetic and diamagnetic) complicates MCD spectra. This is because of the presence of additional ligand to metal charge transfer transitions. Since the Gouterman four-orbital model does not account for these types of transitions, it cannot be used to assign bands in MCD or UV-visible spectra. An example of the complication caused by the additional transitions can been seen in the spectra reported for Cl-iron(III)tetraphenylporphyrin, Cl-Fe(III)TPP, reported by Paulat and Lehnert (see Fig. 2.10).[207] In this case, owing to the paramagnetic iron centre, low-temperature MCD was performed to isolate MCD C-terms. To aid assignment of features in the UV-visible and MCD spectra, simultaneous spectroscopic deconvolution was performed to identify underlying bands. This technique has been successfully used on a number of metalloporphyrins and is particularly useful because simultaneous deconvolution of both MCD and UV-visible spectra reduces the ambiguity of fitted bands.[205, 207, 209] Using this approach in combination with TD-DFT calculations (see Section 2.6.1), compositions of the underlying bands could be suggested.

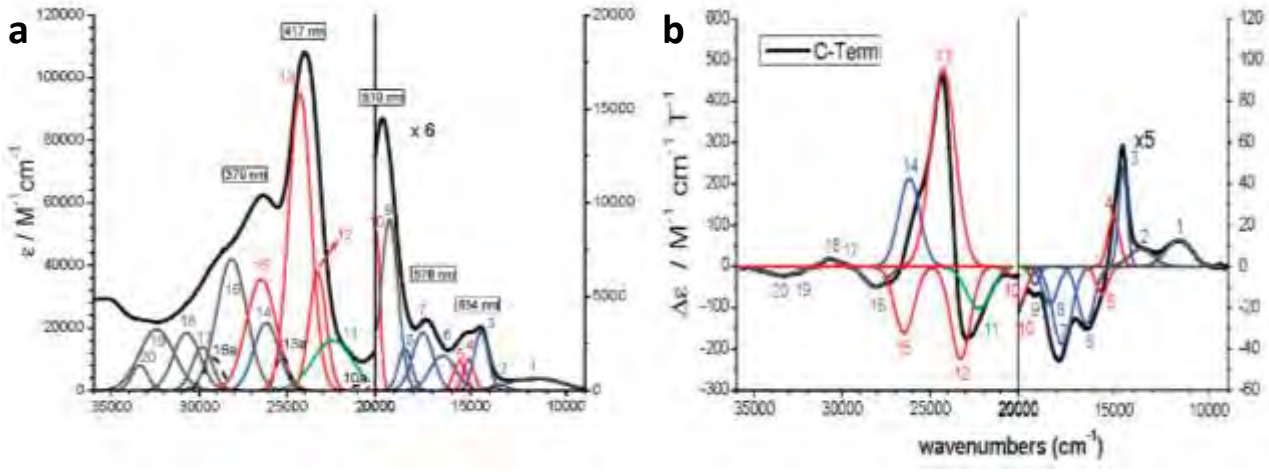


Figure 2.10. UV-visible (**a**) and low-temperature MCD (**b**) spectra of Cl-Fe(III)TPP reported by Paulat and Lehnert.[207] Red, blue, green and grey Gaussian bands have xy-, z-, mixed and undetermined polarisations, respectively. Experimental spectrum is shown in black. Figure adapted with permission from Paulat and Lehnert[207] © 2008, American Chemical Society.

The B-band (or Soret band) was found to be composed of three bands (12, 13 and 15 in Fig. 2.10). While all three bands had significant contributions from $\pi \rightarrow \pi^*$ transitions involved in the Gouterman four-orbital model, greater contribution was observed from transitions originating in lower MOs with b_{2u} symmetry. In addition, by contrast to Zn(II)OEP, there was also a significant $\pi \rightarrow d_\pi$ charge transfer contribution. The Q-band is markedly less prominent in Cl-Fe(III)TPP (see bands 4 and 5 in Fig. 2.10) and while it has a greater contribution from the Gouterman $\pi \rightarrow \pi^*$ transitions than the B-band, $\pi \rightarrow d_\pi$ charge transfer transitions dominate.

Further complexity is introduced by lowering the symmetry of the porphyrin ring system since the ordering of MOs is strongly influenced by porphyrin substituents and axial ligands. This can be seen for instance in the ferrihaem prosthetic group in high-spin cytochrome $P450_{cam}$ reported by Galinato *et al*.[209] In this case, the highest occupied MO indicated from DFT calculations was an axial ligand orbital and was not the a_{1u}/a_{2u} MO as expected for porphyrins using the Gouterman model. Changes to MO energies affect spectroscopic features observed in the UV-visible and MCD spectra (see Fig. 2.11) and make assignment of bands solely by the Gouterman model all but impossible.

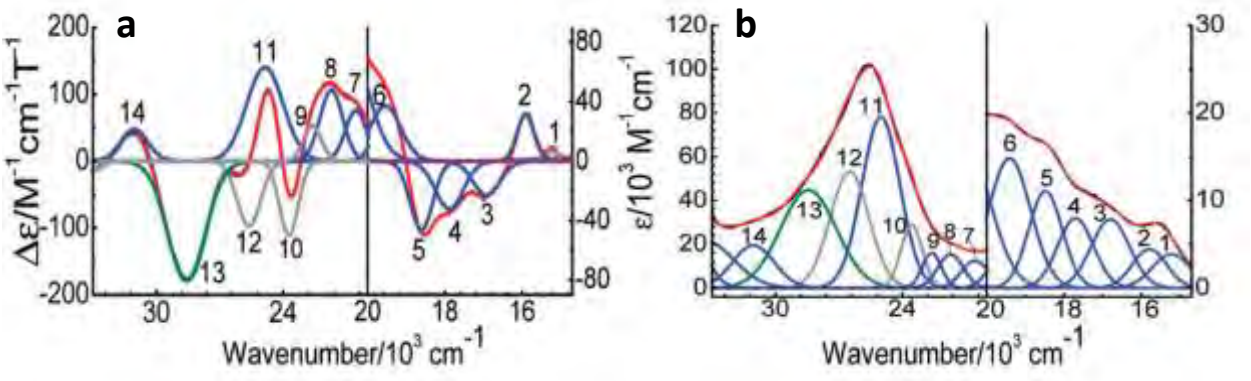


Figure 2.11. Deconvoluted low-temperature MCD (**a**) and UV-visible (**b**) spectra of cytochrome $P450_{cam}$ reported by Galinato *et al*.[209] Blue, red and grey Gaussian bands have xy-, z- and undetermined polarisations, respectively. The experimental spectrum is shown in red. Figure adapted with permission from Galinato *et al*.[209] © 2011, American Chemical Society.

This complication is particularly evident from the spectroscopic assignment of the deconvoluted bands making up the Soret envelope (see bands 10, 11 and 12 in Fig. 2.11) proposed by Galinato *et al*. with the aid of TD-DFT. In this case, there is only a very small contribution from the $\pi \rightarrow \pi^*$ transitions implicated in the Gouterman four-orbital model. Excitations making up the Soret band rather consist of minor $\pi \rightarrow \pi^*$ transitions originating in lower-energy, low-symmetry porphyrin orbitals which terminate in $e_g(\pi^*)$ MOs; and a larger component of various charge transfer transitions originating in porphyrin π- as well as axial sulfur ligand p_z-orbitals which terminate in either d_z^2-, d_π- or d_{xy}-orbitals. On the other hand, excitations comprising the Q-band (bands 5 and 6 in Fig. 2.11) have substantial contributions from Gouterman $\pi \rightarrow \pi^*$ transitions although they are heavily mixed with transitions originating in axial sulfur p_x- and p_z-orbitals which terminate in $e_g(\pi^*)$ or d_{xy} MOs. Lower energy bands (2 to 4 in Fig. 2.11) were charge transfer bands involving transitions that terminated in metal d_π-orbitals and originated in $a_{1u}/a_{2u}(\pi)$ porphyrin MOs or from the axial sulfur p_x-orbital.

2.4 X-Ray Absorption Fine-Structure Spectroscopy

XAFS is a spectroscopic technique which is able to provide information regarding the chemical and physical state of an atom. The coordination number, types of coordinating species as well as details of coordination geometry (distances, angles, etc.), typically within 5 Å, can be determined.[210] Consequently, XAFS is particularly useful when crystal structures are not available. However, this technique is not able to produce absolute structures and instead requires fitting of a model to measured data. Thus development of reliable models is critical. On the other hand, it has often been used in conjunction with crystal structure data because XAFS spectra can be recorded on samples in both the solid-state or in solution which may produce structures different to those obtained under crystallisation conditions.[211]

The XAFS technique is divided into two categories which deal with different parts of XAFS spectra, namely X-ray absorption near-edge spectroscopy (XANES) and EXAFS. While the theory will be discussed for XAFS, this section will focus on the latter from which molecular structure is usually determined.

2.4.1 Theory of XAFS.

The XAFS technique is based on the use of X-rays with energies near or above core electron binding energies which are generated using a synchrotron source. At these energies, the photoelectric effect occurs in a process where matter absorbs an X-ray photon and emits an electron from a core level (for example from 1s or 2p orbitals) as a photoelectron into the continuum (see Fig. 2.12a).[211, 212] At this stage the atom is referred to as being in an excited state which normally decays within a few femtoseconds. The decay of the excited state occurs when a higher-energy electron fills the core hole and releases a fluorescent X-ray photon (see Fig. 2.12b).[211]

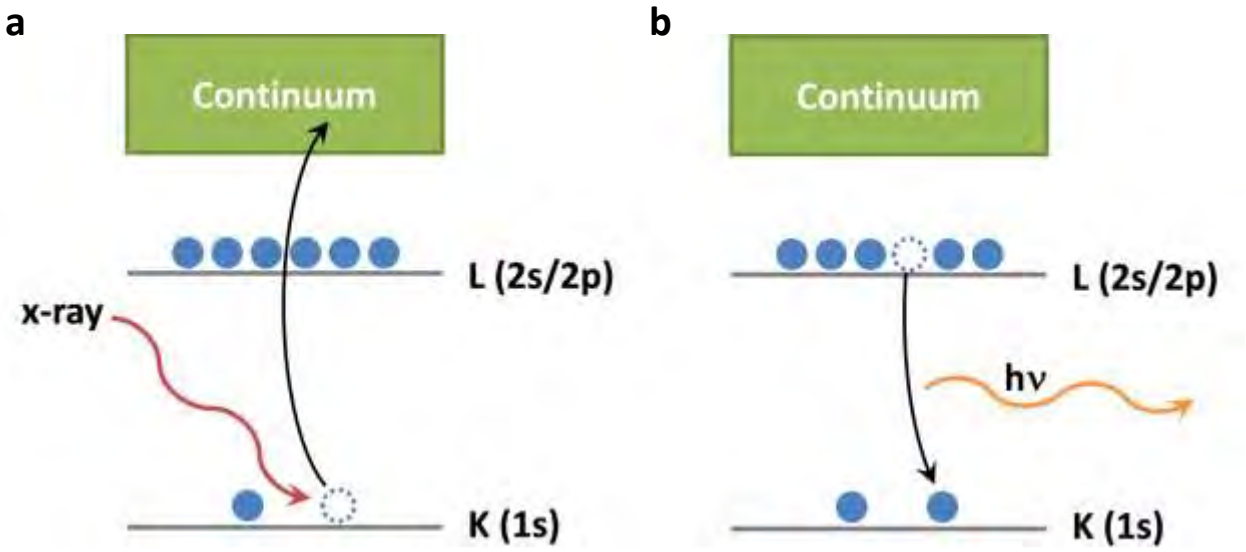


Figure 2.12. (**a**) The photoelectric effect which forms the basis of the XAFS technique. X-rays are absorbed by an atom and core-level electrons are promoted into the continuum producing an excited state. (**b**) The excited state decays when a higher-energy electron fills the core hole and releases a fluorescent X-ray photon with energy, $h\nu$.

The X-ray absorption coefficient, μ, is described by Beer's Law according to eq. 2.4.1, where *I* is the intensity of the X-ray transmitted through the sample or emitted fluorescent photon, I_0 the incident X-ray intensity and *t* the thickness of the sample.[210]

$$I = I_0 e^{-\mu t} \tag{2.4.1}$$

When an absorption event occurs (i.e. the energy of an X-ray photon is equal to the binding energy of a core electron), there is a sharp increase in μ since the electron is promoted into the continuum.[212] This sharp increase is known as the absorption edge. In the case of a hypothetical isolated atom with no nearby neighbours, when μ is plotted as a function of energy, a smooth decay after the absorption edge is observed (see Fig. 2.13a). In reality, however, neighbouring atoms near to the absorber scatter the emitted fluorescent wave, possibly with different mechanisms (see Fig. 2.13b and c) and thus produce a series of oscillations in the plot after the absorption event. It is these interference patterns from which structural information can be obtained.[212] In the XANES regime (typically within 30 eV of the absorption edge), oxidation state and coordination chemistry (octahedral, tetrahedral, etc. coordination) can usually be obtained.[212] At higher energies, in the EXAFS regime, details of coordination number, distances, etc. can be determined using eq. 2.4.2 and 2.4.3.[210, 213, 214] In eq. 2.4.2, S_0^2 is the amplitude reduction factor, *k* the wave number of the photoelectron, *N* is the number of equivalent neighbouring atoms, *R* the distance to the neighbouring atom, σ^2 the disorder in distance, $\lambda(k)$ the mean free path of the electron, $f(k)$ the scattering amplitude and $\delta(k)$ the phase-shift. In eq. 2.4.3, E_0 is the absorption edge energy and *m* the electron mass.

$$\chi(k) = \sum_j \frac{N_j}{kR_j^2} S_0^2 f_i(k) e^{-2k^2\sigma_j^2} e^{\frac{-2R_j}{\lambda(k)}} \sin\left[2kR_j + \delta_j(k)\right] \tag{2.4.2}$$

$$k = \sqrt{\frac{4\pi m(E - E_0)}{h}} \tag{2.4.3}$$

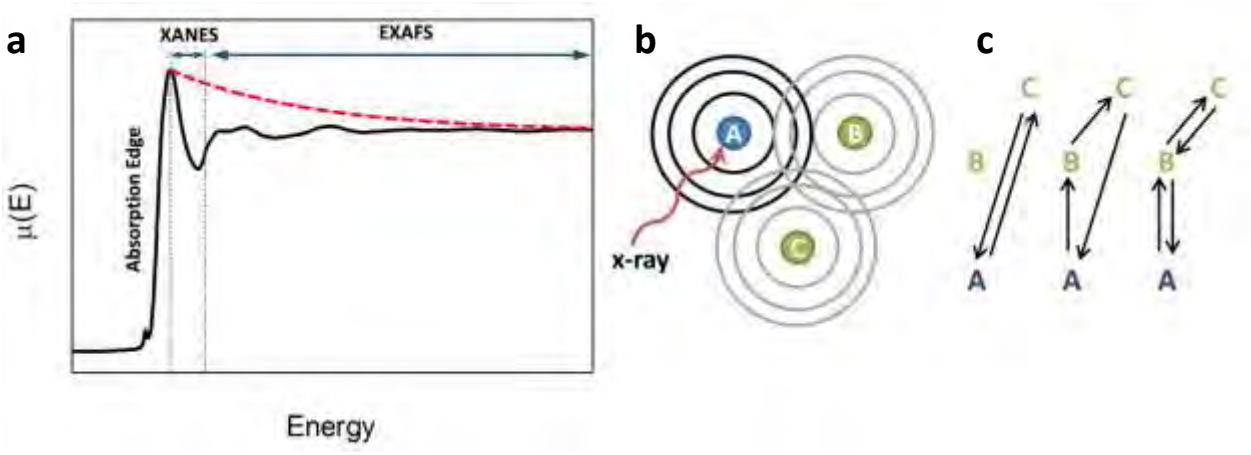


Figure 2.13. (**a**) Plot of absorption coefficient, μ, as a function of energy, *E*. For a hypothetical atom with no nearby neighbours, a smooth decay is observed after the absorption edge (red dashed line). In reality, scattering of the emitted wave is caused by nearby atoms leading to oscillations being observed at energies higher than the absorption edge (solid black line). XAFS spectra are grouped into two regions, XANES at energies within 30 eV of the absorption edge and EXAFS at higher energies. (**b** and **c**) The emitted wave (black rings) are scattered by nearby atoms giving rise to scattered waves (grey rings). Scattering can follow different pathway mechanisms (A=absorber, B and C=scatters).

Scattering factors *f(k)* and $\delta(k)$ can be calculated for a given model using multiple-scattering theory which is implemented in XAFS analysis software such as FEFF.[212] Knowing *f(k)* and $\delta(k)$ allows refinement of S_0^2, *N*, *R* and σ^2 factors in a fitting process where the deviation of the calculated from experimental EXAFS spectrum data is minimised. Two functions, the *R*-factor (residual function) and χ^2 (chi-squared), are usually used for this minimisation process and are described by eq. 2.4.4 and 2.4.5, respectively, where ε_i is the standard deviation at the *i*-th point.[210] Restraints are routinely used to prevent fitted parameters from producing unrealistic values.

$$R = \frac{\sum_{i=1}^{N} |Data_i - Model_i|^2}{\sum_{i=1}^{N} |Data_i|^2} \quad (2.4.4)$$

$$\chi^2 = W \sum_{i=1}^{N} \frac{|Data_i - Model_i|^2}{{\varepsilon_i}^2} \quad (2.4.5)$$

In general, computed and experimental spectra are reported in *k*-space but a more useful form of the data is obtained through Fourier transformation where an effective radial distribution function, called *R*-space, is produced (see Fig. 2.14).[212] In this form, the distance of scattering atoms from the absorbing atom is more clearly seen.

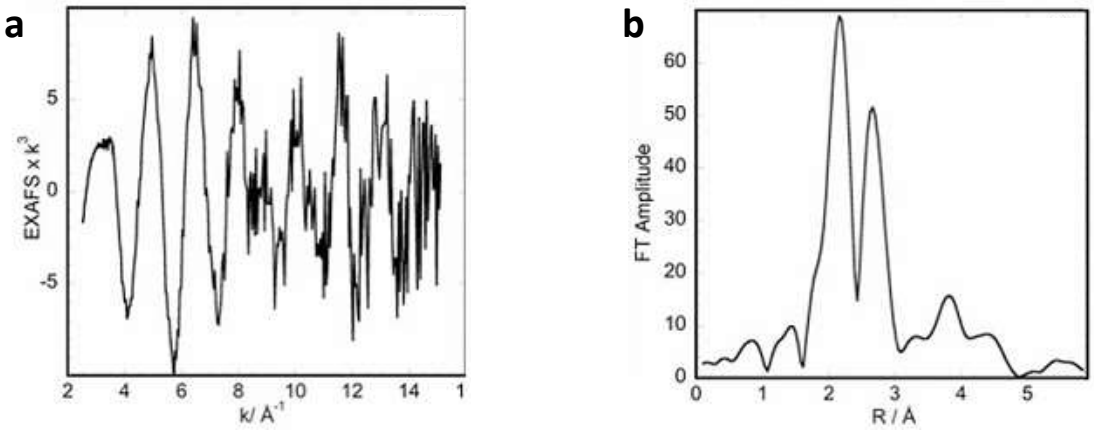


Figure 2.14. An example of the experimental EXAFS spectra of the iron protein in *Klebsiella pneumoniae* reported by Asocone and Strange[215] in (**a**) *k*- and (**b**) *R*-space. Figure adapted with permission from Asocone and Strange[215] © 2009, International Union of Crystallography.

2.4.2 Application of EXAFS to ferrihaem.

Relatively few EXAFS studies have been performed on free ferrihaem species. One of the earliest reports used this technique to confirm the μ-propionato substructure of synthetic haemozoin, specifically by identifying the presence of the iron-carboxylate bond.[216] More recently, EXAFS was used to probe the structure of mesoporphyrin and deutroporphyrin analogues of μ-propionato ferrihaem in the solid state.[217] The four major features of the EXAFS spectrum of synthetic haemozoin were well reproduced by both analogues (see Fig. 2.15a). These included the peak around 2 Å representing the first coordination sphere (see Fig, 2.15b); the peak around 3.1 Å which arises from C_α atoms and the carboxyl carbon of the axial ligand; the shoulder around 3.4 Å representing C_{meso} atoms; and the peak around 4.3 Å arising from C_β atoms. Furthermore, it was shown that EXAFS spectra could distinguish between monomeric and dimeric forms of ferrihaem with the calculated Fe-O bond length of the former species being slightly shorter than that of the dimer (1.85 vs. 1.88 Å).[217]

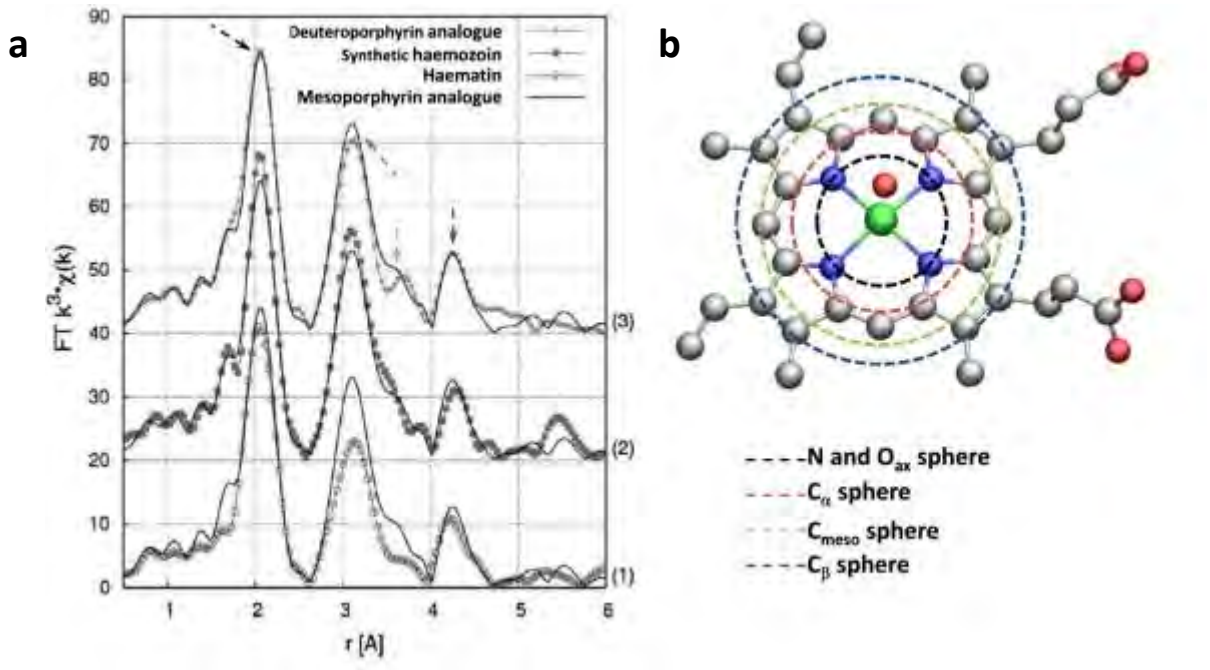


Figure 2.15. (**a**) Comparison of EXAFS spectra in *R*-space of haematin, synthetic haemozoin and its mesoporphyrin and deuteroporphyrin analogues. Figure adapted with permission from Walczak *et al*.[217] © 2010, Elsevier. (**b**) Coordination spheres giving rise to peaks in the EXAFS spectrum are labelled by arrows with corresponding colours in (a).

In subsequent studies, the same group used EXAFS spectroscopy to probe interaction of CQ with the mesoporphyrin analogue in DMSO, acetic acid and aqueous mixtures thereof.[217] All EXAFS spectra showed differences as compared to the mesoporphyrin analogue spectrum recorded in the solid state, particularly in the region around 3 Å (see Fig. 2.16). Qualitative analysis of the spectra, however, showed the dimer remained in present in DMSO but significant dissociation of carboxylate binding occurred in acetic acid and mixed-aqueous solutions. In the case of the latter two solutions, little change in structure was found upon addition of CQ, indicating that CQ did not bind in a similar manner as reported for aryl methanol antimalarial drugs, quinine and quinidine. Furthermore, based on the data obtained, these authors suggested that CQ instead disrupts solvation as well as axial ligand binding to the iron centre in acidic aqueous solution.[152] By contrast, in DMSO, evidence of complex formation was observed where three atoms of the CQ quinoline ring were fitted between 2.5 and 3.8 Å of the iron centre (see Fig. 1.18 in Section 1.4.1).[218]

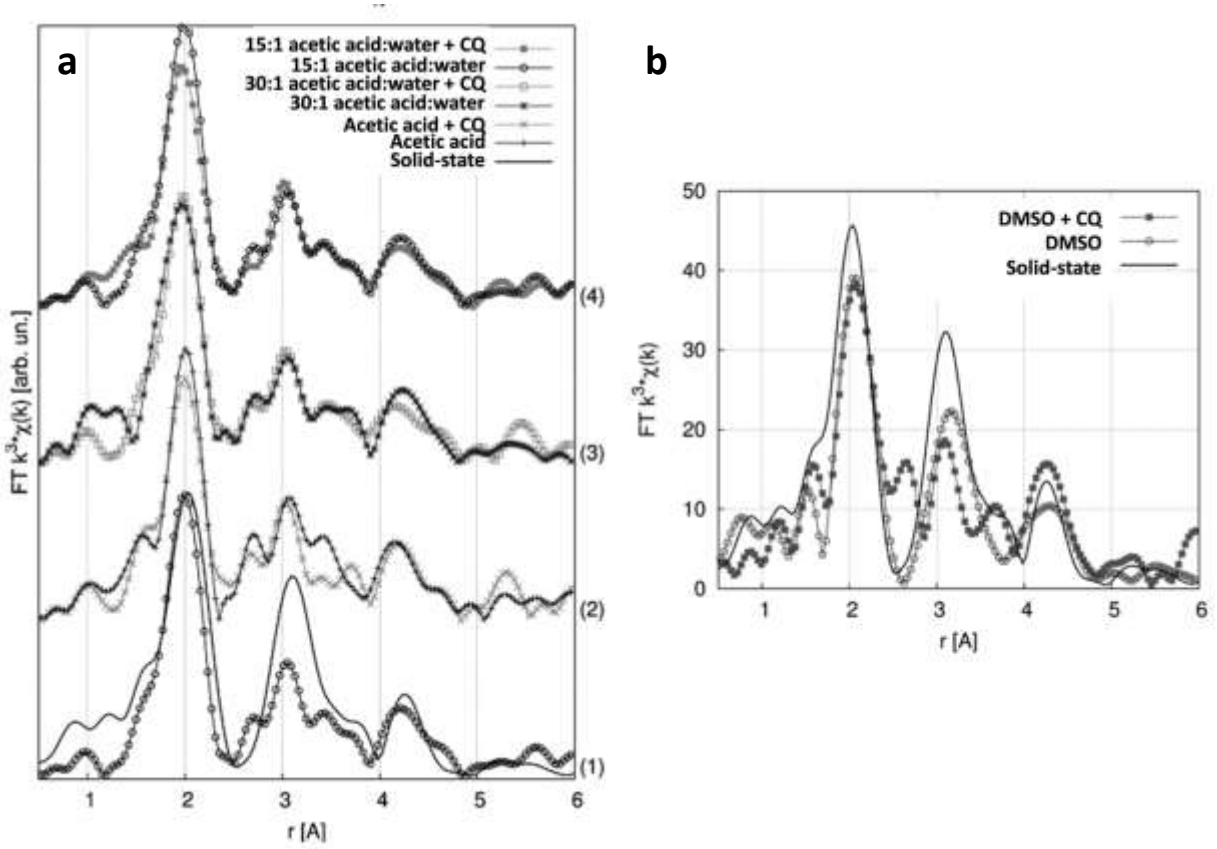


Figure 2.16. (**a**) EXAFS spectra of the mesoporphyrin analogue of synthetic haemozoin in the solid-state, acetic acid (with and without CQ) and acetic acid/water mixtures (with and without CQ). Figure adapted with permission from Walczak *et al*.[152] © 2011, American Chemical Society. (**b**) EXAFS spectra of the deuteroporphyrin analogue of synthetic haemozoin in the solid-state and in DMSO (with and without CQ). Figure adapted with permission from Walczak *et al*.[218] © 2011, American Chemical Society.

2.5 Magnetic Susceptibility

All substances possess magnetic properties. When placed in a magnetic field, the magnetic induction, *B*, is given by equation 2.5.1 where *H* is the strength of the field, $4\pi I$ the contribution by the substance in the field in which *I* is the intensity of magnetisation. Dividing eq. 2.5.1 by *H* gives eq. 2.5.2 where *P* is the permeability and κ is the magnetic susceptibility per unit volume. By convention, magnetic susceptibility is expressed as mass susceptibility, χ_g, and is calculated from eq. 2.5.3 where ρ is the density.[219] This quantity can be further related to the magnetic moment, μ_B, through eq. 2.5.4 where *MW* is the molecular weight of the material and *T* the temperature.

$$B = H + 4\pi I \quad (2.5.1)$$

$$P = 1 + 4\pi\kappa \quad (2.5.2)$$

$$\chi_g = \frac{\kappa}{\rho} \quad (2.5.3)$$

$$\mu_B = 2.84\sqrt{\chi_g \times MW \times T} = 2.84\sqrt{\chi_M \times T} \quad (2.5.4)$$

The value of χ_g (and hence κ and *P*) can be either negative or positive. In the case of the former, this represents a reduction in the density of lines of force while increased density is the cause of the latter (see Fig. 2.17).[219] Materials that produce $\chi_g < 0$ are termed diamagnetic and those that produce $\chi_g > 0$ are known as paramagnetic.

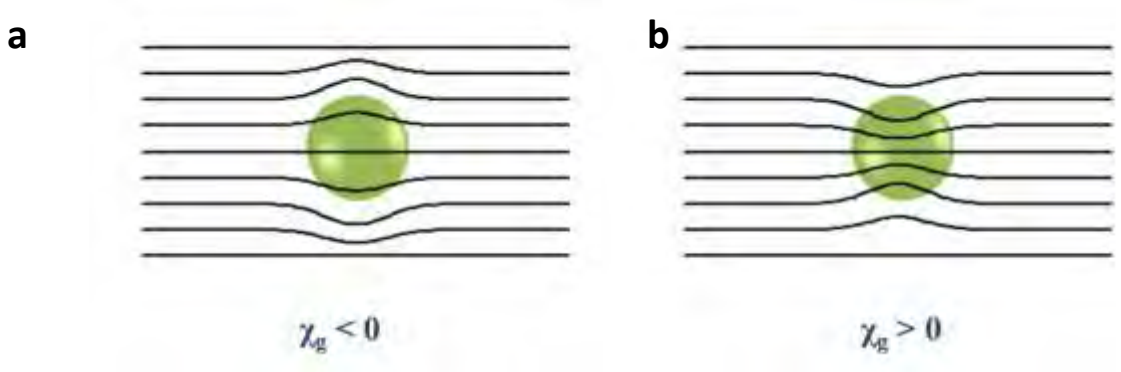


Figure 2.17. Effects on magnetic field (solid lines) of (**a**) diamagnetic and (**b**) paramagnetic materials (green spheres). Decreased field line density is observed in (a) while increased density is observed in (b).

All iron(III)porphyrins are paramagnetic (χ_g > 0) since they contain five valence d-electrons, however, the magnitude of χ_g varies depending on the spin-state of the molecule. Low- (S = $^1/_2$), intermediate- (S = $^3/_2$) and high-spin (S = $^5/_2$) states produce theoretical spin-only magnetic moments of 1.7, 3.9 and 5.9 μ_B respectively. Further modulation of χ_g can arise from inter- or intramolecular interactions of paramagnetic ions within a material in a process known as an "exchange" mechanism.[219] Substances that have electron spins aligned in the same direction increase χ_g and are termed ferromagnetic. Conversely, pairing of electron spins reduces the value of χ_g. This is termed antiferromagnetic coupling and is temperature dependent. At 0 K, complete spin pairing should occur, leading to diamagnetism. As the temperature is raised, however, thermal energy begins to overcome the magnetic interaction causing electron spin-decoupling and hence increases χ_g.[219, 220] The temperature-dependence of χ_g (in the form of μ_B) for μ-oxo dimeric ferrihaem is illustrated in Fig. 2.18 and is described by eq. 2.5.5.[220]

$$\chi_g = \frac{Ng^2\mu^2}{kT} \frac{e^{J/kT} + 5e^{3J/kT} + 14e^{6J/kT} + 30e^{10J/kT} + 55e^{15J/kT}}{1 + 3e^{J/kT} + 5e^{3J/kT} + 7e^{6J/kT} + 9e^{10J/kT} + 11e^{15J/kT}} \qquad (2.5.5)$$

Where *N* is Avogadro's number, *g* is the electron g-value, μ is the Bohr magneton (9.274×10^{-21} erg G^{-1}), *k* is the Boltzmann constant (1.381×10^{-16} erg K^{-1}) and *J* is the exchange coupling constant. Values for *g* and *J* were reported by Stanek and Dziedzic-Kocurek[220] as 2.1 ± 0.2 and -110 ± 15 cm^{-1}, respectively.

There has been a misconception in the literature regarding the magnetic properties of μ-oxo ferrihaem where the observed low magnetic moment has been attributed to a low-spin state.[170] This is incorrect. The temperature-dependence of χ_g clearly indicates antiferromagnetic coupling exists between iron centres, leading to a magnetic moment that is coincidentally similar but not identical to that expected for a low-spin complex at 298 K (2.0 vs. 1.7 μ_B, see dashed red line in Fig. 2.18).

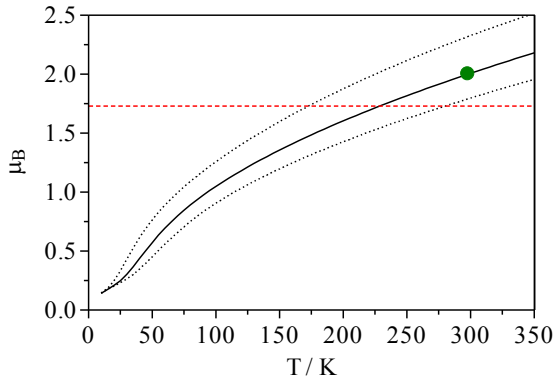


Figure 2.18. Predicted temperature dependence on the magnetic moment, μ_B, of μ-oxo ferrihaem determined from eq. 2.5.5 and the *J*-coupling constant of -110 ± 15 cm^{-1} reported by Stanek and Dziedzic-Kocurek.[220] Two standard deviations of the predicted dependence are depicted by dotted lines. Predicted values μ_B for antiferromagnetic and low-spin species are indicated by green circle and red dashed line respectively.

A variety of methods have been developed to determine the magnetic susceptibility of materials in both the solid- and solution-states. Examples of solid-state techniques include the superconducting quantum interference device (SQUID), Faraday and Gouy methods while solution-state techniques include the Quincke and Evans method.[219, 221, 222] Discussion will be limited to two techniques employed in this work, namely the Evans NMR method and Evans balance which is based on a modification of the Gouy method.

2.5.1 Solution-phase magnetic susceptibility – Evans NMR method.

Determination of magnetic susceptibility using NMR spectrometry was first described by Evans in 1959 and thus this technique is referred to as the Evans method. In a standard NMR spectrum, peaks obtained at a given chemical shift represent the Larmor frequency of precession by specific nuclei. The Larmor precession is sensitive to the chemical environment of the nuclei and consequently is dependent on the bulk susceptibility of the medium in which it is dissolved.[222] Introduction of paramagnetic ions to the bulk medium strongly influences the chemical environment and thus cause chemical shifts to alter. It is this difference in chemical shift between the reference and paramagnetic sample peaks from which the magnetic susceptibility of the paramagnetic substance can be obtained (see eq. 2.5.6).[223]

$$\chi_g = \frac{-3\Delta f}{4\pi f m} - \chi_0 \qquad (2.5.6)$$

Where χ_g is the mass magnetic susceptibility ($cm^3\ g^{-1}$), Δf the difference in chemical shift (Hz), f the NMR operating frequency (Hz) and m the molality of the paramagnetic sample ($g\ cm^{-3}$). The second term in eq. 2.5.6 represents the diamagnetic component of the mass magnetic susceptibility, χ_0. It should be noted that eq. 2.5.6 is often reported in which the second term is added to the first term using a positive value for χ_0. While this gives the correct answer, by convention diamagnetic susceptibility is reported as a negative value and thus it is more correct to subtract χ_0, as shown in eq. 2.5.6. The diamagnetic correction term is routinely approximated using Pascal's constants.[224]

2.5.2 Solid-state magnetic susceptibility – Evans balance.

As mentioned above, the Evans balance is a modification of the Gouy balance method. The latter involves measuring the change in mass of a sample in the presence and absence of a magnetic field.[219] The Evans balance employs the same equipment configuration as the Gouy method, however, the change in mass of the magnet in the presence and absence of a sample is measured instead.[225] In the case of the Evans balance, the change in mass of the magnet, ΔR, is related to χ_g of the sample by eq. 2.5.7 where L is the sample length (in cm), m the mass of the sample and C the balance calibration constant.[225] The value of C is determined using the standard mercury(II)tetrathiocyanatocobaltate(II) which has a known χ_g value of $16.44 \times 10^{-6}\ cm^3\ g^{-1}$.[226] Similar to that described for the solution-phase Evans NMR method, diamagnetic correction factors also need be taken into account.

$$\chi_g = \frac{C \times L \times \Delta R}{m \times 10^9} \qquad (2.5.7)$$

2.6 Density Functional Theory

DFT is a QM modelling method. QM modelling methods rely on solving the Schrödinger equation (eq. 2.6.1) for a particular molecule with wavefunction, ψ, in order to obtain the energy, *E*, of that wavefunction. Most QM methods approximate ψ as a LCAO and thus ψ is expressed in terms of MOs. By minimizing *E*, subject to constraints such as wavefunction antisymmetry and orthogonality of orbitals, the lowest energy MOs are produced from which a low energy molecular conformation can be obtained.

$$\hat{H}\psi = E\psi \qquad (2.6.1)$$

While most other QM modelling methods rely on determining the energy from ψ, DFT uses the electron density, ρ, which is related to the molecular wavefunction for an *n*-electron (closed shell) system through eq. 2.6.2.[227, 228] Calculating ρ instead of ψ greatly reduces the computational cost of the calculation as ρ is a function of position and is only dependent on three variables (x, y, z) whereas the wavefunction of a *n*-electron system is dependent on 4*n* variables, three spatial (x, y, z) and one spin coefficient (α or β) for each electron.[228]

$$\rho = \sum_{i=1}^{n} n_i |\psi_i|^2 \qquad (2.6.2)$$

DFT is founded on two theorems proposed by Hohenberg and Kohn.[229] The first theorem states that any ground state property of a molecule, such as energy, is a functional of the ground state electron density, ρ_0. The second theorem states that a trial electron density function will give an energy greater than or equal to the true ground state energy. These two theorems are expressed mathematically in eq. 2.6.3, where E_t and E_0 are the trial and true energies, F_t and F_0 the trial and true electron density functionals, and ρ_t and ρ_0 the trial and true electronic density, respectively.[228]

$$E_t = F_t[\rho_t] \geq E_0 = F_0[\rho_0] \quad (2.6.3)$$

It should be evident from eq. 2.6.3 that the true energy, E_0, of a system can be directly calculated provided the true energy functional, F_0, and true electronic density, ρ_0, are known. In reality, however, this is not possible as neither the correct true energy functional nor an accurate ρ is known.[228] The former issue is addressed by employing the Kohn-Sham (KS) approach which separates the electronic energy, E, into terms that can be accurately calculated without using DFT.[230] The unknown energy functional is reduced to a single term that does not contribute greatly to the overall energy, thus decreasing the error in the total energy. This is achieved as shown in eq. 2.6.4 where E is expressed as the sum of the nucleus-electron attraction energies, V_{Ne}, electron-electron repulsion energies, V_{ee} and electron kinetic energies, T, all three of which are functionals of the electronic density, ρ.[228]

$$E = V_{Ne}[\rho] + V_{ee}[\rho] + T[\rho] \quad (2.6.4)$$

The nucleus-electron attraction energy is described by eq. 2.6.5, where Z_A/r_{iA} is the potential energy of interaction of electron i with nucleus A at varying distance r, and $v(r)$ is the external potential for the attraction of electron i to all nuclei. If ρ is known, $V_{Ne}[\rho]$ is easily calculated.

$$V_{Ne}[\rho] = \int \left[\rho(r_i) \sum_{nuclei\ A} \left(-\frac{Z_A}{r_{iA}} \right) \right] dr_i = \int \rho(r)\, v(r) dr \quad (2.6.5)$$

In order to obtain equations which describe the electron-electron repulsion and kinetic energies (eq. 2.6.6 and 2.6.7, respectively), the introduction of a fictitious reference system of non-interacting electrons is required. By definition, the electron density of this reference system gives the same electron density as the real system.[228] The electron-electron repulsion and kinetic energy differences between the reference and real systems are $\Delta V_{ee}[\rho]$ and $\Delta T[\rho]$, respectively. The reference system electron-electron repulsion, $V_{ee}[\rho]_{ref}$, can be described classically as the sum of all repulsion energy between pairs $\rho(r_1)$ and $\rho(r_2)$, separated by distance r_{12}. Multiplication by one half prevents double counting. Thus if ρ is

known, $V_{ee}[\rho]_{ref}$ can be readily calculated. The kinetic energy of the reference system, $T[\rho]_{ref}$, is described by eq. 2.6.7 for a closed-shell system.[227] The ψ^{KS} terms are known as KS orbitals and are comprised of non-interacting electrons. In analogy with other QM modelling methods, the orbitals are described by a linear combination of basis functions which are provided by the user in a DFT calculation. Further discussion of basis functions is detailed in Section 2.6.3.

$$V_{ee}[\rho] = \Delta V_{ee}[\rho] + V_{ee}[\rho]_{ref} = \Delta V_{ee}[\rho] + \frac{1}{2}\iint \frac{\rho(r_1)\rho(r_2)}{r_{12}} dr_1 dr_2 \quad (2.6.6)$$

$$T[\rho] = \Delta T[\rho] + T[\rho]_{ref} = \Delta T[\rho] - \frac{1}{2}\sum_{i=1}^{2n} \langle \psi_i^{KS} | \nabla_i^2 | \psi_i^{KS} \rangle \quad (2.6.7)$$

Incorporating eq. 2.6.5 – 2.6.7 into 2.6.4, gives eq. 2.6.8, the last two terms of which are transformed by the unknown functional into energy, known as the exchange-correlation energy, E_{xc}.[228] Thus eq. 2.6.8 can also be rewritten as eq. 2.6.9. In practice, the unknown E_{XC} functional is approximated by a DFT functional which is supplied by the user for a given calculation, however, further details regarding these terms will be discussed in Section 2.6.2.

$$E = \int \rho(r)\, v(r) dr + \frac{1}{2}\iint \frac{\rho(r_1)\rho(r_2)}{r_{12}} dr_1 dr_2 - \frac{1}{2}\sum_{i=1}^{2n} \langle \psi_i^{KS} | \nabla_i^2 | \psi_i^{KS} \rangle + \Delta T[\rho] + \Delta V_{ee}[\rho] \quad (2.6.8)$$

$$E = \int \rho(r)\, v(r) dr + \frac{1}{2}\iint \frac{\rho(r_1)\rho(r_2)}{r_{12}} dr_1 dr_2 - \frac{1}{2}\sum_{i=1}^{2n} \langle \psi_i^{KS} | \nabla_i^2 | \psi_i^{KS} \rangle + E_{XC}[\rho] \quad (2.6.9)$$

In order to minimize the energy of the system, E, KS equations have been developed (eq. 2.6.10 – 2.6.12, written with respect to electron number 1). These were derived by relating ρ to KS orbitals through eq. 2.6.10, and differentiating E with respect to ψ^{KS}. In eq. 2.6.11, ε^{KS} is the KS energy levels, v_{XC} the exchange-correlation potential and the terms in square brackets are referred to as the KS operator, $\hat{h}^{KS}$.[228]

$$\rho = \rho_{ref} = \sum_{i=1}^{2n} \left|\psi_i^{KS}(1)\right|^2 \tag{2.6.10}$$

$$\left[-\frac{1}{2}\nabla_i^2 - \sum_{nuclei\ A} \frac{Z_A}{r_{iA}} + \int \frac{\rho(r_2)}{r_{12}} dr_2 + v_{XC}(1)\right] \psi_i^{KS}(1) = \varepsilon_i^{KS} \psi_i^{KS}(1) \tag{2.6.11}$$

$$v_{XC} = \frac{\delta E_{XC}[\rho(r)]}{\delta \rho(r)} \tag{2.6.12}$$

The KS equations are used to calculate the electron density, ρ, from which ε^{KS} and ψ^{KS} are calculated. By minimizing ε^{KS} and ψ^{KS}, the molecular energy E is thus indirectly minimized. A more detailed description of this minimization cycle, with reference to Fig. 2.19, is described below.

Step 1: A guess value of the electronic density, ρ, is made based on the initial starting structure.

Step 2: The exchange-correlation potential, v_{XC}, is calculated employing eq. 2.6.12 using the guessed value of ρ and the user defined functional E_{XC}.

Step 3: The KS operator, $\hat{h}^{KS}$, is calculated from eq. 2.6.11 using the values of v_{XC} and ρ.

Step 4: The KS operator, $\hat{h}^{KS}$, and user defined basis functions, φ, are used to obtain coefficients, c, and KS energy levels, ε^{KS}. The coefficients are used in the calculation of KS MOs, ψ^{KS}, according eq. 2.6.13, where m is the number of functions in the basis set.

$$\psi_i^{KS} = \sum_{s=1}^{m} c_{si}\varphi_s \qquad i = 1, 2, 3 \ldots m \tag{2.6.13}$$

Step 5: Comparison of ε^{KS} and ψ^{KS} values calculated in step 4 with values calculated in the previous cycle is made. In the case of the first cycle, no comparison is made since there are no previously calculated values. Should the difference between the two values not be within a specified tolerance, the cycle continues to step 6.

Step 6: A new electronic density, ρ, is calculated from ψ^{KS} using eq. 2.6.10.

Step 7: The improved electronic density guess, ρ, is used to calculate a new exchange-correlation potential, v_{XC} using eq. 2.6.12.

The cycle is then repeated until the difference between two successive ψ^{KS} and ε^{KS} values is below a set tolerance, after which the molecular energy, E, is calculated using eq. 2.6.9 in step 8. The minimized geometry of the system can then be determined by taking the second derivative of E with respect to geometric parameters (e.g. bond lengths and angles).[228]

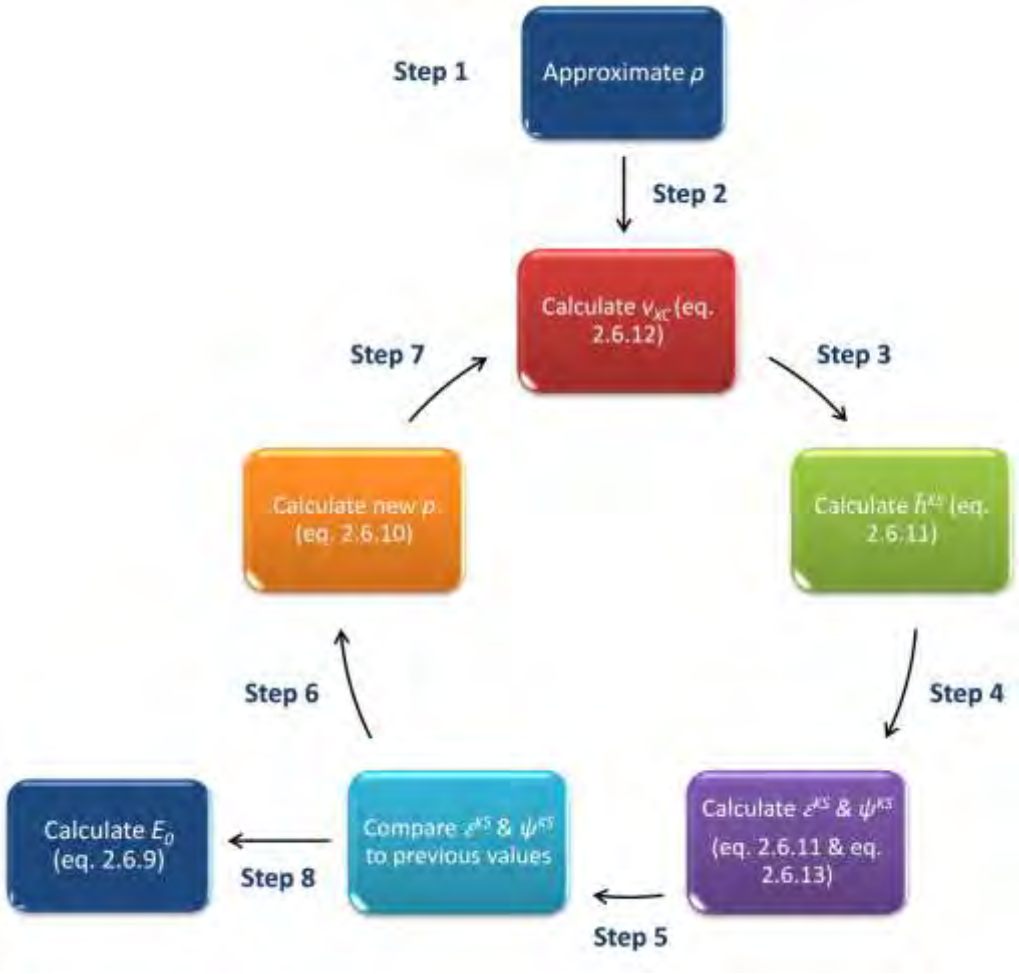


Figure 2.19. Diagrammatic representation of the DFT energy minimization procedure. Information regarding each step is detailed in the text.

2.6.1 Time-dependent DFT.

While DFT has been successfully applied to ground state systems, the formalism of this method is not appropriate to describe time-dependent phenomena or excited states.[231] Instead, the related method known as TD-DFT overcomes the limitations of conventional DFT in this regard. The basis for TD-DFT involves the introduction of an external time-dependent potential, $v_{ext}(r,t)$, to the system.[232, 233] Incorporation of $v_{ext}(r,t)$ thus requires application of the time-dependent Schrödinger equation, eq. 2.6.14 (see eq. 2.6.1 for the corresponding time-independent form).

$$\hat{H}(r,t)\psi(r,t) = i\frac{\partial}{\partial t}\psi(r,t) \qquad (2.6.14)$$

To solve eq. 2.6.14, time-dependent analogues of the two Hohenberg-Kohn theorems which form the foundation of DFT (see Section 2.6 above) were developed by Runge and Gross.[234] These theorems relate the time-dependent energy of a system as a functional of the time-dependent electron density, $\rho(r,t)$, and show that for a given external potential, a unique solution to $\rho(r,t)$ is obtained. As in the case of conventional DFT, the true functional is unknown. Thus a time-dependent form of the KS approach is used to determine the time-dependent energy of the system and leads to eq. 2.6.15 which is analogous to the time-independent form in eq. 2.6.11.[232] The first term of the Hamiltonian describes the kinetic energy of the electrons while the second term is the effective potential, $v_{eff}(r,t)$. The latter encompasses three terms (see eq. 2.6.16), the first being the external potential (which includes the nucleus-electron attraction potential), the second is the electron-electron repulsion and the third is the exchange-correlation potential. The adiabatic local density approximation is applied to $v_{XC}(r,t)$, the basis of which is that the density varies only slowly with time and thus allows conventional DFT functionals to be used with TD-DFT.[231, 232]

$$\left[-\frac{1}{2}\nabla^2 + v_{eff}(r,t)\right]\psi^{KS}(r,t) = i\frac{\partial}{\partial t}\psi^{KS}(r,t) \qquad (2.6.15)$$

$$v_{eff}(r,t) = v_{ext}(r,t) + \int \frac{\rho(r_2,t)}{r_{12}}dr_2 + v_{XC}(r,t) \qquad (2.6.16)$$

2.6.2 Density Functionals.

As mentioned above, a major complication in DFT is that the true form of the exchange-correlation functional is not known but is necessary in calculating the energy of a system (see eq. 2.6.9). In order to overcome this obstacle, various density functionals have been developed as an approximation to the unknown term and it is this approximation of the true functional that introduces the main source of error in a DFT calculation.[235] The general approach to deriving the density functionals is to separate the exchange and correlation energies. Since both energy terms are functionals of the electron density, density functionals are comprised of an exchange functional which describes same-spin electron interactions and a correlation functional which relates to interactions between electrons of opposite spin.[228, 235] For example, in the well-known density functional B3LYP, B3 refers to the Becke 88 3-parameter exchange functional[236] and LYP to the correlation functional developed by Lee, Yang and Parr.[237]

Density functionals can be classified into different types based on the methods used in their development. The accuracy of the different types in relation to each other has often been described using the biblical concept of Jacob's ladder, where Hartree-Fock theory represents the Earth and the true density functional (and thus chemical accuracy) represents Heaven. There have been various interpretations of this description in the literature[228, 238], but according to the interpretation of Sousa *et al.*[235], density functionals can be classed into five types (see figure 2.20) and are listed below from lowest to highest rung: (i) local density approximation (LDA)/local spin density approximation (LSDA); (ii) generalised gradient approximation (GGA); (iii) meta-GGA; (iv) hybrid-GGA and hybrid meta-GGA; and (v) fully non-local theory. As expected, the complexity of these methods increases from lower to higher rungs, the details of which are briefly mentioned below. At present, full non-local theory is not yet practically applicable to molecular calculations and thus will not be discussed.

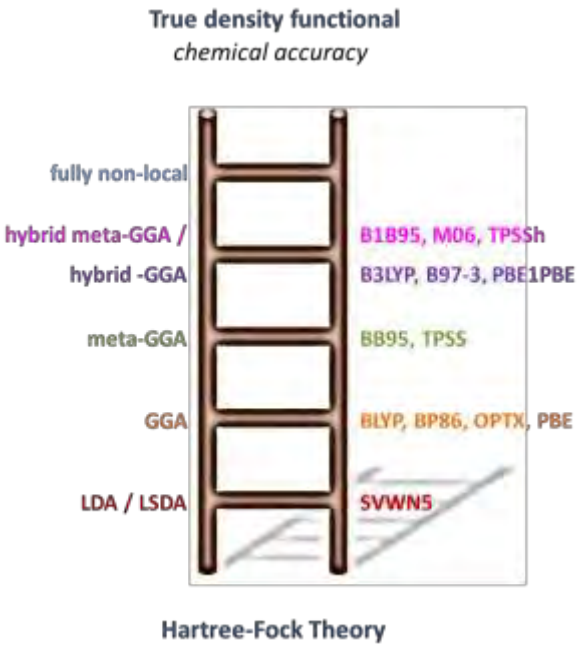


Figure 2.20. Jacobs ladder as described by Sousa *et al*.[235] for the development of DFT functionals. Example functionals are listed on the right.

The first class relates the exchange-correlation energy only to the electron density, ρ, at a specific point and makes the assumption that the energy at this specific point is the same as in a homogenous electron gas with the same ρ.[150] LSDA differs from LDA in that the electron density is separated for α- and β-spin electrons giving ρ^{α} and ρ^{β} respectively, allowing calculation of open-shell systems.[228, 235] LDA and LSDA functionals are often used in solid-state physics as the density of solids tend to be close to homogenous but higher-rung functionals have largely replaced these functionals for molecular calculations where electron density is more inhomogeneous.[150, 228]

The second class, GGA, is similar to the first (LDA/LSDA) where the exchange-correlation energy is dependent on ρ, however differs in that the gradient of electron density, $\nabla\rho$, is also taken into account.[228] Incorporation of $\nabla\rho$ better describes the inhomogeneous electron density of molecules and thus GGA functionals have been observed to perform better than LDA/LSDA functionals in calculating total energies, atomization energies, structural energy differences and energy barriers.[150, 235]

Because this inclusion of gradient corrections leads to greater accuracies in calculated properties, the third class of functionals, meta-GGA, goes one step further and also relates the exchange-correlation energy to the second derivative of electron density, $\nabla^2\rho$.[228] These functionals have been shown to produce more accurate atomization energies.[235]

The fourth class of functionals differs from their lower-rung counterparts in that a weighted percentage of Hartree-Fock (HF) exchange is included. The percentage weighting factor is generally fitted semiempirically by optimizing a set of coefficients to experimental data for a representative set of molecules. In the case of hybrid-GGA functionals, the HF exchange is included in GGA functionals and for hybrid meta-GGA functionals, HF exchange is included in meta-GGA functionals. While these functionals generally provide more accurate results for molecular properties, their use in solid-state physics is less successful.[235]

2.5.3 Basis sets.

As detailed in Section 2.6 above, calculation of the energy of a system is reliant on a description of the molecular wavefunction, ψ. This description is obtained through a LCAO which, in turn, is described by user-defined basis sets. Basis sets are mathematical functions which approximate atomic orbitals. As shown in eq. 2.6.17, the linear combinations of these functions, φ, are used to describe MOs, ψ.[227, 228] One approach, developed by Slater, makes use of basis functions of the general form in eq. 2.6.18, where *a* is the normalization constant, *b* the orbital exponent and *r* the radius (Å).[227, 239]

$$\psi = c_1\varphi_1 + c_2\varphi_2 + c_3\varphi_3 + \cdots + c_i\varphi_i \quad (2.6.17)$$

$$\varphi = ae^{(-br)} \quad (2.6.18)$$

$$\varphi = ae^{(-br^2)} \quad (2.6.19)$$

While these Slater-type orbitals (STO) are good approximations to atomic wavefunctions, they are not routinely used as calculation of certain two-electron integrals is computationally expensive. To reduce this computational cost, a second approach was developed where orbitals were described using Gaussian functions (eq. 2.6.19) and hence these are known as Gaussian-type orbitals (GTO). While a single Gaussian function is more mathematically tractable than the Slater function, the former is a poorer approximation to the atomic wavefunction than the latter.[228] Figure 2.21a illustrates the difference between the GTO approximating a STO using a single Gaussian (STO-1G) and the STO described by the Slater function using eq. 2.6.18.[228] To improve agreement with the STO, several Gaussian functions are used instead of just one, forming a contracted basis function with the general form described by eq. 2.6.20. An example of such a contracted basis function is STO-3G which is comprised of three primitive Gaussian functions (see figure 2.21b) each having a contraction coefficient, *d*. As can be seen from figure 2.21, this contracted basis function is a much better approximation to the STO described by the Slater function.

$$\varphi_i = d_1 e^{-b_1 r^2} + d_2 e^{-b_2 r^2} + d_3 e^{-b_3 r^2} \qquad (2.6.20)$$

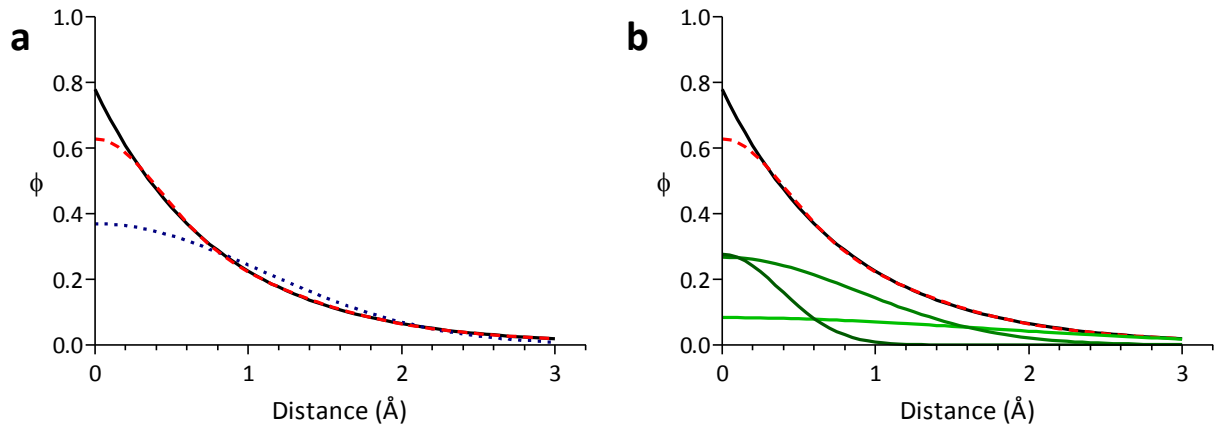


Figure 2.21. (**a**) Slater (solid black line), STO-1G (dotted blue line) and STO-3G (red dashed line) functions for a hydrogen atom. (**b**) Decomposition of the contracted STO-3G function into its three primitive Gaussian constituents (green solid lines).

In practice, the above mentioned GTO basis set, STO-3G, is the smallest that is used in *ab initio* calculations and is termed a minimal basis set. Minimal basis sets have one basis function for every atomic orbital required in describing a given atom.[227] For example, a hydrogen atom will have one basis function representing the 1s orbital, while a carbon atom is described by five basis functions representing the 1s, 2s and three 2p orbitals. It should be noted that atoms in a specific row are described by the same number of basis functions regardless of whether the orbitals are occupied or not. For example, in row 2 of the Period Table, lithium is described by the same five basis functions as carbon even though the 2p orbitals of the former are unoccupied.[228]

While these minimal basis sets have the advantage of low computational cost, they produce somewhat inaccurate molecular geometries and energies. A second class of basis set, known as split-valence basis sets, generally has an improved accuracy over minimal basis sets which has made them one of the most popularly employed in DFT calculations. The difference between the minimal and split-valence basis sets is that the latter separates the valence orbitals into parts. This separation makes for a more realistic description of the electron distribution as a greater flexibility in adjusting contributions to the molecular orbitals is possible. Basis sets which further separate core orbitals in two ("double zeta basis sets"), three ("triple zeta basis sets") or more parts have been also been developed although are more computationally expensive than the split-valence counterparts.[228]

An example of a split-valence basis set is 6-31G in which the valence shell is separated into an inner and outer part. Thus a hydrogen atom described by this basis set will have two basis functions, 1s' and 1s'' (inner and outer), and a carbon atom will have nine basis functions representing the core (one 1s), inner valence shell (one 2s' and three 2p') and outer valence shell (one 2s'' and three 2p''). In this basis set, the core orbitals of an atom are each described by one contracted basis function comprised of six primitive Gaussian functions, orbitals of the inner shell each described by one basis function comprised of three primitive Gaussian functions and orbitals of the outer shell each described by one Gaussian function. An example of a larger split-valence basis set is 6-311G which is similar to 6-31G except in that the valence shell is separated into three parts.

To further increase accuracy, but at the expense of speed, split-valence basis sets are often supplemented with diffuse and polarisation functions. Polarisation functions introduce additional functions describing unoccupied orbitals, allowing electron delocalisation to be polarized along a particular direction and increasing the flexibility in varying contributions to MOs.[227, 228] For example, for atoms in row two of the periodic table which are described by the 6-31G basis set, in addition to the nine basis functions representing 1s, 2s', 3x2p', 2s'' and 3x2p'', a further six functions representing d-orbitals are also present. In this case, the inclusion of d-orbitals with the 6-31G basis set is specified by adding "*" or "(d)" to the end, i.e. 6-31G* or 6-31G(d). Addition of valence f-orbitals to atoms in row two and below or p-orbitals to hydrogen is also possible and can be done in an analogous manner as described above using 6-31G(df,p) for example.

Diffuse functions are Gaussians which fall off very slowly and are added to better describe the behaviour of valence electrons not involved in bonding.[228] Basis sets which included diffuse functions are identified by "+", such as in 6-31+G, and generally introduce one basis function comprised of a single Gaussian for every valence orbital. For example, the above basis set contains a total of thirteen basis functions for an atom in row two, nine representing 1s, 2s', 3x2p', 2s'' and 3x2p'', and four diffuse functions, 2s+ and 3x2p+.

Table 2.6.1 lists a selection of basis sets and the total number of basis functions present for a given group of elements. In most cases, the basis sets contain a greater number of basis functions than the number of occupied atomic orbitals for a particular atom. For example, in the case of a carbon atom described by the 6-31G basis set, nine basis functions effectively give rise to nine atomic orbitals, only three of which will be filled. The remaining six orbitals are unoccupied (virtual) atomic orbitals. Similarly, unoccupied molecular orbitals are produced for molecules described by this type of basis set.

Table 2.6.1. Basis set functions describing atoms using STO-3G, 6-31G, 6-31G* and 6-31+G* basis sets.

	STO-3G	**6-31G**	**6-31G***	**6-31+G***
H – He	1 x 1s	1 x 1s’	1 x 1s’	1 x 1s’
	1 function	1 x 1s’’	1 x 1s’’	1 x 1s’’
		2 functions	*2 functions*	*2 functions*
Li – Ne	1 x 1s	1 x 1s	1 x 1s	1 x 1s
	1 x 2s, 3 x 2p	1 x 2s’, 3 x 2p’	1 x 2s’, 3 x 2p’	1 x 2s’, 3 x 2p’
	5 functions	1 x 2s’’, 3 x 2p’’	1 x 2s’’, 3 x 2p’’	1 x 2s’’, 3 x 2p’’
		9 functions	6 x 3d	6 x 3d
			15 functions	1 x 2s+, 3 x 2p+
				19 functions
Na – Ar	1 x 1s	1 x 1s	1 x 1s	1 x 1s
	1 x 2s, 3 x 2p	1 x 2s, 3 x 2p	1 x 2s, 3 x 2p	1 x 2s, 3 x 2p
	1 x 3s, 3 x 3p	1 x 3s’, 3 x 3p’	1 x 3s’, 3 x 3p’	1 x 3s’, 3 x 3p’
	9 functions	1 x 3s’’, 3 x 3p’’	1 x 3s’’, 3 x 3p’’	1 x 3s’’, 3 x 3p’’
		13 functions	6 x 3d	6 x 3d
			19 functions	1 x 3s+, 3 x 3p+
				23 functions
K – Ca	1 x 1s	1 x 1s	1 x 1s	1 x 1s
	1 x 2s, 3 x 2p	1 x 2s, 3 x 2p	1 x 2s, 3 x 2p	1 x 2s, 3 x 2p
	1 x 3s, 3 x 3p	1 x 3s, 3 x 3p	1 x 3s, 3 x 3p	1 x 3s, 3 x 3p
	1 x 4s, 3 x 4p	1 x 4s’, 3 x 4p’	1 x 4s’, 3 x 4p’	1 x 4s’, 3 x 4p’
	13 functions	1 x 4s’’, 3 x 4p’’	1 x 4s’’, 3 x 4p’’	1 x 4s’’, 3 x 4p’’
		17 functions	6 x 3d	6 x 3d
			23 functions	1 x 4s+, 3 x 4p+
				27 functions
Sc – Zn	1 x 1s	1 x 1s	1 x 1s	
	1 x 2s, 3 x 2p	1 x 2s, 3 x 2p	1 x 2s, 3 x 2p	
	1 x 3s, 3 x 3p	1 x 3s, 3 x 3p	1 x 3s, 3 x 3p	
	1 x 4s, 3 x 4p	1 x 4s’, 3 x 4p’	1 x 4s’, 3 x 4p’	
	5 x 3d	1 x 4s’’, 3 x 4p’’	1 x 4s’’, 3 x 4p’’	
	18 functions	6 x 3d’	6 x 3d’	
		6 x 3d’’’	6 x 3d’’’	
		29 functions	*29 functions*	
Ga – Kr	1 x 1s	1 x 1s	1 x 1s	
	1 x 2s, 3 x 2p	1 x 2s, 3 x 2p	1 x 2s, 3 x 2p	
	1 x 3s, 3 x 3p	1 x 3s, 3 x 3p	1 x 3s, 3 x 3p	
	1 x 4s, 3 x 4p	1 x 4s’, 3 x 4p’	1 x 4s’, 3 x 4p’	
	5 x 3d	1 x 4s’’, 3 x 4p’’	1 x 4s’’, 3 x 4p’’	
	18 functions	6 x 3d	6 x 3d	
		23 functions	*23 functions*	

As the number of basis functions in a basis set increases, the time required to complete the *ab initio* calculation increases as well. This is a particular problem for heavier atoms (K – Kr and below) because there are a large number of core electrons. While these core electrons are somewhat chemically unimportant, they are required to properly describe valence orbitals through electron-electron repulsion interactions. In an effort to reduce the time required for calculations involving heavy atoms, effective core potential (ECP) basis sets were developed. These basis sets replace the core orbitals with a potential, $U_{ECP}(r)$, which mimics the average effect of the core electrons on the valence electrons. $U_{ECP}(r)$ can be described, for example, by Gaussian functions of the form in eq. 2.6.21.[228, 240]

$$U_{ECP}(r) = \sum_i a_i r^{n_i} e^{-\alpha_i r^2} \qquad (2.6.21)$$

Here, a_i, n_i and α_i are parameters which are optimised using least squares fitting and are dependent on angular momentum (s-, p-, d-, etc.). As expected, the larger the number of Gaussian functions used, the more accurate the description of the potential becomes but a corresponding increase in computational time also occurs. In general, between two and seven Gaussian functions are used.[240] Since these are much fewer than the number of basis functions usually used to describe core electrons, a significant reduction in computing time is obtained for ECPs in comparison to all-electron basis sets. An example of a popular ECP basis set is LANL2DZ.[241-244]

2.7 Molecular Mechanics Modelling

MM modelling is a computational technique in which the electronic motions of atoms are ignored.[245] Molecules are instead approximated classically as balls attached via springs and energies calculated using parametric functions parameterised to reproduce experimental or higher-level theory computations.[240] These parameterised functions are known as force fields and form the basis of MM modelling (see Section 2.7.1). By removing calculations relating to electrons (i.e. the wavefunction), a substantial decrease in required computing time is produced. Consequently, modelling of larger systems than that possible using *ab initio* is feasible. In addition, by ignoring electronic motions, dynamics of an atom or molecule can be calculated using classical mechanics (see Section 2.7.2).[240, 245] A number of software packages are available to perform MM or MD simulations, each having different means of calculating energies, etc. For the purposes of this chapter, explanation of this technique will be focused on how the software package CHARMM treats calculations since this was the program used throughout this thesis for MM/MD modelling.[246]

2.7.1 Force fields.

The MM energy of a system, E_{sys}, is calculated by summing individual terms relating to bonded and non-bonded interactions (see eq. 2.7.1 and Fig. 2.22). In the case of the former this comprises of functions used to describe energies associated with bond stretching (E_{bonds}), bending (E_{angles}), rotation around single bonds ($E_{torsions}$) and out of plane bending motions ($E_{impropers}$). The non-bonded energy term includes electrostatic (E_{elec}) and van der Waals (E_{VDW}) interactions. Functions used to approximate each energy term are described by eq. 2.7.2.[247, 248]

$$E_{sys} = E_{bonds} + E_{angles} + E_{torsions} + E_{impropers} + E_{elec} + E_{VDW} \tag{2.7.1}$$

$$E_{sys} = \sum_{bonds} K_b(b - b_0)^2 + \sum_{angles} K_\theta(\theta - \theta_0)^2 + \sum_{dihedrals} K_\Phi(1 + \cos(n\Phi - \delta)) + \sum_{\substack{improper \\ dihedrals}} K_\varphi(\varphi - \varphi_0)^2 + \sum_{non-bonded} \left(\frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}\right) + \varepsilon_{ij}^{min}\left[\left(\frac{R_{min,ij}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{min,ij}}{r_{ij}}\right)^{6}\right] \tag{2.7.2}$$

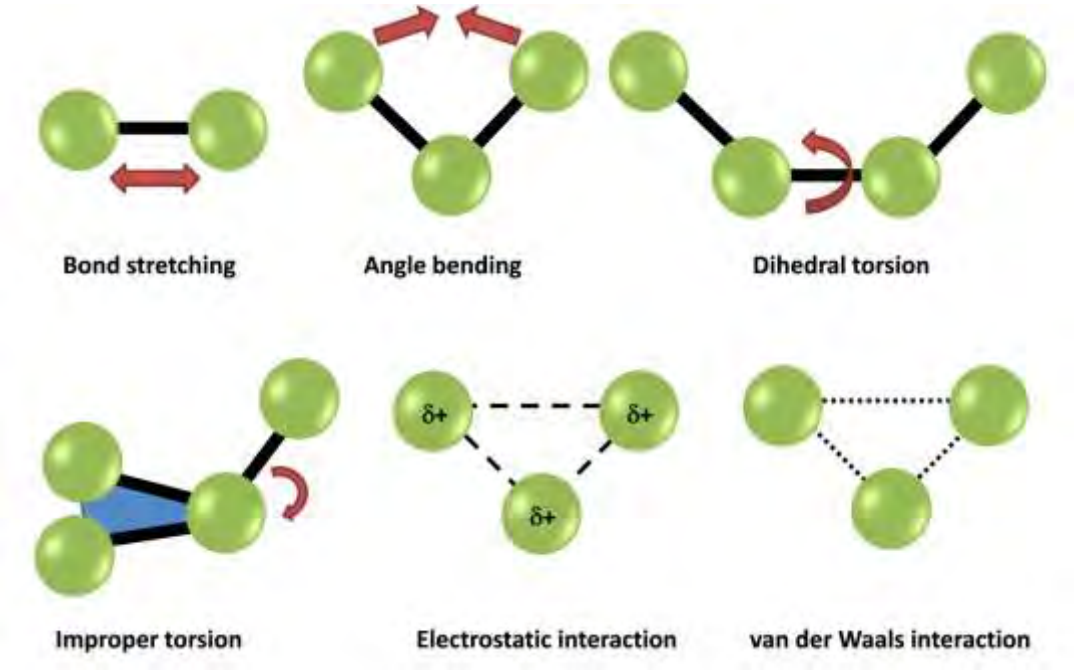


Figure 2.22. Bonded and non-bonded interactions contributing to the MM energy calculated using eq. 2.7.2.

Symbols highlighted red in eq. 2.7.2 represent values that must be parameterised for each type of interaction (see below for details regarding these values and how parameterisation is conducted). It is important to note that each term is specific to a particular interaction and must be parameterised separately. For example, the C-C bond stretching parameters will be different from C-H and C=C values. To distinguish between different type of bonds, angles, etc. and to ensure the correct parameters are used to describe interactions, atoms are given different labels. For example, aromatic carbon atoms in benzene are labelled CA in CHARMM, vinyl carbon atoms CE1 and the pyrrole carbons in the haem moiety CPA and CPB for the C_α and C_β positions respectively (see Fig. 2.14 for positions). While each type of interaction requires a particular set of parameters, often parameters are transferable. This transferability is important in MM calculations as it can circumvent the need to parameterise new molecules.

Bond stretching and angle bending. The potential energy curve for a hypothetical bond between two atoms is shown in Fig. 2.23. Since bonds don't deviate significantly from their equilibrium positions (well minimum in Fig. 2.22), in MM modelling they are usually approximated using Hooke's law.[245] This is described by eq. 2.7.3 where K_b is the bond stretching constant (force constant) and b_0 the reference bond length. These two values are parameterised to reproduce equilibrium bond lengths (obtained from crystal structures or higher-level *ab initio* calculations) and/or bond stretching frequencies from infrared data.[247] Since angles similarly do not deviate markedly from their equilibrium values, the same Hook's Law function is employed (eq. 2.7.4). In this case, K_θ and θ_0 represent the force constant and reference angle.[240]

$$E_{bond} = K_b(b - b_0)^2 \quad (2.7.3)$$

$$E_{angle} = K_\theta(\theta - \theta_0)^2 \quad (2.7.4)$$

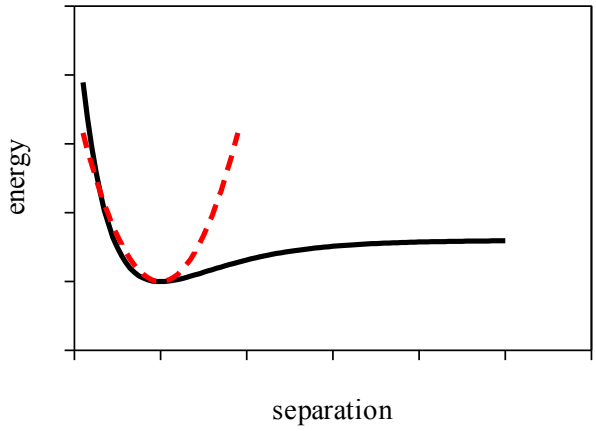


Figure 2.23. Potential energy curve for a hypothetical bond (solid black line) as the distance between two atoms is varied. Hook's law is used to approximate bond stretching in MM calculations (red dashed line).

Torsion and improper dihedrals. The energy associated with rotation around a bond is often characterised by a series of maxima and minima. For example, in the case of ethane, staggered conformations corresponding to ±60 and 180° are lower in energy than eclipsed conformations at 0 and ±120° (see Fig. 2.24).[245] To describe this torsion energy profile, MM employs the function in eq. 2.7.5, where K_Φ is the torsional force constant, n the multiplicity and δ the phase factor. The number of minimum points is given by n while δ determines the angles at which minima are obtained.[245] These parameters are usually optimised to reproduce torsion energies obtained from higher-level *ab initio* calculations.[247] Improper dihedrals describe out-of-plane bending motions and consist of four atoms. The nomenclature used to represent these dihedrals, for example in formaldehyde, is C*-H-H-O where the asterisk indicates the central atom to which all others are bonded. These improper dihedrals are described by eq. 2.7.6 with K_φ and φ_0 the force constant and reference angle respectively.

$$E_{torsion} = K_\Phi[1 + \cos(n\Phi - \delta)] \quad (2.7.5)$$

$$E_{improper} = K_\varphi(\varphi - \varphi_0)^2 \quad (2.7.6)$$

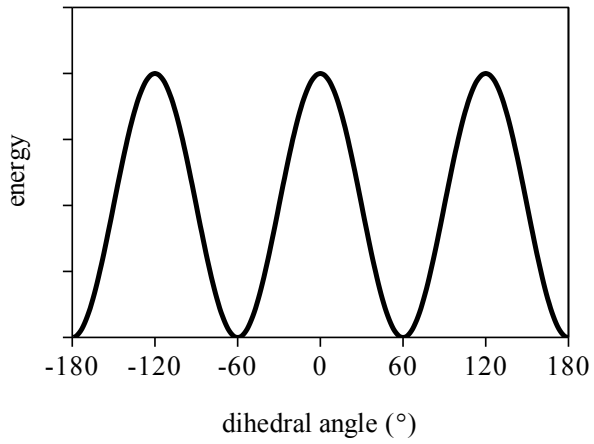


Figure 2.24. Hypothetical potential energy curve for rotation around the C-C bond in ethane.

Van der Waals interactions. Non-covalent interactions which are not electrostatic in nature make up van der Waals interactions. These interactions occur between atoms and molecules and are negligible at long distances.[245] As separation between atoms/molecules is decreased there is an increase in attractive force which reaches a minimum at a particular distance, R_{min}. Decreasing the separation further produces higher energies because of Pauli repulsion between electrons and internuclear repulsion. This is shown graphically in Fig. 2.25a. MM calculations make use of the Lennard-Jones 12-6 function (eq. 2.7.7) to describe these interactions, where R_{min} is the distance at which the energy is a minimum and ε_{min} is the well depth. The two terms comprising eq. 2.7.7 involving r^{-12} and r^{-6} describe the repulsive and attractive energies respectively(see Fig. 2.25b).[240, 245]

$$E_{VDW} = \varepsilon_{min,ij}\left[\left(\frac{R_{min,ij}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{min,ij}}{r_{ij}}\right)^{6}\right] \qquad (2.7.7)$$

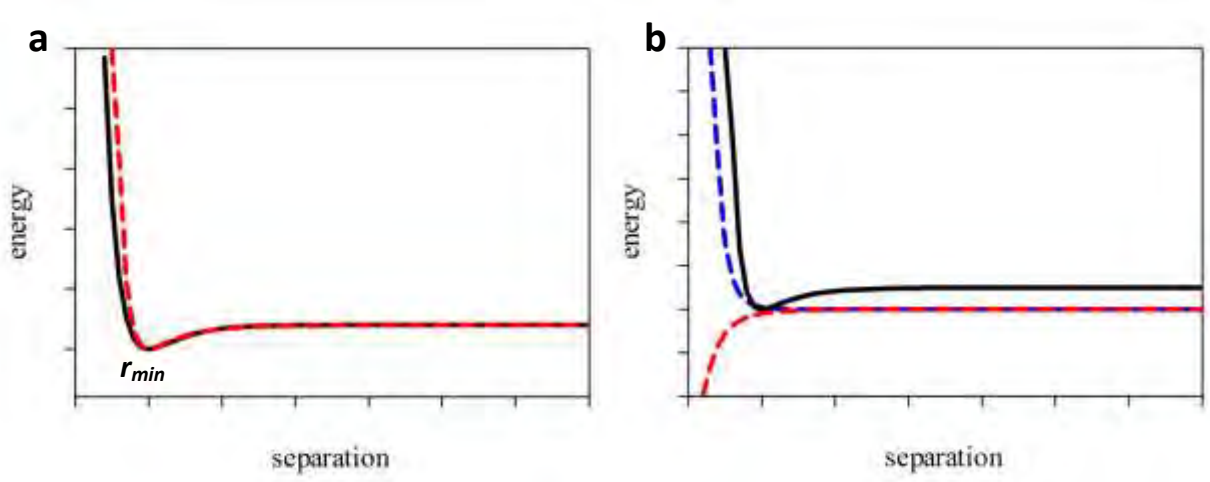


Figure 2.25. (**a**) Van der Waals energy as a function of distance between two atoms or molecules (solid line). This is approximated using the Lennard-Jones 12-6 function (eq. 2.7.7) and produces the red dashed curve. (**b**) The Lennard-Jones function (solid black line) is comprised of attractive (red dashed line) and repulsive (blue dashed line) components which are represented by the r^{-6} and r^{-12} terms in eq. 2.7.7.

Electrostatic interactions. Unequal distribution of electrons within a molecule gives rise to positive and negative regions.[245] The differences in electron distribution can be modelled in MM by assigning point charges to atoms within the molecule such that the sum of all charges reproduces the entire charge on the molecule. The interaction energy between these point charges (or atomic charges) is calculated using Coulomb's law (eq. 2.7.8), where q_i and q_j are the point charges on atoms *i* and *j* respectively, ε_0 is the permittivity of vacuum and r_{ij} the distance between atoms *i* and *j*.[240, 245] Assignment of atomic charges can be approached in a variety of ways. The CHARMM parameterisation protocol requires atomic charges to reproduce scaled dipole moments determined using higher-level *ab initio* calculations.[247, 248] In addition, reproduction of QM interaction energies between atoms comprising a molecule and a water molecule (see Fig. 2.26) by MM atomic charges is routinely employed.[247] A further criterion can also be enforced in the form of electrostatic potentials. These are defined as the forces acting on a unit of positive charge at a particular point.[245] MM atomic charges should reproduce molecular electrostatic potentials determined from QM calculations. This can be achieved using least squares fitting according to eq. 2.7.9, where w_i is a weighting factor that can be applied to different atoms as a measure of importance, φ_i^{QM} the QM-derived electrostatic potential at a point and φ_i^{MM} the MM-derived electrostatic potential. The value of *R* is minimised by changing atomic charges until the best-fit is obtained.[245]

$$E_{elec} = \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \tag{2.7.8}$$

$$R = \sum_{i=1}^{N_{points}} w_i\left(\varphi_i^{QM} - \varphi_i^{MM}\right)^2 \tag{2.7.9}$$

Figure 2.26. Water interaction between atoms of 4-aminopyridinium used in parameterisation of atomic charges.

2.7.2 Molecular dynamics.

The expressions described above allow for the calculation of the energy of a molecule without explicitly taking into account electronic motion. This drastically reduces the computing power required and enables calculation of time-dependent positions and momenta of molecules.[245] These types of time-dependent calculations are termed molecular dynamics simulations. Configurations of molecules are calculated by integrating Newton's second law (see eq. 2.7.10). The result is a trajectory which describes how the position and velocity of a particle changes with time.[240, 245]

$$F = ma \tag{2.7.10}$$

$$-\frac{dV_i}{dr_i} = m_i \frac{d^2 r_i}{dt^2} \tag{2.7.11}$$

Owing to the large number of particles with motions coupled to one another, analytical integration of Newton's second law is not possible. Instead, the integration is approached using the finite difference method in which it is divided up into small parts separated by a fixed time interval, Δt.[245] The force, F, at each time step, t, for a particle, i, can be calculated from eq. 2.7.11, where V is the potential energy at position r. The acceleration, a, of the particle can thus be calculated from F and combined with the velocity, v, and position, r, at the current time step to determine the new position and velocity at $t+\Delta t$. Forces at this new position can then be calculated and thus velocity and position at $t+2\Delta t$ can be determined, etc. Integration by the finite difference method can be achieved using a number of algorithms, all of which approximate position, velocity and acceleration by a Taylor series expansion (see eq. 2.7.12 – 2.7.14).[240] In these expressions, v is the first derivative of positions with respect to time, a the second derivative, b the third, etc.

$$r(t+\Delta t) = r(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 + \frac{1}{6}b(t)\Delta t^3 + \frac{1}{24}c(t)\Delta t^4 + \cdots \tag{2.7.12}$$

$$v(t+\Delta t) = v(t) + a(t)\Delta t + \frac{1}{2}b(t)\Delta t^2 + \frac{1}{6}c(t)\Delta t^3 + \cdots \tag{2.7.13}$$

$$a(t+\Delta t) = a(t) + b(t)\Delta t + \frac{1}{2}c(t)\Delta t^2 + \cdots \tag{2.7.14}$$

An example of an integration algorithm is the Verlet algorithm which is one of the most widely used integrators in MD simulations.[245, 249] This method determines the positions at the next time step, *r(t+Δt)*, from the acceleration at the current time, *a(t)*, and positions at the current and previous step, *r(t)* and *r(t–Δt)* respectively. This is achieved by adding eq. 2.7.15 and 2.7.16 to give eq. 2.7.17. While calculation of positions is relatively simple, this method has a disadvantage that velocities at the current time step are not explicitly calculated and cannot be determined until the positions at the next have been determined (eq. 2.7.18). Furthermore, a loss of precision using this integrator can occur because the first two terms in eq. 2.7.17 are much larger than the last.[240, 245]

$$r(t+\Delta t) = r(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 + \cdots \quad (2.7.15)$$

$$r(t-\Delta t) = r(t) - v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 - \cdots \quad (2.7.16)$$

$$r(t+\Delta t) = 2r(t) - r(t-\Delta t) + a(t)\Delta t^2 \quad (2.7.17)$$

$$v(t) = \frac{r(t+\Delta t) - r(t-\Delta t)}{2\Delta t} \quad (2.7.18)$$

The Leap-frog algorithm is a modification of the Verlet algorithm which overcomes the disadvantages mentioned above.[250] It does this by explicitly calculating velocities and avoids addition of large and small terms.[245] The method first calculates *v(t + ½ Δt)* from *v(t − ½ Δt)* and *a(t)* using eq. 2.7.19, and subsequently determines *r(t+Δt)* from *v(t + ½ Δt)* and *r(t)* using eq. 2.7.20. Eq. 2.7.21 is then used to calculate *v(t)*. Velocities thus leap-frog over positions followed by positions leap-frogging over velocities. It is this leap-frogging effect, however, that is the major disadvantage of this method in that velocities and positions are not synchronised.[245]

$$r(t+\Delta t) = r(t) + v\left(t+\frac{1}{2}\Delta t\right)\Delta t \quad (2.7.19)$$

$$v\left(t+\frac{1}{2}\Delta t\right) = v\left(t-\frac{1}{2}\Delta t\right) + a(t)\Delta t \quad (2.7.20)$$

$$v(t) = \frac{1}{2}\left[v\left(t+\frac{1}{2}\Delta t\right) + v\left(t-\frac{1}{2}\Delta t\right)\right] \quad (2.7.21)$$

A final integrator of note is the velocity Verlet algorithm which calculates positions, velocities and accelerations at the same time but also has a high level of precision.[251] This algorithm calculates *r(t+Δt)* from *r(t)*, *v(t)* and *a(t)* using eq. 2.7.22, followed by calculation of *v(t+½Δt)* from *v(t)* and *a(t)* using eq. 2.7.23. New forces are calculated from *r(t+Δt)* from which *a(t+Δt)* is obtained and finally, *v(t+Δt)* is calculated from *v(t+½Δt)* and *a(t+Δt)* using eq. 2.7.24.[245] While this method overcomes the disadvantages of the above-mentioned algorithms, it requires more computing time.

$$r(t+\Delta t) = r(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 \tag{2.7.22}$$

$$v\left(t+\frac{1}{2}\Delta t\right) = v(t) + \frac{1}{2}a(t)\Delta t \tag{2.7.23}$$

$$v(t+\Delta t) = v\left(t+\frac{1}{2}\Delta t\right) + \frac{1}{2}a(t+\Delta t)\Delta t \tag{2.7.24}$$

Knowing how a system changes position and velocity over time allows for the calculation of time-dependent properties, *X(r,p)*. These properties are determined as ensemble averages, $\langle X \rangle_M$, where an ensemble is a collection of configurations (see eq. 2.7.25). According to the ergodic hypothesis, the ensemble average is approximated as being the same as the time average, with the statistical error associated with the latter, σX, being proportional to the square of the number of configurations (see eq. 2.7.26).[240, 252] Consequently, larger ensembles produce lower errors. However, the accuracy of the average is dependent on how well the ensemble represents the property being calculated.[245] These calculated properties can thus be compared to experimentally-measured values, although agreement is sometimes difficult to obtain since the latter are time-averages determined for a much larger number of conformations (approx. 10^{23}) and over longer periods of time than calculated in MD simulations.

$$\langle X \rangle_M = \frac{1}{M}\sum_i^M X(r_i, p_i) \tag{2.7.25}$$

$$\sigma X \propto \frac{1}{\sqrt{M}} \tag{2.7.26}$$

CHAPTER 3

EXPERIMENTAL INVESTIGATION OF THE CQ-FERRIHAEM INTERACTION

3.1 Introduction

The interaction of CQ and ferrihaem has been previously studied using a wide range of experimental techniques. These included UV-visible,[124, 163] IR,[181] Raman,[171, 253] EXAFS,[152, 218] NMR[168, 169, 172] and Mössbauer[254] spectroscopies as well as ITC[125, 160] and mass spectrometry.[170] Investigations have been conducted in the both the solid and solution states with the latter often spanning wide pH ranges and incorporating various organic solvents such as DMSO and methanol or detergents. Because of the variability in experimental conditions used as well as the frequent reliance on a single experimental technique, many conflicting results have been obtained from these investigations. An example of such a contradiction is observed for the reported binding stoichiometries of the CQ-ferrihaem complex which have ranged from 1:1 to 1:8.[124, 125, 161, 172] Additionally, there has been controversy over the form of ferrihaem present in the complex where some authors suggest ferrihaem is monomeric while others have proposed it is a μ-oxo dimeric species.[77] Investigations have been hampered by the strong dependence of ferrihaem speciation on solvent, pH and ionic strength which have very likely contributed to the conflicting results obtained.[75, 76] Furthermore, some authors have suggested that changes to ferrihaem speciation in the solid state occur during the drying process and that the species observed in aqueous solution is not the same as that obtained in the solid state at more biologically relevant pH (~5).[181] Consequently, there is much confusion surrounding the interaction of CQ and ferrihaem in both the solution and solid states.

This chapter details the experimental investigation of the CQ-ferrihaem complex in both aqueous solution and the solid precipitated at pH 5. The interaction of CQ and ferrihaem has been comprehensively studied using a wide variety of complementary experimental techniques, taking care to avoid solvent-induced ferrihaem speciation by performing experiments in aqueous solution, thus better reproducing biological conditions. The main aim of this work was to provide a clearer understanding of the interaction of CQ with ferrihaem which could then be used in subsequent modelling studies and hence reduce the confusion surrounding this topic.

3.2 Methods

3.2.1 General.

All commercially available materials were purchased from Sigma-Aldrich with the exception of haemin (Cl-ferrihaem, Fluka) and D_2O (Merck). All glassware was washed scrupulously as reported previously in order to prevent build-up of ferrihaem.[75] UV-visible spectra were recorded on a Varian Cary 100 or Shimadzu UV1800 UV-visible spectrophotometer. MCD spectra were recorded on a Chirascan-Plus CD spectrophotometer with MCD attachment (calibrated at 0.977 T) using step scans with step size of 1 nm, time per point of 0.5 s and bandwidth of 1 nm. IR spectra were recorded on a PerkinElmer Spectrum 100 spectrophotometer using an ATR attachment. NMR measurements were performed on a Bruker Ultrashield 400 Plus spectrometer at 303 K. Mass spectrometry was performed on an Agilent 6530 QTOF LC-MS system coupled with an Agilent 290 Infinity UHPLC using an Eclipse Plus C18 column (50 × 2.1 mm, 1.8 μm). The pH of solutions was determined using a Jenway 3510 or Crison MicropH 2000 pH meter.

3.2.2 Preparation of CQ (free base).

Unless otherwise stated, all experiments involving CQ were performed with the diphosphate salt. The free base form of CQ used in some cases was prepared from the diphosphate salt by adding 2M NaOH to an aqueous solution of CQ (0.2 M) which was extracted using dichloromethane. The organic layer was separated, dried using $MgSO_4$ and removed under vacuum. The resultant oil was washed with a small volume of diethyl ether which was removed under vacuum to afford the free base as a white solid.

3.2.3 Preparation of ferrihaem species and the CQ-ferrihaem complex.

The solid precipitate of ferrihaem was obtained by adding to a haematin solution (11.5 mM in 0.1 M NaOH), 1.75 M acetic acid in 50 μL aliquots until a measured pH of 5 was obtained. The precipitate

was centrifuged at 4000 rpm for 20 minutes and the supernatant discarded. The pellet was extensively washed by repeated resuspension in water and centrifugation. After the final washing, the precipitate was left to dry in a desiccator over phosphorous pentoxide. The CQ-ferrihaem complex was prepared in the same manner except CQ free base was added to give an initial concentration of 58.5 mM. The μ-oxo dimer of ferrihaem was prepared as a tetrasodium salt according to a method modified from Brown *et al*.[64] This procedure involved adding NaOH (5 M, 0.5 mL) to a haematin solution (79 mM in DMSO), followed by excess acetone. The precipitate obtained was collected by filtration, washed with acetone and air dried. To remove traces of water, the sample was further dried at 100 °C. To obtain solid HO-ferrihaem as a disodium salt, an aqueous solution of haematin (23 mM in 0.07 M NaOH) was lyophilised using a thermosavant YOD 230 lyophiliser and one equivalent of NaCl produced as a side-product.

3.2.4 Spectrophotometric titrations.

Solutions used in titrations were buffered to pH 7.4 using 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, 0.02 M) and additions to the working solution were made using a 50 μL Hamilton syringe. Titrations were performed by adding aliquots of an aqueous ferrihaem solution (1 mM in 0.02 M HEPES) to an aqueous CQ solution (0.02 mM in 0.02 M HEPES). The former was prepared by diluting a haemin stock solution (20 mM in 0.1 NaOH) to the required concentration with 0.02 M HEPES, while the latter was made by diluting a stock CQ solution (0.08 mM in 0.02 M HEPES) with the same buffer solution. Titration data were analysed using the HypSpec software package according to a 1:2 CQ:ferrihaem model as determined from Job plot experiments (see Section 3.3.1).[255]

3.2.5 Job plots.

To construct Job plots, the UV-visible spectra of solutions (prepared in the same solvent system as used for spectrophotometric titrations) of varying CQ and ferrihaem mole fraction were recorded. The total combined concentration of CQ and ferrihaem working solutions was constrained to 0.2 mM and made up using stock CQ (2 mM in 0.2 M HEPES) and ferrihaem solutions (2 mM in 0.015 mM NaOH). Fifteen working solutions were prepared with ferrihaem mole fractions of 0.00, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95 and 1.00. These solutions consisted of (i) a combined

volume of CQ and ferrihaem of 0.1 mL (based on the individual volumes needed for required mole fractions); (ii) an aliquot of HEPES buffer (0.2 M, pH 7.4) equal to the difference in volume between 0.1 mL and the volume of CQ added; (iii) an aliquot of NaOH (0.15 M) equal to the difference in volume between 0.1 mL and the volume of ferrihaem added; and (iv) 0.8 mL water. Solutions gave measured pH values in the range 7.5 – 7.6. UV-visible spectra were recorded in a 0.1 cm pathlength cuvette at 25.0 or 30.0 ± 0.2 °C and analysed at 370 nm.

3.2.6 MCD spectra.

All MCD spectra of ferrihaem species were recorded by diluting a stock haemin solution (1 mM in 0.1 M NaOH) into various solvent systems. To obtain the spectrum of H_2O-ferrihaem, 15 μL of the haemin stock solution was added to 1.885 mL aqueous acetone (5.64 M), followed by the addition of 0.1 mL nitric acid (0.2 M). This produced a final ferrihaem concentration in solution of 7.5 μM with a measured pH of 2. The spectrum of $(imidazole)_2$-ferrihaem was recorded on a solution prepared by diluting an imidazole stock (0.2 M, pH 7.4) and haemin stock solution to 0.02 M and 10 μM, respectively, with HEPES buffer (0.02 M, pH 7.4). The μ-oxo ferrihaem dimer spectrum was recorded on a solution consisting of 15 μL haemin stock in an aqueous DMSO solution (5.64 M, 2 mL). A solution of ferrihaem π-π dimer was prepared by diluting the haemin stock solution to 15 μM with HEPES buffer (0.02 M, pH 7.4). The CQ-ferrihaem complex was made as described for the π-π dimer with the addition of CQ diphosphate (1 mM in water) to the solution to give a final CQ concentration of 15 μM.

3.2.7 Magnetic susceptibility.

Magnetic susceptibility measurements of ferrihaem species in solution were determined using the Evans NMR method (see Section 2.5.1). The solution used to record the magnetic moment of μ-oxo ferrihaem consisted of 3 mg haematin dissolved in NaOD (0.3 M in D_2O, 1.2 mL), to which 0.8 mL d_6-DMSO was added. For H_2O-ferrihaem, 3 mg haematin was dissolved in 1.7 mL acidified d_6-acetone (0.2 M nitric acid) and diluted to 2 mL with D_2O. Reference solutions for both μ-oxo ferrihaem and H_2O-ferrihaem were made in the same manner as the sample solution with the exclusion of ferrihaem. The magnetic susceptibility of π-π dimeric ferrihaem was determined using a solution comprising haematin

(3 mg in 0.01 M NaOD, 1.89 mL) and phosphoric acid (0.1 M in D_2O, 0.11 mL), the measured pD of which was 7.2. D_2O was used as the reference solvent. To determine the magnetic moment of the CQ-ferrihaem complex, a haematin solution (3 mg in 0.01 M NaOD, 0.525 mL) was added to a CQ solution (12.2 mg in D_2O, 1.575 mL) giving a measured pD of 7.5. The reference solution (with a measured pD of 7.5) was prepared by adding NaOD (0.1 M, 0.49 mL) to the same CQ solution (1.51 mL) used for the sample. Solid-state magnetic susceptibility of ferrihaem samples was measured using an Evans balance (see Section 2.5.2). The magnetic susceptibility of the standard, mercury(II)tetrathiocyanatocobaltate(II), reported by Figgis and Nyholm (16.44×10^{-6} cm^3 g^{-1}) was used as a calibrant.[226] Samples were measured five times and the average reported. For both solid- and solution-state magnetic moment determination, diamagnetic corrections of -5.09×10^{-7} and -4.64×10^{-7} cm^3 g^{-1} for ferrihaem and CQ, respectively, were used. The latter was determined from Pascal's constants, while the former was calculated from the experimental value determined for PPIX of 2.95×10^{-4} cm^3 mol^{-1} and Pascal's constants for Fe(III) and H_2O/O^{2-}.[224]

3.2.8 Diffusion.

Diffusion coefficients of CQ and the CQ-ferrihaem complex in aqueous solution (pH 7.4) were measured using the diffusion cell method reported by Linder *et al*.[256] according to the protocol described by de Villiers *et al*.[75] In the case of CQ, a working solution (0.6 mM in 0.02 M HEPES, pH 7.4) was prepared by diluting a stock CQ solution (6 mM in 0.2 M HEPES, pH 7.4) with water. To determine the diffusion coefficient of CQ-ferrihaem a working solution consisting of haematin stock solution (3 mM in 0.1 M NaOH, 2.5 mL) and the same CQ stock (2.5 mL) solution in 20 mL water was made. The final concentration of ferrihaem and CQ in this solution was 0.3 mM and 0.6 mM, respectively. Working solutions were placed in the lower chamber and buffer solution (0.02 M HEPES, pH 7.4) in the upper chamber. Solutions were allowed to diffuse from the lower to upper chamber over 1 or 2 hours for CQ and CQ-ferrihaem respectively. Concentrations of CQ and CQ-ferrihaem in the upper chamber were determined using absorbance measurements at 342 (CQ) and 385 nm (CQ-ferrihaem) and extinction coefficients obtained from Beers Law plots of the working solutions (18 560 for CQ and 42 760 M^{-1} cm^{-1} for the complex). Four replicates were determined for CQ at 297 K and eight replicates for CQ-ferrihaem, four at 297 K and the remaining four at 295 K. Diffusion coefficients were normalised to 298 K using temperature and viscosity corrections according to the Stokes-Einstein relationship.[257]

3.3 CQ-Ferrihaem Interaction in Solution

3.3.1 Stoichiometry and binding of CQ to ferrihaem.

The binding stoichiometry of the CQ-ferrihaem complex in aqueous solution (pH 7.4) was probed using a Job plot. Under aqueous conditions, ferrihaem has been shown to exist in equilibrium between the monomeric and π-π dimeric form and is described by the dimerisation constant logK_D = 6.82 ± 0.02 at pH 7.4.[75] Owing to the concentration dependence of ferrihaem species, complications arising from changes in speciation as the mole fraction was varied were minimised by using high ferrihaem concentrations. At the lowest concentration used (20 μM), 94 % of ferrihaem existed as the dimeric species[75] and from the third lowest concentration onwards, it was more than 97 % dimerised. The Job plot determined for the CQ-ferrihaem complex at 25 °C is displayed in Fig. 3.1 and indicates that binding occurred in a 1:2 CQ:ferrihaem stoichiometry. The same ratio was obtained at 30 °C.

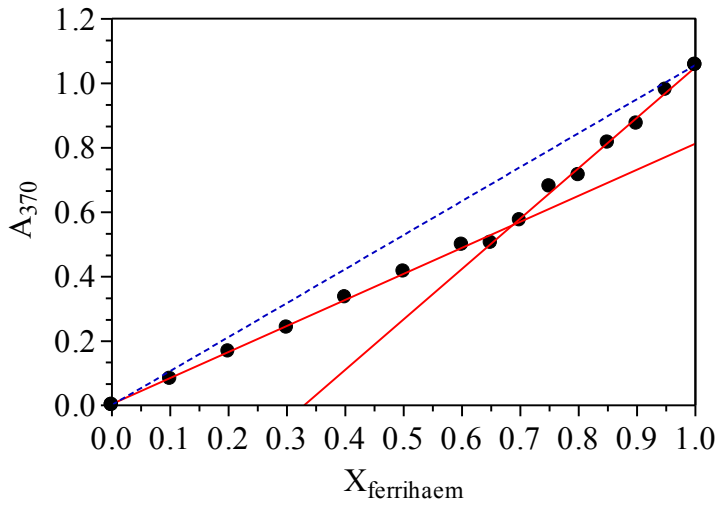


Figure 3.1. Typical Job plot obtained for the CQ-ferrihaem complex in aqueous solution pH 7.4 (circles). Dashed blue line represents the theoretical line of dilution of ferrihaem (Beers law). Average intersection of the two solid red lines was 0.68 ± 0.02 (n = 5).

The observed stoichiometry of the CQ-ferrihaem complex was in agreement with ratios reported by Chou *et al*.[99] probed using equilibrium dialysis in aqueous solution (pH 7.4); Constantinidis and Saterlee[164] for urohaemin determined using a UV-visible Job plot (pH 6); Leed *et al*.[172] and Schwedhelm *et al*.[170] using NMR relaxation investigations (pH 6.5); and Crespo *et al*.[163] using a UV-visible Job plot (pH 7.4). A number of studies have reported higher ratios (1:4 to 1:8), however, all these investigations were performed using ITC with high phosphate buffer concentrations (e.g. 0.25 M).[125, 159, 160] It is possible that the high ionic strength of these experimental conditions may have caused ferrihaem aggregation. Indeed, elevated ionic strengths have previously been shown by Asher *et al*. to induce higher aggregation of μ-oxo ferrihaem dimer.[76]

The stoichiometry of interaction was used in conjunction with spectrophotometric titration to determine the association constant, K_{obs}, of CQ with ferrihaem. Performing the titration by adding CQ to ferrihaem produced data that were not in agreement with the stoichiometry identified by Job plots, where data points fell below the strong binding limit at low CQ concentration (see Fig.3.2).

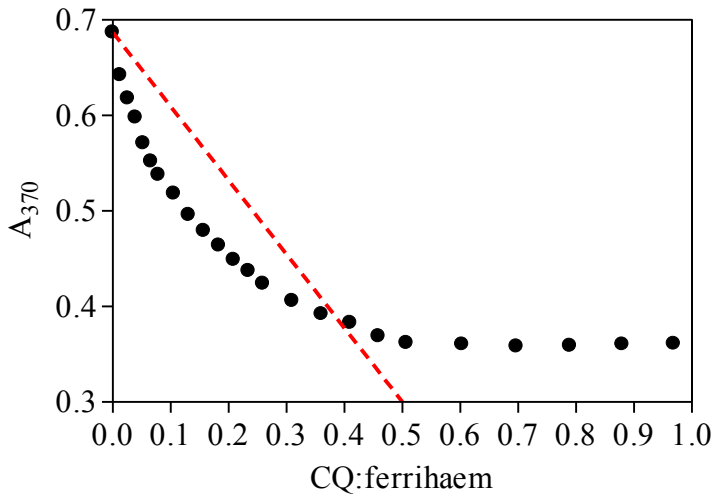


Figure 3.2. Typical spectrophotometric titration curve obtained when CQ was added to ferrihaem in aqueous solution (pH 7.4). Dashed red line indicates the strong binding limit for a 1:2 (0.5:1) CQ:ferrihaem molar ratio.

For this apparent increase in stoichiometry to have occurred, a lower concentration of ferrihaem must have existed in solution and could possibly have resulted from precipitation or adsorption onto the walls of the cuvette. Given no precipitation was observed and the fact that such adsorption has previously been reported for ferrihaem,[75] the latter of the two was more likely. A further complication of performing the titration in this manner arose from the fact that analysis of the data was mathematically intractable because the concentration of the free CQ appears in the denominator as well as the numerator in the expression describing free ferrihaem. Thus for strong binding and at low CQ concentrations, the denominator tended to zero. Interestingly, similar experimental difficulties were experienced by Crespo *et al*. when the titration was performed in the same way.[163]

The data could only be fitted when performing the titration in reverse, i.e. titrating a ferrihaem solution into CQ. Data were analysed by fitting the entire spectrum (300 – 800 nm) using the HypSpec software package. The binding model employed accounted for both complex formation ($CQ{\cdot}M_2$) and dimerization (M_2) in equations 3.1 to 3.4:

$$2M \overset{K_D}{\rightleftharpoons} M_2 \tag{3.1}$$

$$K_D = \frac{[M_2]}{[M]^2} \tag{3.2}$$

$$2M + CQ \overset{K_{obs}}{\rightleftharpoons} CQ \cdot M_2 \tag{3.3}$$

$$K_{obs} = \frac{[CQ{\cdot}M_2]}{[M]^2[CQ]} \tag{3.4}$$

Unconstrained attempts to fit the data were not successful owing to the large number of fitted parameters. However, given that the value of K_D as well as spectra of CQ and the ferrihaem π-π dimer (M_2) were independently known, they could be employed as constants in the fitting procedure. Using this approach produced a fitted log K_{obs} value of 14.1 ± 0.4 (n = 12). The percentage of monomeric ferrihaem (*M*) obtained from the fit was on average only 1.4 % and never exceeded 2.5 % at any point throughout the titration. This suggested that a simplified model described by equation 3.5 which ignores

the very small amount of monomeric species present could instead be used to fit the data. Employing the simpler model permitted free-fitting of all parameters from which the equilibrium constant *K′* could be determined. *K′* can be related to the K_{obs} parameter of the more complex model via equations 3.6 and 3.7:

$$M_2 + CQ \overset{K'}{\rightleftharpoons} CQ \cdot M_2 \tag{3.5}$$

$$K' = \frac{[CQ \cdot M_2]}{[M_2][CQ]} \tag{3.6}$$

$$K' \times K_D = \frac{[CQ \cdot M_2]}{[M_2][CQ]} \times \frac{[M_2]}{[M]^2} = \frac{[CQ \cdot M_2]}{[CQ][M]^2} = K_{obs} \tag{3.7}$$

The value for log K_{obs} obtained using the simplified model was 13.3 ± 0.2 (log *K′* + log K_D = [6.5 ± 0.2] + [6.82 ± 0.02]) which compared well with that calculated using the more complex model but displayed a smaller error.

Fitting of spectrophotometric titration data over the entire spectrum was a more stringent method of determining parameters than single wavelength fitting. This method required all variables to reproduce data points at each step throughout the titration at every wavelength. An added advantage of the fitting procedure was that the spectrum of each species involved in the binding model was computed. In the case of the simplified model described above, the fitted spectra of CQ, ferrihaem π-π dimer and the CQ-ferrihaem complex were determined and are shown in Fig. 3.3a – c. It should be noted that the spectra of CQ and CQ-ferrihaem complex were directly observable at the beginning and end of the titration respectively (see Fig 3.3b and c), while at no point was the spectrum of the π-π dimer dominant (see Fig. 3.3a). Despite this fact, the fitted spectrum of the latter was in remarkably good agreement with the experimental UV-visible spectrum independently recorded for the π-π dimer. The spectrum of CQ was similarly well reproduced. Minor deviations observed in the fitted spectra around 400 nm are to be expected given the exclusion of the monomeric ferrihaem species from the simplified model. The fitted spectrum of the CQ-ferrihaem complex (Fig. 3.2c) was in good agreement with the experimental UV-visible spectrum of ferrihaem in the presence of excess CQ. No difference to the predicted spectrum of this species was observed when fitting the data to the more complex model described by equations 3.1 to 3.4.

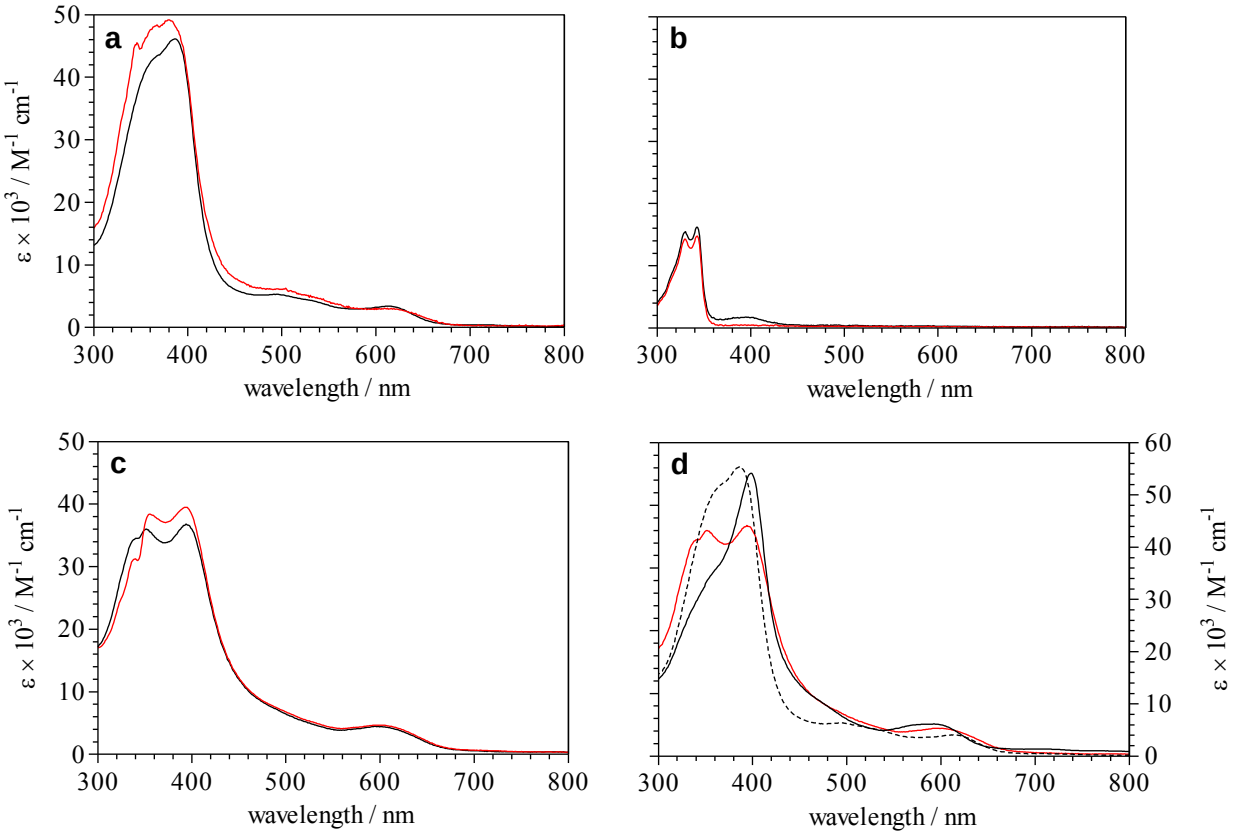


Figure 3.3. Predicted (red) and experimental (black) spectra of (**a**) ferrihaem π-π dimer; (**b**) CQ and (**c**) CQ-ferrihaem complex. (**d**) Comparison of the predicted CQ-ferrihaem complex spectrum with experimental spectra of the ferrihaem π-π dimer (black dashed line) and μ-oxo ferrihaem dimer (black solid line). The complex and π-π dimer are plotted on the left axis, while the μ-oxo dimer is plotted on the right. All predicted spectra were calculated from spectrophotometric titration data obtained by addition of ferrihaem to CQ in aqueous solution at pH 7.4.

3.3.2 Identifying the species of ferrihaem in the CQ-ferrihaem complex.

While the stoichiometry of interaction and strength of binding for CQ-ferrihaem was successfully determined, an important consideration still remained, namely, the form of ferrihaem in the CQ-ferrihaem complex. Previous reports by de Villiers *et al*. showed the utility of UV-visible spectra in

identification of ferrihaem species.[75] Comparison of the predicted spectra of the complex obtained from fitting of spectrophotometric data to independently measured UV-visible spectra of the π-π and μ-oxo ferrihaem dimer was thus made (see Fig. 3.3d). The fitted spectrum of the CQ-ferrihaem complex shared similarity with that of the μ-oxo ferrihaem dimer, albeit with approx. 20% hypochromism, while it was markedly different from that of the π-π dimer (see Fig. 3.3d). Particularly good agreement was observed in the long wavelength region with both the complex and μ-oxo dimeric ferrihaem species displaying a similar single broad spectroscopic envelope at 600 and 590 nm, respectively. The broad low-energy shoulder of the Soret band was also in good agreement. The Soret band and high-energy shoulder were somewhat less similar. As mentioned above, the Soret peak had a lower intensity than the corresponding band maxima of μ-oxo ferrihaem while the high-energy shoulder was obscured by overlapping CQ peaks.

While the spectrum of the CQ-ferrihaem complex predicted from spectrophotometric titration appeared most closely related to that of the μ-oxo dimeric species of ferrihaem, the broad features of the UV-visible spectrum prevented definitive identification of the form of ferrihaem in the complex. On the other hand, spectra obtained using MCD spectroscopy are often better resolved as the presence of positive and negative Gaussian-type bands (B- and C-terms) and Gaussian derivative bands (A-terms) are possible (see Section 2.3.3 for description of these bands). By contrast, only positive Gaussian-type bands are observed in UV-visible spectra. The MCD spectrum of the complex was determined and compared to a variety of ferrihaem species. These included high-spin monomeric H_2O-ferrihaem and the low-spin bis-imidazole complex,[183] $(imidazole)_2$-ferrihaem, as well as π-π dimeric and μ-oxo dimeric ferrihaem species. MCD spectra of the two monomeric and three dimeric ferrihaem species are displayed in Fig. 3.4 and 3.5 respectively, alongside the corresponding UV-visible spectra for comparison. Owing to the marked differences in MCD spectra across all ferrihaem species, it was clear that this could be used in combination with UV-visible spectroscopy as a fingerprint technique to more easily identify ferrihaem speciation.

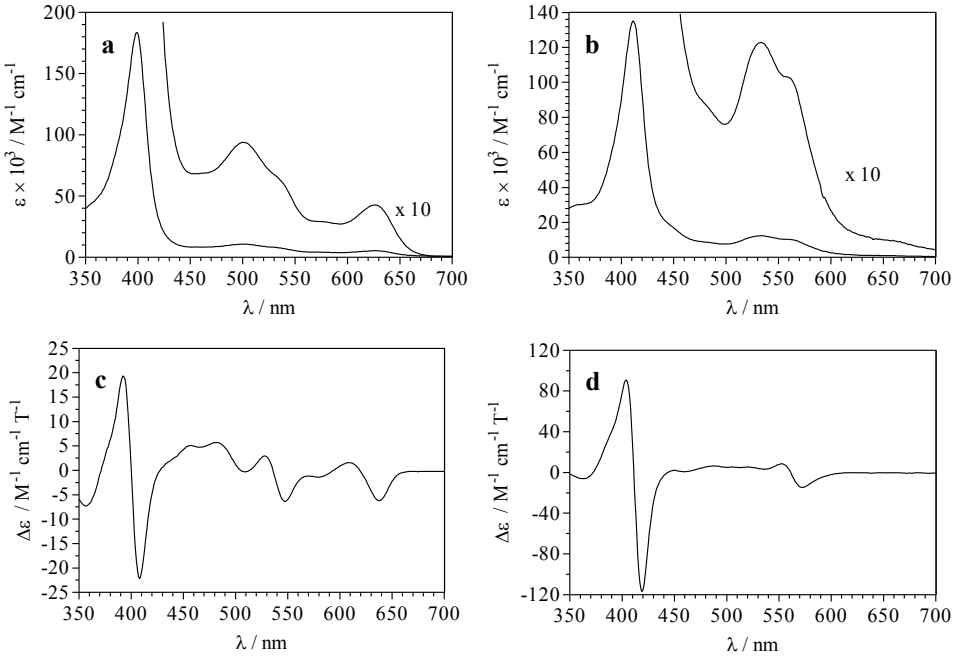


Figure 3.4. UV-visible (top) and MCD (bottom) spectra of monomeric ferrihaem species. (**a** and **c**) H_2O-ferrihaem recorded in aq. acetone solution, pH 2. (**b** and **d**) Low-spin $(imidazole)_2$-ferrihaem recorded in 40% aq. DMSO, pH 7.4. UV-visible and MCD spectra were recorded at 25 °C and room temperature, respectively.

Monomeric H_2O- and $(imidazole)_2$-ferrihaem species were both characterised by an intense derivative-shaped feature in their MCD spectra. In the case of H_2O-ferrihaem, the band was less intense than observed for $(imidazole)_2$-ferrihaem and its centre somewhat blue-shifted (400 vs. 410 nm). The lower energy (longer wavelength) region was markedly different between the two species. The $(imidazole)_2$-ferrihaem complex displayed one weak derivative-shaped feature centred around 562 nm which corresponded to the low energy envelope characteristic of low-spin species in the UV-visible spectrum.[183, 258] By contrast, the corresponding region of H_2O-ferrihaem displayed a number of notable features. These included the fairly broad positive peak near 480 nm with prominent shoulder around 450 nm as well as two derivative-shaped features centred at 535 and 634 nm which appeared to correspond to the Q- and charge transfer bands in the UV-visible spectrum as assigned by Makinen and Churg for H_2O-metMb (see Section 2.2.1).[196]

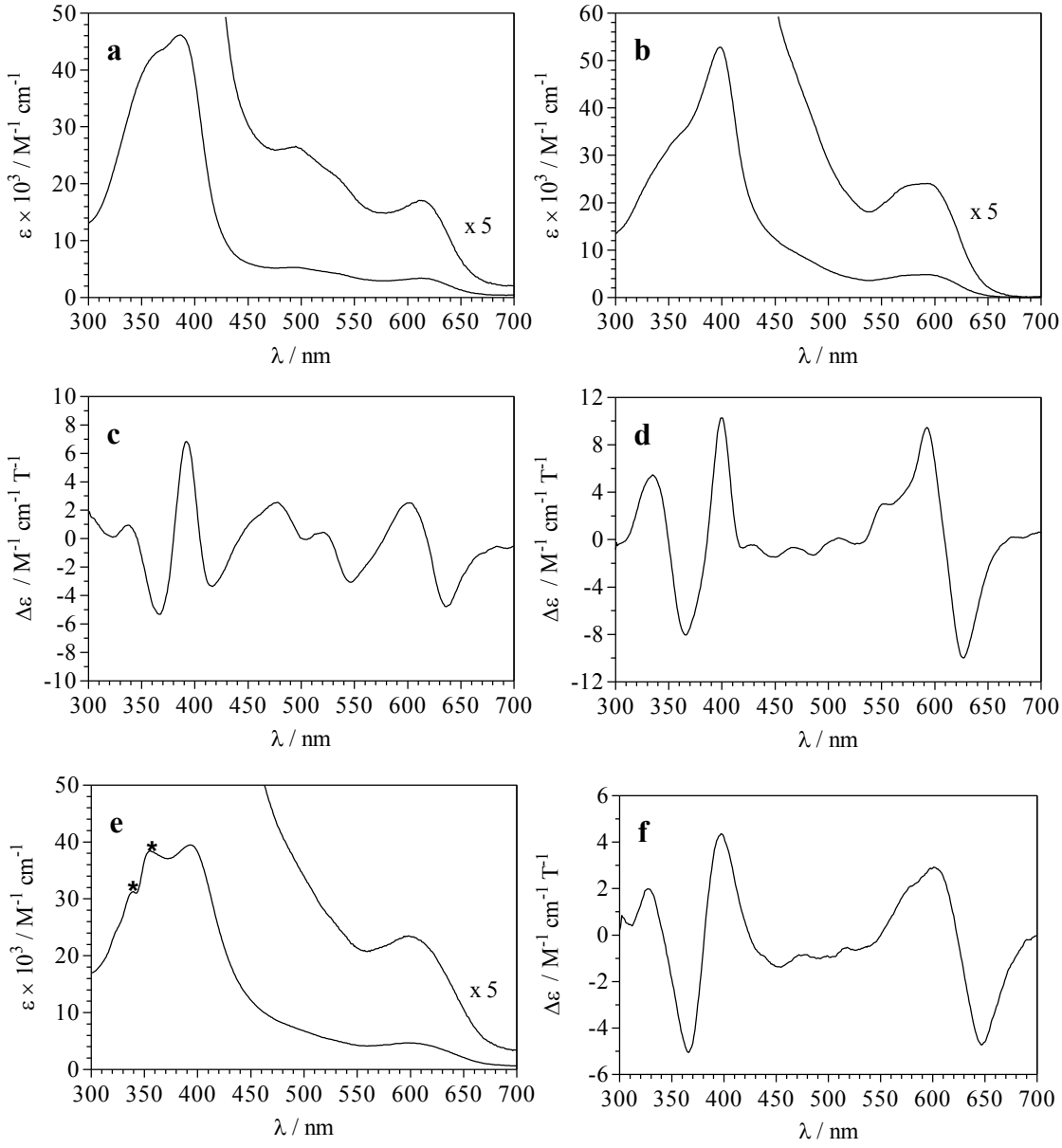


Figure 3.5. UV-visible and MCD spectra of dimeric ferrihaem species and the CQ-ferrihaem complex. (**a** and **c**) Ferrihaem π-π dimer recorded in aqueous solution, pH 7.4. (**b** and **d**) The μ-oxo ferrihaem dimer recorded in an alkaline 40% aq. DMSO solution. (**e** and **f**) CQ-ferrihaem complex recorded in aqueous solution, pH 7.4. Asterisks denoted peaks due to CQ. UV-visible and MCD spectra were recorded at 25 °C and room temperature, respectively.

The MCD spectrum of the π-π dimeric species displayed some similarity to that of H_2O-ferrihaem in the long wavelength region where the same broad positive Gaussian band and two derivative-shaped peaks were observed. In the case of the latter features, these were shifted to shorter wavelength relative to their counterparts in H_2O-ferrihaem (535 vs. 530 nm and 634 vs. 620 nm). A marked difference was observed, however, in the Soret region between the π-π dimeric and monomeric H_2O-ferrihaem species where the former had a significantly less intense and seemingly reversed Gaussian derivative-shaped peak. A similar feature was observed in the same region for μ-oxo ferrihaem. At higher wavelengths, however, the spectrum of μ-oxo ferrihaem was markedly different from the monomeric and dimeric species already described. An intense positive Gaussian derivative-shaped feature centred on 609 nm dominated at lower energy, while a series of weak Gaussian-shaped bands existed between this feature and the Soret peak. Of the spectra reported in Fig. 3.4 and 3.5, the MCD spectrum of the CQ-ferrihaem complex mostly closely resembled that of μ-oxo ferrihaem. More specifically, the complex displayed a similar apparently reversed derivative-shaped peak in the Soret region and intense positive derivative-shaped feature at longer wavelength with a fairly featureless region consisting of weak Gaussian-shaped bands between the two. Thus both UV-visible and MCD spectroscopy suggested that the form of ferrihaem in the complex was μ-oxo dimeric.

Further evidence in support of CQ inducing the μ-oxo dimeric species was obtained by measuring the magnetic susceptibility of the complex using the Evans NMR method (see Section 2.5.1). In aqueous solution (pH 7.5), the magnetic moment was determined as 2.25 ± 0.02 μ_B. This value was significantly different to that obtained for ferrihaem in the absence of CQ (4.8 ± 0.1 μ_B) which agreed well with the magnetic moment previously reported for the π-π dimeric species (4.6 μ_B). The low magnetic moment can be accounted for by one of two possible explanations. The first is that the complex was low-spin (S = ½). Low-spin complexes arise when two strong-field ligands coordinate to the iron centre of ferrihaem, an example of which is $(imidazole)_2$-ferrihaem.[183, 258] Such a situation could not occur in the CQ-ferrihaem complex since the mole fraction of ferrihaem in the case of a low-spin complex would be 0.33 which was inconsistent with that observed experimentally (0.68, see Section 3.3.1). Furthermore, the characteristic low-spin features observed in the UV-visible and MCD spectra of $(imidazole)_2$-ferrihaem were not reproduced in the corresponding spectra of the complex. On the basis of these two pieces of evidence, it was concluded that the complex was not low-spin.

The second more likely explanation of the low magnetic moment was that antiferromagnetic coupling between two iron centres facilitated by coordination of a bridging ligand was present. Given the magnitude of the magnetic moment, a single bridging atom would be required to provide the necessary orbital overlap for an efficient superexchange mechanism.[259] In theory, chloride met the requirements for the observed antiferromagnetic coupling, however, this ligand could be discounted in this case based on its deliberate exclusion from experimental conditions by using haematin rather than haemin. This left only two possible ligands remaining, namely oxide and hydroxide. Complexes involving the latter ligand are less common than those involving the former and reported *J*-coupling constants of μ-hydroxo-bridged iron porphyrins tend to be weaker than those of μ-oxo dimers.[260-262] Indeed, theoretical investigations by Chen *et al*. supported this observation where protonation of oxide-bridged μ-oxo dimers was predicted to reduce *J*-coupling constants.[263] Consequently the antiferromagnetic coupling observed for the complex was unlikely to arise from ligation by hydroxide and thus left the μ-oxo species as the only feasible candidate. This conclusion was supported by the close agreement of magnetic moments measured for μ-oxo ferrihaem (2.0 ± 0.1 μ_B) and the CQ-ferrihaem complex (2.25 ± 0.02 μ_B). Furthermore, both values corresponded well with predicted values for μ-oxo ferrihaem based on the *J*-coupling constant reported by Stanek and Dziedzic-Kocurek using SQUID measurements (see Fig. 3.6).[220]

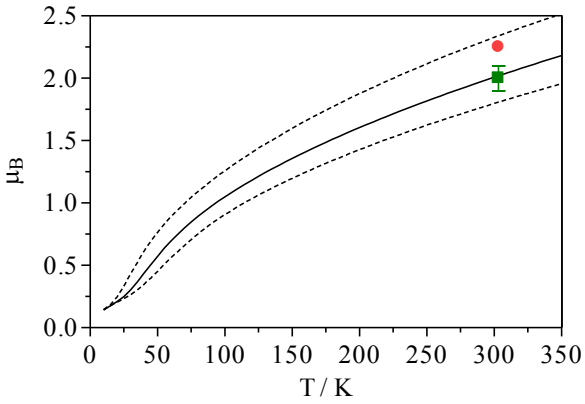


Figure 3.6. Predicted temperature dependence of the magnetic moment, μ_B, of μ-oxo ferrihaem determined from eq. 14 and the *J*-coupling constant of -110 ± 15 cm^{-1} reported by Stanek and Dziedzic-Kocurek.[220] Two standard deviations of the predicted dependence are depicted by dashed lines. Value determined for μ-oxo ferrihaem and the CQ-ferrihaem complex in solution are represented by the green square and red circle, respectively.

While several studies have previously suggested that CQ interacts with μ-oxo ferrihaem,[125, 161, 167, 168, 170, 172] many have based this fact on the presumption that the μ-oxo dimer was the dominant form of ferrihaem in aqueous solution. However, three relatively recent studies have convincingly shown that π-π dimeric ferrihaem is instead the major species.[75, 76, 163] Despite this evidence, there is still a widespread misconception that μ-oxo ferrihaem dominates in aqueous solution, with even a very recent review stating this.[77] It was thus important to make the explicit distinction that the spectroscopic and magnetic susceptibility results of the current study showed that CQ induces μ-oxo dimer formation in aqueous solution from the π-π dimeric species and that CQ did not simply bind to pre-formed μ-oxo ferrihaem.

3.3.3 Determining the aggregation state of the CQ-ferrihaem complex.

The findings reported above strongly suggested that CQ induced the μ-oxo dimeric species from the π-π ferrihaem dimer in aqueous solution. The aggregation state of the CQ-(μ-oxo dimer) complex, however, could not be inferred from these experiments. There has been some disagreement in the literature regarding this. Initially, Moreau *et al*. suggested that the complex consisted of large aggregates of alternating CQ and μ-oxo dimer molecules.[167] Later work by Schwedhelm *et al*., however, proposed that it existed in a 2:4 CQ:ferrihaem complex.[170] In an attempt to probe whether the complex existed as monomeric, dimeric or high-order aggregates, the diffusion coefficient, *D*, was measured using a diffusion cell method. The value determined for the CQ-(μ-oxo dimer) complex is listed in Table 3.1. Furthermore, the diffusion coefficient measured for CQ as well as those reported for related ferrihaem species are also included for comparison.[75] All values were normalised to 25 °C using temperature and viscosity corrections according to the Stokes-Einstein relationship. This relationship has also sometimes been used to determine molecular size. However, since it was developed for macroscopic colloid particles it is not generally applicable at the molecular level, particularly in strongly solvating solvents such as water. Instead, several empirical relationships have been developed to relate diffusion coefficients to molecular size. These include the Othmer-Thakar, Hayduk-Laudie, Wilke-Chang, Scheibel and more recently Gustafson-Dickhut equations.[264-268]

Table 3.1. Diffusion coefficients (*D*) normalized to 298 K and calculated molar volumes, *V*.

	D ($\times 10^{-10}$ m^2 s^{-1})	V (cm^3 mol^{-1})
$(CN)_2$-ferrihaem [a]	2.2 ± 0.2	485
HO-ferrihaem [a, b]	1.4 ± 0.1	920
μ-oxo ferrihaem [a]	1.6 ± 0.1	907
CQ-ferrihaem [c]	0.6 ± 0.2	1217 (1:2)
		2434 (2:4)
		3652 (3:6)
		4869 (4:8)
CQ [c]	3.3 ± 0.3	288 [d]

[a] Values reported by de Villiers *et al*.[75]; [b] π-π dimer, pH 10; [c] Determined in this work; [d] Predicted using ACD ChemSketch.

The empirical linear relationship previously reported by Gustafson and Dickhut related the logarithm of the diffusion coefficient, log*D*, and logarithm of molar volume, log*V*, for polycyclic aromatic organic molecules.[268] This relationship was later further extended by de Villiers *et al*. to describe the monomeric bis-cyano ferrihaem complex, $(CN)_2$-ferrihaem, as well as the π-π ferrihaem dimer in aqueous solution.[75] From this relationship, the molar volume could be calculated for a known diffusion coefficient, from which the degree of aggregation could be estimated. To further improve the statistical reliability of the empirical correlation, additional diffusion coefficients reported by Neisner and Heintz,[270] as well as the value determined for CQ in this work were incorporated into the plot. The linear regression fitted to the enlarged data set was statistically significant ($r^2 = 0.95$, $P < 0.0001$) and is shown in Fig. 3.7 as a solid red line. Values of π-π and μ-oxo dimeric ferrihaem were purposely excluded from the fit in order to instead serve as test molecules. It was clear form Fig. 3.7 that these values fell on the linear regression and indicated that the correlation could indeed be used to estimate the degree of aggregation of the CQ-(μ-oxo dimer) complex.

Four degrees of aggregation of the CQ-(μ-oxo dimer) complex were considered, namely the monomeric (1:2), dimeric (2:4), trimeric (3:6) and tetrameric (4:8) states. The molar volume of each was estimated by dividing the molecular weight (1567, 3134, 4701 and 6268 g mol^{-1}) by the predicted density of the complex. The density was estimated from the weighted average of ferrihaem (1.3753 g cm^{-3}, as reported by de Villiers *et al.*[75]) and CQ (predicted using the ACD ChemSketch software package). The logarithm of each predicted molar volume was incorporated onto the plot in Fig. 3.7 (filled triangles). The value corresponding to the tetrameric (2:4) aggregation state was found to lie closest to the linear regression. This suggested that the CQ-(μ-oxo dimer) complex in aqueous solution exists in a form where two CQ and two μ-oxo dimer molecules are closely associated, i.e. in the form [CQ-(μ-oxo dimer)]$_2$. On the other hand, given the experimental error and proximity of the dimeric (1:2) and hexameric (3:6) data points to the linear regression, these aggregation states could not be entirely excluded. However, larger aggregates, such as those suggested by Moreau *et al.*, could be discounted.[167]

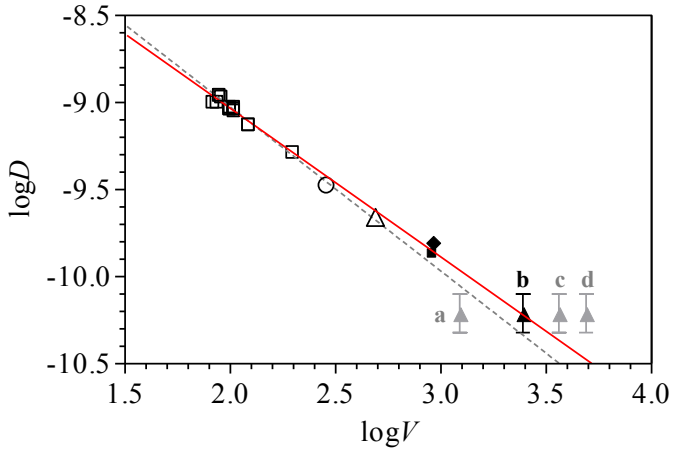


Figure 3.7. Dependence of log*D* (at 298 K) on log*V*. Values for aromatic molecules were obtained from Gustafson and Dickhut, and Neisner and Heintz (unfilled squares).[268, 270] Diffusion coefficients of $(CN)_2$-ferrihaem, π-π and μ-oxo ferrihaem dimer (unfilled triangle, filled square and filled diamond respectively) were reported by de Villiers *et al.* and have been normalised to 298 K.[75] Values for CQ (unfilled circle) and the CQ-ferrihaem complex were determined in this study. Aggregation state models for the CQ-ferrihaem complex are represented by filled triangles and include (**a**) 1:2, (**b**) 2:4, (**c**) 3:6 and (**d**) 4:8. The solid red line represents the linear regression fitted to the entire aromatic molecule data set, CQ and $(CN)_2$-ferrihaem. The dashed line represents the linear regression previously reported which only incorporates data reported by Gustafson and Dickhut and $(CN)_2$-ferrihaem. Filled data points were excluded from regression analysis. The models for the CQ-ferrihaem complex that did not fall on the solid regression line are shown in grey.

While the 2:4 aggregation state identified in this study was in agreement with that proposed by Schwedhelm *et al*., the structural model suggested by these authors to account for this aggregation state is incorrect.[170] As described in Section 1.5, these authors proposed two CQ capped μ-oxo dimers reciprocally coordinated to each other through a propionate side chain to the iron centre of the other. This structure was proposed to account for the low magnetic moment they erroneously attributed to the complex being low-spin. As noted in Section 3.3.2 above, the low magnetic moment instead arises from antiferromagnetic coupling of iron centres. In any case, the model proposed would likely not form a low-spin complex since the oxide and propionate groups are weak-field ligands and low-spin complexes of ferrihaem require strong-field ligands.

3.4 CQ-Ferrihaem Interaction in the Solid State

3.4.1 Stoichiometry of the CQ-ferrihaem complex.

Owing to the insolubility of ferrihaem in aqueous solution at pH 5, the investigation of ferrihaem and its interaction with CQ under these conditions was restricted to the solid state. The rapid precipitation that occurred at this pH thus prevented spectrophotometric titration and Job plot experiments from being conducted. Interestingly, while both ferrihaem and the CQ-ferrihaem complex readily precipitated, there was still a detectable concentration of the latter in solution. This observation supported the findings of Ursos *et al*. who reported that antimalarial drugs affect the pH dependent solubility of ferrihaem.[155] To determine whether the ferrihaem sample precipitated in the presence of CQ existed as a CQ-ferrihaem complex or of ferrihaem alone, mass spectrometry was employed. A molecular ion was detected at 320.19 m/z in the sample precipitated in the presence of CQ but was not observed in the sample obtained in the absence of CQ. This molecular ion corresponds to CQH^+ (320.88 g mol^{-1}) and hence indicated that the former exists as a CQ-ferrihaem complex.

To probe the stoichiometry of CQ:ferrihaem of the complex in the solid state, a portion of the precipitate was dissolved under the same conditions used to determine the Job plot in Section 3.3.1 (aqueous pH 7.4). The ratio of UV-visible absorbance values at 340 and 400 nm (A_{340}/A_{400}) as well as 370 and 400 nm (A_{370}/A_{400}) were measured in order to circumvent unknown concentration values and compared to the corresponding ratios determined from Job plot data (see Fig. 3.7). These particular ratios were chosen as they showed a large response at high and low ferrihaem mole fractions (A_{370}/A_{400} and A_{340}/A_{400} respectively). From Fig. 3.8, it was clear that the absorbance ratios determined from the solid CQ-ferrihaem complex most closely agreed with the Job plot data obtained when there were two ferrihaem molecules per CQ. This was the same stoichiometry determined for the complex in aqueous solution using the Job plot experiment and thus indicated that the complex precipitated in the same binding ratio as that in solution.

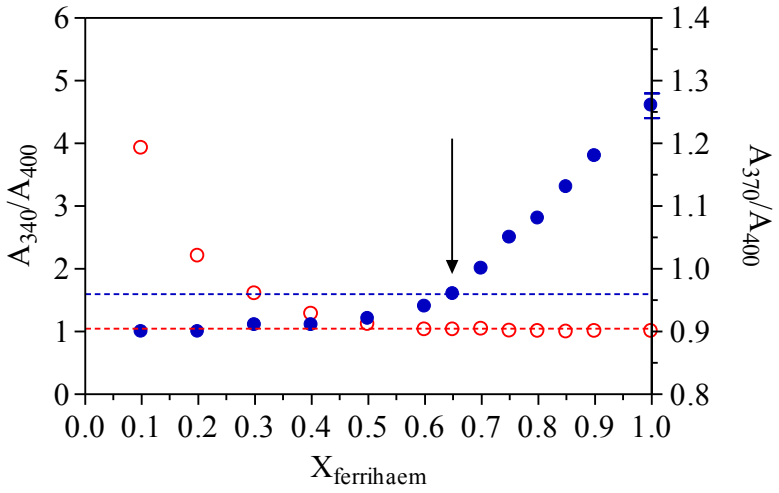


Figure 3.8. UV-visible absorbance ratios at 340 and 400 nm (A_{340}/A_{400}) and 370 and 400 nm (A_{370}/A_{400}) determined from Job plot experiments (see Section 3.3.1) as a function of mole fraction of ferrihaem. Data for A_{340}/A_{400} (blue solid circles) and A_{370}/A_{400} (red unfilled circles) are plotted on the left and right axes, respectively. Absorbance ratios calculated from the UV-visible spectrum obtained from dissolving the solid CQ-ferrihaem complex in aqueous solution (pH 7.4) are represented by dashed lines. The arrow identifies the mole fraction of ferrihaem to which both absorbance ratios most closely correspond and indicates the complex exists in a 1:2 CQ:ferrihaem stoichiometry.

3.4.2 Identifying the species of ferrihaem in the CQ-ferrihaem complex.

As detailed in Section 3.3.2, multiple techniques were used to identify the ferrihaem species that existed in the CQ-ferrihaem complex in aqueous solution. Specifically, UV-visible and MCD spectroscopy as well as Evans NMR measurements were employed. Once again, owing to the insolubility of ferrihaem, these techniques could not be used and instead the speciation of ferrihaem in the solid state was probed using IR spectroscopy and an Evans balance (see Section 2.5.2).

The IR spectrum recorded of the CQ-ferrihaem complex is shown in Fig. 3.9. For comparison purposes, the IR spectra of ferrihaem precipitated at low and high pH as well as that of CQ (free base), Cl-ferrihaem and μ-oxo ferrihaem are included in Fig 3.9. At first glance, the spectrum of the complex seemed to be a simple combination of the spectra obtained for CQ (free base) and ferrihaem precipitated at pH 5 (see Fig. 3.9a, b and d). By contrast, it was markedly different to that obtained for μ-oxo ferrihaem (Fig. 3.9e) which was characterised by two intense features at 1410 and 880 cm^{-1}. The latter peak was characteristic of μ-oxo species and arises from υ(Fe-O-Fe) stretching.[64] The former band has not been previously assigned. Given the position of the band and the fact that the μ-oxo dimer was isolated as a tetrasodium salt, this peak likely corresponded to the carbonyl stretching of the propionate side chains. A similar peak was observed at the same position in the spectrum of HO-ferrihaem (Fig. 3.9c) which was isolated as a disodium salt. The apparent absence of the υ(Fe-O-Fe) peak from the IR spectrum of the complex supported observations recently made Asghari-Khiavi *et al*. who suggested that the ferrihaem species in the CQ-ferrihaem complex was not the μ-oxo dimer.[181]

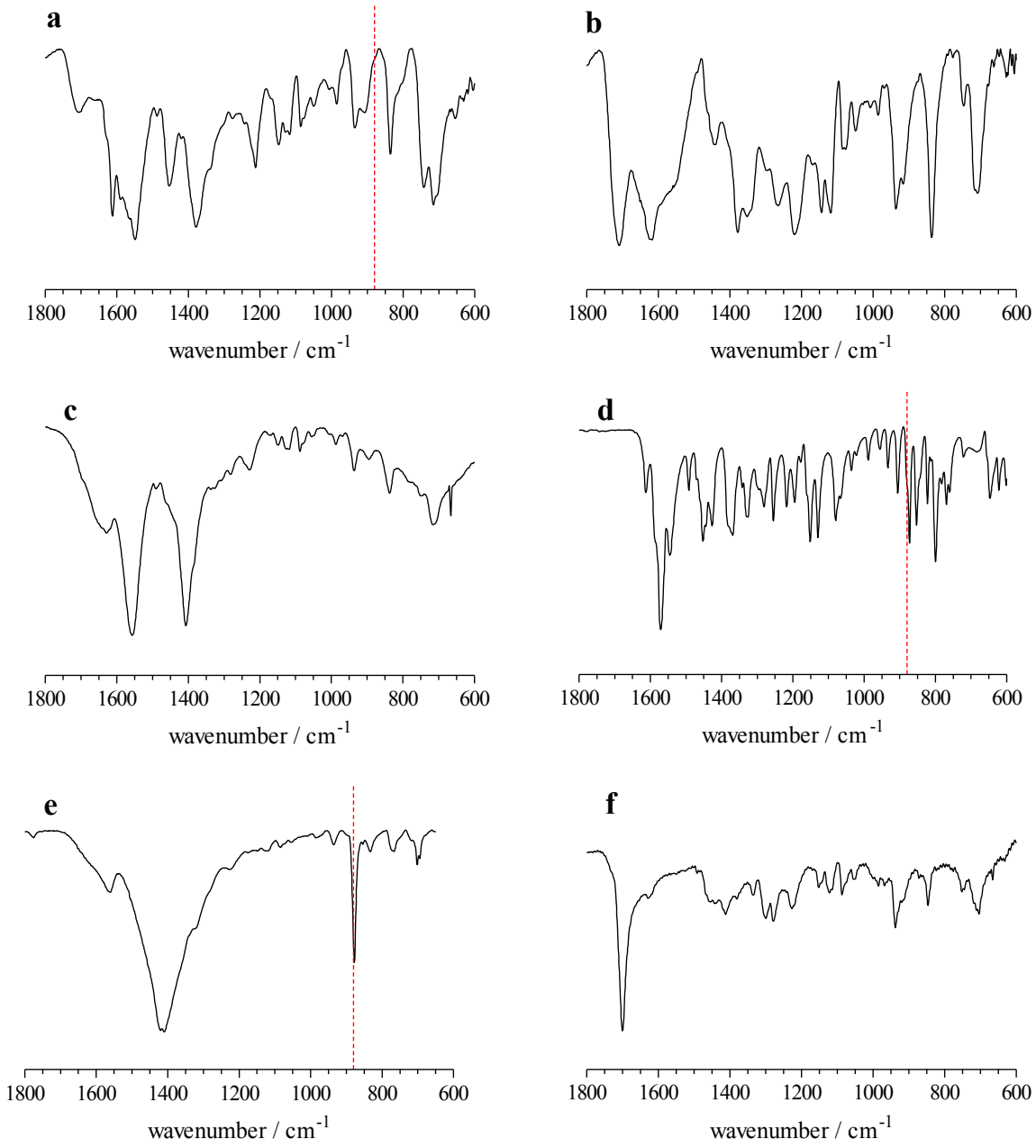


Figure 3.9. IR spectra of ferrihaem species, CQ and the CQ-ferrihaem complex. (**a**) CQ-ferrihaem complex precipitated from aqueous solution, pH 5; (**b**) ferrihaem precipitated from aqueous solution, pH 5; (**c**) ferrihaem obtained from alkaline solution; (**d**) CQ (free base); (**e**) μ-oxo ferrihaem tetrasodium salt precipitated from alkaline acetone solution; (**f**) Cl-ferrihaem. Vertical red lines in (a), (d) and (e) mark 880cm^{-1}, which is the position of the υ(Fe-O-Fe) peak in μ-oxo ferrihaem.

Based solely on the evidence presented above, one may be tempted to conclude that the ferrihaem species in the CQ-ferrihaem complex was the same as that obtained at pH 5 in the absence of ferrihaem and that this species was not the μ-oxo dimer. However, caution must be exercised when performing such simple qualitative comparisons. While it was certainly apparent that the complex and low pH ferrihaem precipitate share similar spectroscopic features, this was because both are isolated at the same pH. Under these conditions, the ferrihaem moiety is expected to have one deprotonated (propionate) and one protonated (propionic acid) side chain which result in peaks around 1400 and 1700 cm^{-1}, respectively. This assignment was confirmed by the spectra of Cl-ferrihaem (Fig. 3.9f) and ferrihaem isolated at high pH. An intense band around 1700 cm^{-1} only was observed in the spectrum of Cl-ferrihaem which has two propionic acid side chains, while no prominent feature was present at 1400 cm^{-1}. The opposite was seen for HO-ferrihaem where no band exists at 1700 cm^{-1} but one was present at 1410 cm^{-1}. By contrast, however, μ-oxo ferrihaem was isolated under alkaline conditions and has four propionate side chains. Thus differences in the regions describing carbonyl stretching of propionate and propionic acid are expected between the complex and μ-oxo ferrihaem tetrasodium salt.

The characteristic υ(Fe-O-Fe) peak of μ-oxo ferrihaem is often used as a marker for this species. Close inspection of the spectrum of the CQ-ferrihaem complex over the range 650 to 1000 cm^{-1} showed no prominent peak at 880 cm^{-1} but rather a shoulder at this position (see Fig. 3.10). Given the broad features of this region of the spectrum and the fact that CQ has an overlapping band at 873 cm^{-1}, the υ(Fe-O-Fe) band could not be used as a definitive marker in this case. Further analysis of the low-wavenumber region revealed that there were in fact no unique features that could be used to identify the form of ferrihaem present because bands were either common to all ferrihaem species or overlapped with CQ peaks. Thus the use of IR spectroscopy to identify the form of ferrihaem was not conclusive. On the other hand, however, the IR spectrum of the complex provided additional evidence against the structural model suggested by Schwedhelm *et al*.[170] The proposed model involved reciprocal coordination of propionate side chain to a neighbouring μ-oxo dimer, similar to that observed in β-haematin. In the case of β-haematin, intense IR bands at 1660 and 1210 cm^{-1} are observed which arise from carbonyl υ(C=O) and υ(C-O) stretching of the coordinated propionate respectively and are characteristic of this species. The model proposed by Schwedhelm *et al*. should have similar peaks in the IR spectrum but were clearly absent from that recorded for the CQ-ferrihaem complex (Fig. 3.9a).

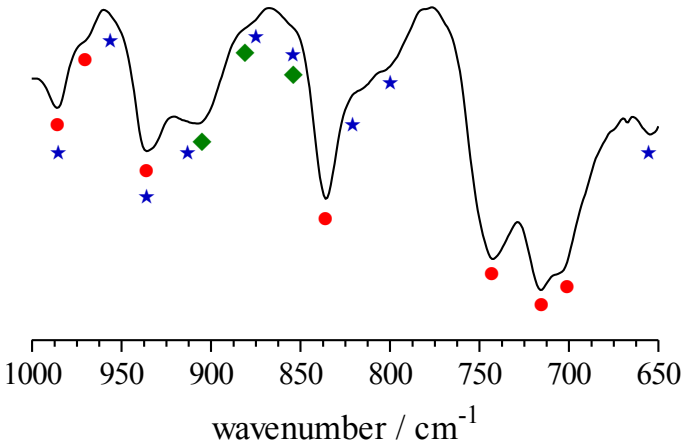


Figure 3.10. IR spectrum of the CQ-ferrihaem complex precipitated at pH 5. Red dots indicate the position of features common to all ferrihaem species presented in Fig. 3.9, blue stars represent features present in CQ (free base) and green diamonds are features present in the spectrum of μ-oxo ferrihaem only.

By contrast, magnetic susceptibility measurements of the solid samples were more definitive (see Table 3.2). Cl-ferrihaem was used as a control and the value obtained (6.2 ± 0.2 μ_B) compares well with the predicted value for high-spin (S = $^5/_2$) species (5.92 μ_B). The magnetic moment determined for the CQ-ferrihaem complex precipitated at pH 5 was in excellent agreement with that determined for the complex in aqueous solution (see Section 3.3.2 and Table 3.2), both of which compare well with that obtained for the μ-oxo species (2.0 ± 0.1 μ_B). Furthermore, the value determined for the complex was within experimental error of the predicted value at 298 K reported by Stanek and Dziedzic-Kocurek for μ-oxo ferrihaem.[220] These results suggested that the form of ferrihaem in the CQ-ferrihaem complex was the same as observed in solution and in both cases was μ-oxo ferrihaem. This conclusion was supported by previous Mössbauer studies undertaken on the complex precipitated at low pH.[254]

Table 3.2. Magnetic moments per Fe(III) center of ferrihaem species (μ_B) as determined in solution by the Evans NMR method and in the solid state using an Evans balance.

	Ferrihaem species	μ_B
Solution-state		
	ferrihaem [a, b]	4.8 ± 0.1
	μ-oxo ferrihaem [c]	2.0 ± 0.1
	CQ-ferrihaem [d, b]	2.25 ± 0.02
Solid-state		
	Cl-ferrihaem	6.2 ± 0.2
	ferrihaem [e, f]	4.0 ± 0.2
	CQ-ferrihaem [d, f]	2.3 ± 0.1

[a] π-π dimer, axial ligand H_2O/HO^-; [b] aqueous, pH 7.4; [c] 5.64 M aqueous DMSO, pH 10; [d] 1:2 CQ:ferrihaem; [e] π-π dimer, axial ligand H_2O; [f] precipitated from aqueous solution at pH 5.

Given the convincing evidence provided by magnetic susceptibility measurements that the μ-oxo dimer existed in the solid CQ-ferrihaem complex, the apparent absence of characteristic υ(Fe-O-Fe) in the IR spectrum recorded on the very same sample was surprising. As mentioned above, based on the broad features in the IR spectrum over the range 650 to 1000 cm^{-1} and overlapping CQ peaks, it could only be concluded that the peak at 880 cm^{-1} could not be used as a definitive marker for μ-oxo dimer formation for this system. A similar situation was reported by Ercolani *et al*. for two interchangeable forms of μ-oxo dimer of iron(III)phthalocyanin.[271] Despite definitive magnetic and compositional evidence for μ-oxo dimer formation for both species, only one displayed the characteristic υ(Fe-O-Fe) stretching band in the IR. Lever and co-workers also reported a similar absence of IR stretch in the case of some μ-oxo manganese(III)phthalocyanins.[272] Neither of the authors provided an explanation for the disappearance of this band from the IR spectrum. In the case of the CQ-ferrihaem, however, it is suspected that the IR peak is present but as a shoulder which was not clearly resolved owing to overlap with CQ peaks. It would be interesting to perform further experiments to probe this using techniques such as RR and EPR spectroscopy which could not be conducted in this thesis owing to a lack of equipment.

3.4 Conclusions

UV-visible and MCD spectroscopy as well as magnetic susceptibility measurements indicated that CQ induces μ-oxo dimer formation in the CQ-ferrihaem complex from the π-π dimeric species in solution. The same species was present in the complex precipitated at pH 5 although the characteristic υ(Fe-O-Fe) peak at 880 cm^{-1} was not observed. The absence of this peak was most likely because of overlap with CQ bands. The stoichiometry of interaction was found to be 1:2 CQ:ferrihaem in both aqueous solution and the solid state. The aggregation state of the complex in aqueous solution was found to be tetrameric $[CQ\text{-}(\mu\text{-oxo ferrihaem})]_2$ although owing to experimental errors dimeric CQ-(μ-oxo ferrihaem) and hexameric $[CQ\text{-}(\mu\text{-oxo ferrihaem})]_3$ species could not be entirely discounted. Higher order aggregates were not consistent with the findings.

While the results of this study provided a clearer description of the interaction of CQ and ferrihaem in aqueous solution, they were unable to provide a structural description of the complex.

CHAPTER 4

QUANTUM MECHANICAL CALCULATIONS OF FERRIHAEM SPECIES

4.1 Introduction

In order to perform meaningful computational MD simulations of ferrihaem species, accurate structural models of these molecules were required. At present, only six crystal structures of ferrihaem exist where this prosthetic group is not bound by a protein. These include three ferrihaem-antimalarial drug complexes,[78, 156] Cl-ferrihaem (haemin)[178] and two structures of β-haematin.[53, 154] Additionally, a seventh closely related crystal structure of ferrihaem dimethyl ester has been reported in the μ-oxo dimeric form.[72] On the other hand, no reported crystal structures exist for monomeric H_2O- and HO-ferrihaem which make up the dominant π-π dimer species in aqueous solution. Therefore quantum mechanical calculations were required to produce geometries and atomic charges of ferrihaem species for use in structural model parameterisation (see Chapter 5).

Relatively few QM investigations have been conducted on the ferrihaem species of interest in this study. Reports of DFT calculations in the literature have been limited to either Cl-ferrihaem[273, 274] or μ-propionato ferrihaem[275-277] and its derivatives[278]. These studies utilised QM-derived data to obtain a clearer understanding of the geometric structure, vibrational frequencies, electronic transitions and thermodynamic stability of ferrihaem species. This chapter details the DFT methods used to obtain structures of monomeric H_2O- and HO-ferrihaem species, as well as dimeric μ-oxo and μ-propionato ferrihaem molecules, with the aim to use them for parameterisation of MM ferrihaem models (see Chapter 5). Given the strong dependence of experimental investigations of ferrihaem and its interaction with antimalarials on IR,[122, 126, 216, 254, 279] RR[171, 182, 280-283] and UV-visible[75, 76, 163, 177, 183] spectroscopic measurements, theoretical vibrational stretches and electronic excitations were also calculated. These were compared with experiment both as a test of the QM structures and to gain deeper insight into the origins of the observed spectroscopic features.

4.2 Methods

4.2.1 General.

All DFT and TD-DFT calculations were performed using the Gaussian09 software package[284] at the Centre for High Performance Computing (CHPC), South Africa. Geometry optimisation calculations were carried out over 32 Nehalem processors (2.93 GHz) and required approx. 2 weeks continuous calculation to complete in the case of monomeric ferrihaem species, and approx. 4 weeks in the case of μ-oxo ferrihaem. Editing of structures was conducted using the MOLDEN and Avogadro software packages.[285] Analysis and visualisation of calculated IR and UV-visible spectra were performed using the GABEDIT and SWIZARD software packages.[287-289] Computed IR spectra were convoluted using Lorentzian functions with 5 cm^{-1} half-width while calculated UV-visible spectra were convoluted using Gaussian functions with half-widths of 700 cm^{-1}. MOs were analysed using the Chemissian software package and visualised using GABEDIT on an electrostatic potential isosurface of 0.02 au.

4.2.2 DFT geometry optimisation.

Geometry optimisation of ferrihaem and ferriporphine molecules was conducted using the OPBE functional and LANL2DZ basis set[241-244] (OPBE/LANL2DZ). Starting geometries of H_2O-/HO-ferrihaem and ferriporphine molecules were constructed by modifying the crystal structure of β-haematin,[53] while those of μ-oxo ferrihaem and ferriporphine were obtained by modifying the crystal structure of μ-oxo ferrihaem dimethyl ester.[72] The initial geometry of μ-propionato ferrihaem was taken from the crystal structure of β-haematin.[53] Aqua species were modelled in the high-spin ($S = {}^5/_2$) state while the hydroxo species were modelled in both the high- and intermediate-spin states ($S = {}^3/_2$). To correctly model the antiferromagnetic coupling of μ-oxo species, the method of broken symmetry, as described by Noodleman and co-workers[290], was used. This involved geometry optimisation of high-spin states ($S = {}^{10}/_2$) after which electrons on one ferriporphyrin/porphine moiety were flipped and the structure re-optimised for this spin-coupled state. The μ-propionato species was modelled in the high-spin ($S = {}^{10}/_2$) state. Structures were calculated without orbital symmetry constraints and a quadratically convergent

self-consistent field (SCF) procedure was required to ensure convergence.[291] Furthermore, frequencies were determined to ensure a stationary point was obtained and wavefunction stabilities tested.

4.2.3 TD-DFT calculations.

Electronic spectra were calculated for DFT-optimised structures of [H_2O-ferrihaem]$^{1+}$ and [μ-oxo ferrihaem]$^{4-}$ using TD-DFT and the PBE0 functional.[292] The LANL2DZ basis set was used to describe the iron atom while LANL2DZdp basis set which includes diffuse and polarisation functions was used to describe the remaining atoms (PBE0/LANL2DZ:LANL2DZdp).[293] Calculations were performed under the implicit solvent conditions of the SMD water model.[294] Molecular orbitals were calculated using the same functional and basis set combination in the presence of implicit water (SMD). Calculations were conducted for [H_2O-ferrihaem]$^{1+}$ and [μ-oxo ferrihaem]$^{4-}$ used 120 and 360 Intel Westmere processors and required approx. 2 and 4 weeks of continuous computing time to complete, respectively.

4.2.4 UV-visible and MCD spectroscopic deconvolution.

UV-visible and MCD spectra were recorded as detailed in Section 3.2.6. Deconvolution was performed by simultaneous fitting of Gaussian functions according to equation 4.1 using Microsoft Excel with Solver add-in.

$$f(E) = \sum_i a_{i,UV-vis/MCD}\, exp^{\frac{-(E_i - E_{0,i})^2}{{\sigma_i}^2}} \qquad (4.1)$$

In eq. 4.1, E is energy (in cm^{-1}), E_0 is the peak position, σ the bandwidth (in cm^{-1}) and $a_{UV\text{-}vis/MCD}$ the intensity. Simultaneous deconvolution was performed under the rigid shift approximation which assumes that the applied magnetic field has no effect on the spectroscopic bandshape of excitations.[198, 199] Fitted peaks in the UV-visible spectrum were constrained to the same E_0 and σ values as the corresponding peaks in the MCD spectrum.[295, 296] UV-visible and MCD intensities, $a_{UV\text{-}vis/MCD}$, were allowed to vary independently because the mechanism of MCD intensity involves the magnetic dipole moment as well as the transition dipole moment whereas the UV-visible intensity is dependent on the

transition dipole moment only.[198, 199] Spectra were fitted with the minimum number of bands to produce acceptable χ^2 values. Full width at half maxima, ΔE, was calculated using equation 4.2.

$$\Delta E = 2\sigma\sqrt{\ln 2} \quad (4.2)$$

4.2.5 IR spectra.

IR spectra were recorded on a PerkinElmer Spectrum 100 FT-IR spectrophotometer using an ATR attachment. The IR spectra of H_2O-, HO- and μ-oxo ferrihaem were recorded on samples prepared as described in Section 3.2.3. The β-haematin sample used to record an IR spectrum was prepared according to a modified procedure reported previously.[254] Briefly, this involved dissolving haemin (15 mg) in 2.5 mL NaOH (0.1 M) which was neutralised with 0.25 mL HCl (1 M). To the resulting precipitate, 2.31 mL sodium acetate (9.7 M, pH 4.8) was added and the stirred mixture left to incubate for 1 hour at 60 °C. The solid was filtered on cellulose nitrate filter discs (0.22 μm), washed extensively with water and placed in a desiccator over phosphorous pentoxide to dry.

4.2.6 Magnetic susceptibility measurements.

Magnetic susceptibility of H_2O-, HO- and μ-oxo ferrihaem was determined in solution using the Evans NMR method (see Section 2.5.1 for details). NMR spectra were recorded on a Bruker Ultrashield 400 Plus spectrophotometer at 303 K. In the case of H_2O- and μ-oxo ferrihaem, samples were prepared as detailed in Section 3.2.7. To determine the magnetic moment of HO-ferrihaem, haematin (3 mg) was dissolved in 0.1 mL NaOD (1 M) and 0.9 mL MeOD. The reference solution used was prepared in the same manner with the exclusion of haematin.

4.3 Structure Determination of Ferrihaem Species using DFT Calculations

4.3.1 Functional and Basis Set Investigation.

To determine the functional and basis set combination which would produce reliable structural models of ferrihaem species, three functionals and two basis set combinations were tested on the anionic $[H_2O\text{-ferrihaem}]^{1-}$ molecule. Calculations were performed for the high-spin ($S = {}^5/_2$) state that was found for this species using magnetic susceptibility measurements (see Table 4.1). Levels of theory employed included BP86,[236, 297] M06[298] and OPBE.[299, 300] BP86 was chosen because studies on iron(III)OEP by Schiedt and co-workers using this functional produced results in good agreement with experimental nuclear resonance vibrational spectra.[301] M06 was selected based on its reported ability to describe dispersion interactions.[298] OPBE, which combines the OPTX optimised exchange functional[299] with PBE correlation functional,[300] was used since investigations by Swart *et al*. indicated that accurate geometries and spin-state energies were obtained for iron systems using this functional.[302, 303] Owing to the size of the ferrihaem molecules of interest (particularly the dimeric μ-oxo and μ-propionato species), the effective core potential basis set, LANL2DZ, was employed in both basis set combinations investigated in an effort to reduce the computational cost of calculations. The first involved all atoms being described by LANL2DZ. Results from such calculations were compared to the second combination where only the iron atom was described by LANL2DZ and the remaining atoms described by the larger 6-31G(2d,p) basis set which includes polarisation functions. This second basis set combination is termed LANL2DZ:6-31G(2d,p).

Table 4.1. Magnetic moments (μ_B) and suggested spin-states of ferrihaem species in solution determined using the Evans NMR method.

Species	μ_B	Spin-state
H_2O-ferrihaem [a]	5.94	$S = {}^5/_2$
HO-ferrihaem [b]	4.34	$S = {}^5/_2$ and ${}^3/_2$ admixture
μ-Oxo ferrihaem [c]	2.13	$S = {}^5/_2$, antiferromagnetic

[a] In 80% aq. acetone, pH 2.0; [b] In 90% aq. alkaline methanol; [c] In 40% aq. DMSO, pH 12.

Of the six calculations employing the above-mentioned functional and basis set combinations, only that involving OPBE and LANL2DZ, the latter describing all atoms, (OPBE/LANL2DZ) produced an acceptable geometry. In the case of the BP86 functional, despite employing a SCF procedure, convergence was not achieved when calculated using either basis set combination. While convergence was attained with OPBE and M06 functionals employing the LANL2DZ:6-31G(2d,p) basis set combination, these calculations produced unrealistic geometries where the O-H bonds of the axial water ligand were orientated parallel to that of the porphyrin plane (see Fig. 4.1a). Lastly, using M06 in conjunction with LANL2DZ produced a geometry in which a proton transfer from the axial water ligand to propionate side chain occurred (see Fig 4.1b). While this proton transfer is certainly feasible, the M06/LANL2DZ calculation is unable to produce the anionic H_2O-ferrihaem species which has been shown to exist in aqueous solution and thus could not be used.

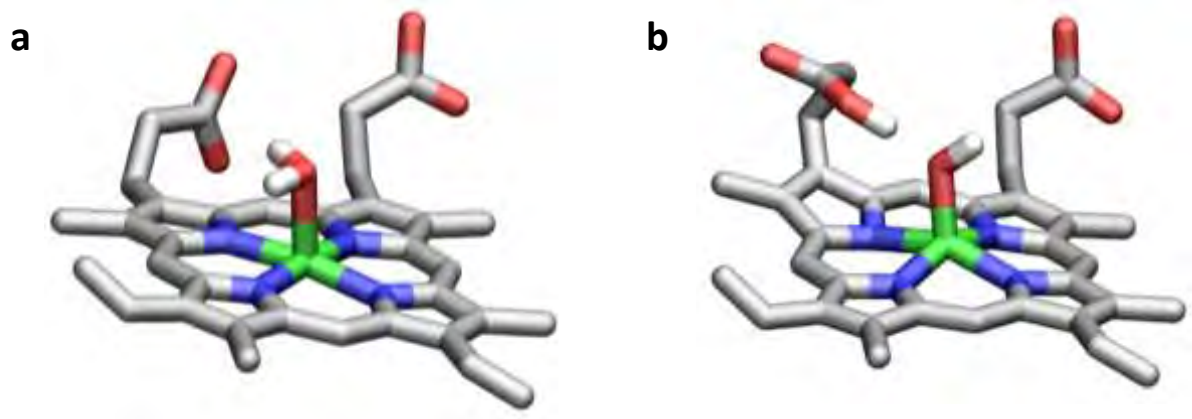


Figure 4.1. Unsatisfactory geometry-optimised structures of $[H_2O\text{-ferrihaem}]^{1-}$. (**a**) Structure obtained using OPBE and M06 functionals in conjunction with LANL2DZ:6-31G(2d,p) basis set combination. Axial O-H bonds were unrealistically orientated parallel to the porphyrin plane. (**b**) Structure obtained using the M06 functional and LANL2DZ basis set where a proton transfer from axial water ligand to propionate side chain was observed. Porphyrin hydrogen atoms have been omitted for clarity.

With OPBE/LANL2DZ identified as the best method with which to calculate the structure of [H_2O-ferrihaem]$^{1-}$, DFT calculations on the remaining species were performed. These included [H_2O-ferrihaem]^{0}, [H_2O-ferrihaem]$^{1+}$, [HO-ferrihaem]$^{2-}$, [μ-oxo ferrihaem]$^{4-}$ and [μ-propionato ferrihaem]^{0}. As in the case of [H_2O-ferrihaem]$^{1-}$, calculations of the neutral and cationic forms of this species were conducted in the high-spin (S = $^5/_2$) state. For [HO-ferrihaem]$^{2-}$, magnetic susceptibility measurements (see Table 4.1 above) suggested that this species existed as a quantum mechanical admixture of high- (S = $^5/_2$) and intermediate-spin (S = $^3/_2$) states. Since DFT calculations cannot describe such a spin-state situation, separate calculations for [HO-ferrihaem]$^{2-}$ were performed for the high- and intermediate-spin states. To correctly describe the observed antiferromagnetic coupling between iron centres of [μ-oxo ferrihaem]$^{4-}$, the method of broken symmetry was adopted (see Section 4.2.2 for details regarding broken symmetry). Calculations on the [μ-propionato ferrihaem]^{0} species were performed for the high-spin (S = $^{10}/_2$) state, as indicated from Mössbauer and EPR data.[254, 304]

4.3.2 H_2O- and HO-Ferrihaem Species.

As seen from Table 4.2, a decrease in Fe-O bond length is observed for H_2O-ferrihaem species upon protonation of the propionate side chains indicating the charge-state of these groups affect the axial water ligand of ferrihaem. While the cationic species displays the closest value to experiment, it must be noted that of the thirteen relevant crystal structures reported in the CSD which were used to obtain the experimental Fe-O value for H_2O iron(III)porphyrin ligands, only four are five-coordinate. The average Fe-O bond length for these four structures is 2.06 ± 0.02 Å which is approx. 0.1 Å shorter than that obtained for [H_2O-ferrihaem]$^{1+}$. A similar deviation from experimental value was reported for the Fe-Cl bond in Cl-iron(III)TPP by Paulat *et al*. when using the LANL2DZ basis set to describe non-metal atoms.[207] These authors found that by describing non-metal atoms by the LANL2DZdp basis set (which includes polarisation functions), the deviation significantly decreased. The same approach was applied to [H_2O-ferrihaem]$^{1+}$, however this produced a structure similar to the unrealistic geometry observed for the anionic species when using the LANL2DZ:6-31G(2d,p) basis set combination (see Fig. 4.1a).

The Fe-O bond length of the structure optimised using OPBE/LANL2DZ is reproduced to within two standard deviations of the experimental values, although it is in poorer agreement with the average value obtained using only the four five-coordinate molecules. It must be emphasized, however, that none of the experimental structures used to determine this value actually involve ferrihaem and thus some structural differences are likely. Indeed, a marked difference of approx. 0.1 Å in the Fe-O bond length is observed for the unsubstituted $[H_2O\text{-ferriporphine}]^{1+}$ species as compared to its cationic ferrihaem counterpart. Conversely, relatively little change was observed between Fe-N, O-Fe-N, N-Fe-N or iron out of plane deviation values. DFT-calculated H_2O-ferrihaem and ferriporphine Fe-N bonds were slightly longer than average experimental values but fell within the reported range of 1.945 – 2.061 Å. O-Fe-N and N-Fe-N angles as well as the iron out of plane deviation of DFT-optimised structures reproduced average CSD values well. Final geometry-optimised structures of all H_2O-ferrihaem species are displayed in Fig. 4.2a-c.

Table 4.2. Selected structural parameters of DFT-optimised H_2O- and HO-ferrihaem and ferriporphine species.

	Spin-State	Fe-O_{ax}	Fe-N	N-Fe-O	N-Fe-N	OOP[a]
$[H_2O\text{-ferrihaem}]^{1-}$	$^5/_2$	2.276	2.077(6)	95(5)	89.6(5)	0.2
$[H_2O\text{-ferrihaem}]^{0}$	$^5/_2$	2.199	2.074(8)	96(2)	89.4(3)	0.2
$[H_2O\text{-ferrihaem}]^{1+}$	$^5/_2$	2.169	2.062(6)	97(2)	89.1(2)	0.3
$[H_2O\text{-ferriporphine}]^{1+}$	$^5/_2$	2.156	2.064(3)	99(1)	88.7(1)	0.3
CSD average		*2.09(6)*	*2.00(4)*	*93(3)*	*89.7(4)*	*0.23(4)*[b]
$[HO\text{-ferrihaem}]^{2-}$	$^5/_2$	1.878	2.13(2)	103.7(3)	86.8(6)	0.5
$[HO\text{-ferrihaem}]^{2-}$	$^3/_2$	1.903	2.06(2)	99.3(8)	88.5(5)	0.3
$[HO\text{-ferriporphine}]^{0}$	$^5/_2$	1.834	2.110(3)	103.4(9)	86.9(3)	0.5
$[HO\text{-ferriporphine}]^{0}$	$^3/_2$	1.882	2.027(3)	97.9(4)	88.9(1)	0.3
CSD average		*1.85(2)*	*2.07(1)*	*102.6(8)*	*87.3(4)*	*0.45(3)*

[a] Out of plane iron deviation. Planes were calculated using the four nitrogen atoms of protoporphyrin IX; [b] For five-coordinate species.

With regard to HO-ferrihaem species, Fe-O bond lengths of DFT-calculated structures were again slightly longer than the average CSD values, with the high-spin ($S = {}^5/_2$) species displaying the closest value to experiment. As similarly noted for H_2O-ferrihaem species, the observed deviation from

experiment may be attributed to the fact that the crystal structures used to obtain the average values are of related systems and none involve HO-ferrihaem specifically. Additionally, in this case, only one of the reported iron(III)porphyrin structures actually contains an axial hydroxide ligand.[305] Owing to the lack of hydroxide-coordinated iron(III)porphyrins in the CSD, crystal structures consisting of alkoxide-coordinated ligands were also included so as to enlarge the data set. In general, the structure of high-spin $[HO\text{-ferrihaem}]^{2-}$ agrees more closely with experiment than that of the intermediate-spin ($S = {}^{3}/_{2}$) species, with the exception of the Fe-N bond length. Minor decreases in Fe-N bond length, O-Fe-N angle and iron out of plane deviation as well as corresponding increases in Fe-O bond length and N-Fe-N angle were observed when the spin-state was changed from high- to intermediate-spin. Owing to the slight differences in structure between the high- and intermediate-spin species, the final geometry-optimised structure of only the former is displayed in Fig. 4.2d. The same structural changes as those observed for HO-ferrihaem species were mirrored by the HO-ferriporphine counterparts.

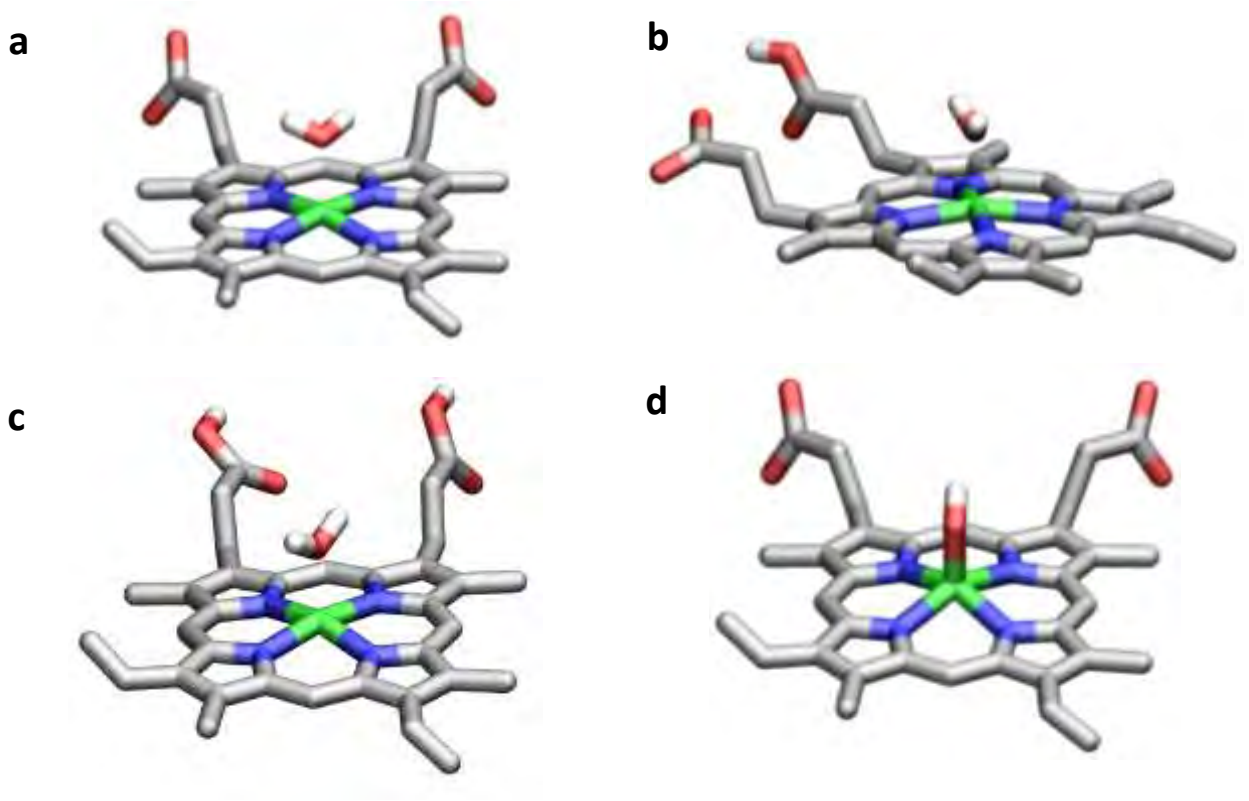


Figure 4.2. High-spin ($S = {}^{5}/_{2}$), geometry-optimised structures of (**a**) $[H_2O\text{-ferrihaem}]^{1-}$; (**b**) $[H_2O\text{-ferrihaem}]^{0}$; (**c**) $[H_2O\text{-ferrihaem}]^{1+}$; and (**d**) $[HO\text{-ferrihaem}]^{2-}$ produced using OPBE/LANL2DZ. Porphyrin hydrogen atoms have been removed for clarity.

4.3.3 μ-Oxo and μ-Propionato Ferrihaem Species.

Table 4.3 lists selected bond lengths, angles and out of plane deviations of iron atoms for DFT-calculated [μ-oxo ferrihaem]$^{4-}$, [μ-oxo ferriporphine]^{0} and μ-propionato ferrihaem species. In the case of [μ-oxo ferrihaem]$^{4-}$, a dramatic decrease was observed in Fe-O bond length of approx. 0.6 Å between the ferromagnetic high-spin (S = $^{10}/_{2}$) and antiferromagnetic forms. The same reduction in Fe-O bond length was also observed for high-spin and antiferromagnetic μ-oxo ferriporphine species. The Fe-O bond length of antiferromagnetic [μ-oxo ferrihaem]$^{4-}$ is somewhat longer than the average Fe-O bond length observed for μ-oxo iron(III)porphyrins in the CSD and falls just above the reported range of 1.740 – 1.786 Å. No other significant change to either ferrihaem or ferriporphine structure was observed when the high-spin species were modelled antiferromagnetically. While Fe-N bond lengths in both ferrihaem and ferriporphine species are similarly longer than experiment, the N-Fe-O, N-Fe-N and Fe-O-Fe angles all fall within the experimental error of the reported average values. The final geometry obtained for [μ-oxo ferrihaem]$^{4-}$ using broken-symmetry DFT calculations to model antiferromagnetic coupling, is shown in Fig. 4.3.

Table 4.3. Selected structural parameters of DFT-optimised μ-oxo ferrihaem, μ-oxo ferriporphine and μ-propionato ferrihaem species.

	Spin-State	Fe-O_{ax}	Fe-N	N-Fe-O	N-Fe-N	Fe-O-Fe/C	OOP [b]
[μ-Oxo ferrihaem]$^{4-}$	$^{10}/_{2}$	1.863(9)	2.137(9)	105(5)	86.0(5)	177.3	0.6
[μ-Oxo ferrihaem]$^{4-}$	(AF) [a]	1.804(1)	2.134(9)	105(5)	86.1(4)	177.9	0.6
[μ-Oxo ferriporphine]^{0}	$^{10}/_{2}$	1.853(2)	2.115(1)	104.1(2)	86.6(1)	179.7	0.5
[μ-Oxo ferriporphine]^{0}	(AF) [a]	1.794(1)	2.114(1)	104.0(2)	86.6(1)	179.9	0.5
CSD average		*1.76(1)*	*2.081(8)*	*104(1)*	*86.8(4)*	*176(4)*	*0.49(3)*
[μ-Propionato ferrihaem]^{0}	$^{10}/_{2}$	1.992(1)	2.094(5)	102(2)	87.4(2)	167.3(8)	0.5
[μ-Propionato ferrihaem]$^{2-}$	$^{10}/_{2}$	1.899(2)	2.11(1)	103(2)	86.9(3)	159.6(6)	0.5
CSD average		*1.893(1)*	*2.060(1)*	*102.9(5)*	*87.3(4)*	*144(3)*	*0.45(3)*

[a] AF = antiferromagnetically coupled, modelled using broken symmetry; [b] Out of plane iron deviation. Planes were calculated using the four nitrogen atoms of protoporphyrin IX.

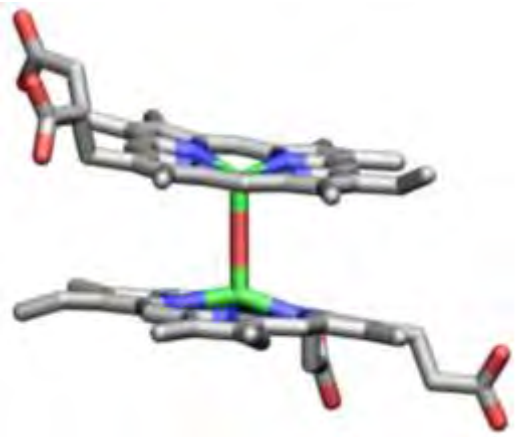


Figure 4.3. Geometry-optimised structure of antiferromagnetic [μ-oxo ferrihaem]$^{4-}$ produced using OPBE/LANL2DZ. Porphyrin hydrogen atoms removed for clarity.

Geometry optimisation of [μ-propionato ferrihaem]^{0} produced a structure in which an intramolecular hydrogen bond between the coordinated propionate side chains and nearby propionic acid groups were observed (see Fig. 4.4a). This differs from the two reported crystal structures of β-haematin where intermolecular hydrogen bonding is observed between either a neighbouring μ-propionato ferrihaem dimer or a DMSO solvent molecule.[53, 154] Since the DFT calculation is conducted on a single μ-propionato ferrihaem molecule with no neighbouring dimers or solvent molecules to hydrogen bond with, the formation of the intramolecular bond can be rationalised. This geometry produces a Fe-O bond length which is markedly longer than that of the average value for the two reported crystal structures (1.99 vs 1.89 Å). However, when the possibility of an intramolecular hydrogen bond was removed by deprotonating the propionic acid side chains to give anionic [μ-propionato ferrihaem]$^{2-}$, the Fe-O bond length of the optimised geometry agreed well with experiment (see Fig, 4.4b for optimised structure). The shortening of Fe-O bond in the anionic species was accompanied by a corresponding lengthening of Fe-N bond which is slightly larger than the average experimental value. Little to no difference in N-Fe-O and N-Fe-N angles or iron out of plane deviation was observed between neutral and anionic species, however both have significantly larger Fe-O-C angles than those reported for β-haematin. The Fe-O-C angle deviation can possibly be explained by crystal packing effects as coordinated propionates of one μ-propionato dimeric unit are in close proximity to the propionic acid group of a neighbouring unit and likely cause the structural differences observed (see Fig. 4.4c).

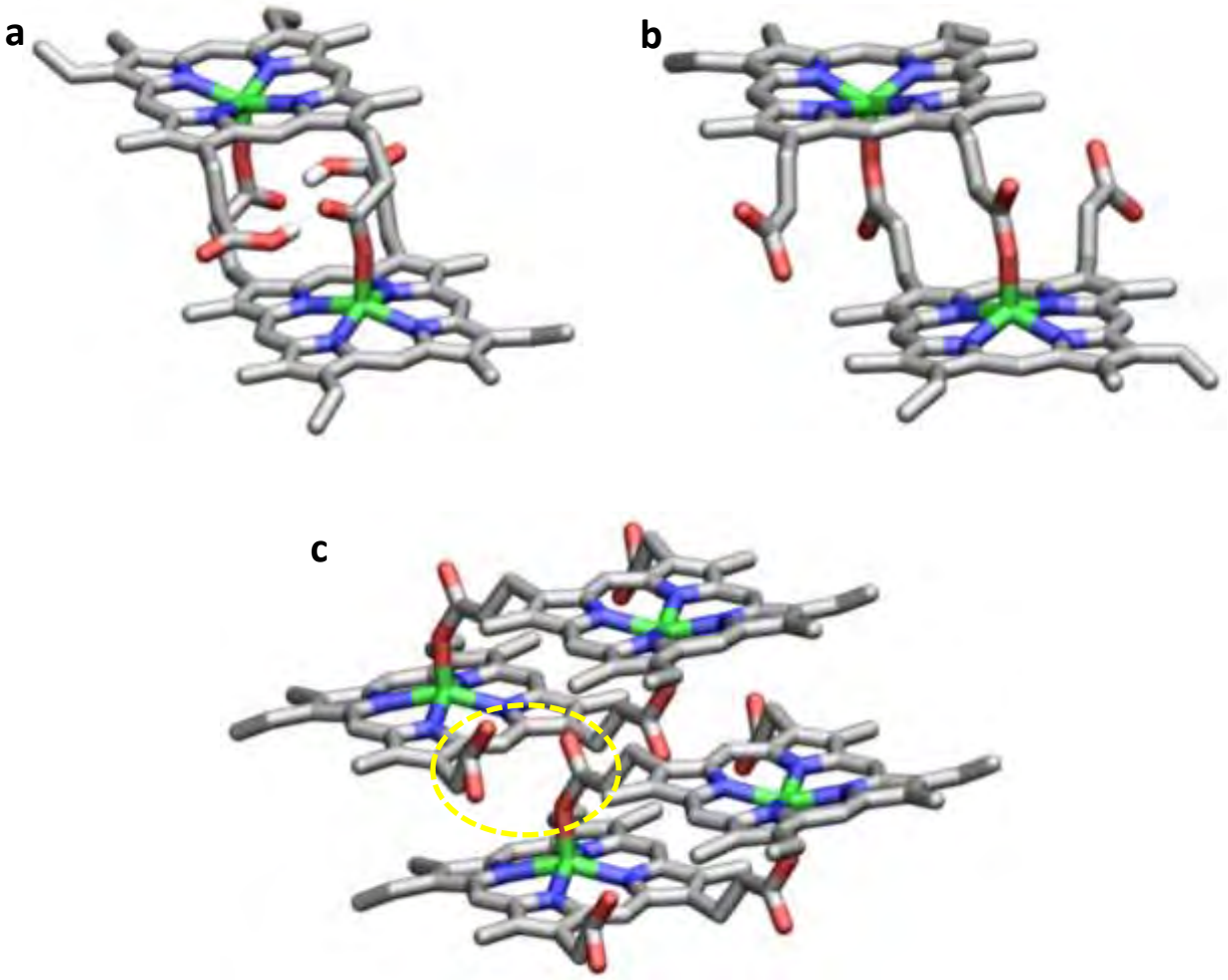


Figure 4.4. Geometry-optimised structures of high-spin (S = $^{10}/_{2}$) (**a**) [μ-propionato ferrihaem]^{0} and (**b**) [μ-propionato ferrihaem]$^{2-}$ produced using OPBE/LANL2DZ. (**c**) Crystal structure of β-haematin reported by Pagola *et al.*[53] Yellow dotted region highlights possible crystal packing interactions that are suspected to influence the Fe-O-C angle and were not reproduced by DFT-calculation. Porphyrin hydrogen atoms have been removed for clarity.

4.4 Vibrational Analysis of Ferrihaem Species

Vibrational spectroscopy, specifically IR, is routinely used to identify and investigate ferrihaem compounds. While detailed computational investigations have been reported for μ-propionato ferrihaem with regard to Raman spectroscopy,[275, 276] no such studies of IR vibrations have been reported for other ferrihaem molecules. Given the extensive experimental data reported in this regard and with DFT-calculated ferrihaem models in hand, the opportunity was presented to compare computed vibrational modes with IR evidence. The main objective to do this was to provide experimental support for the computed structures as well as to investigate whether previous assignments made were indeed supported by first-principle calculations.

Vibrational analysis was performed for ferrihaem species in which experimental IR spectra could be obtained. These included $[H_2O\text{-ferrihaem}]^0$, $[HO\text{-ferrihaem}]^{2-}$, $[\mu\text{-oxo ferrihaem}]^{4-}$ and $[\mu\text{-propionato ferrihaem}]^0$. Theoretical IR spectra are presented in Fig. 4.5 and 4.6 alongside their experimental counterparts.

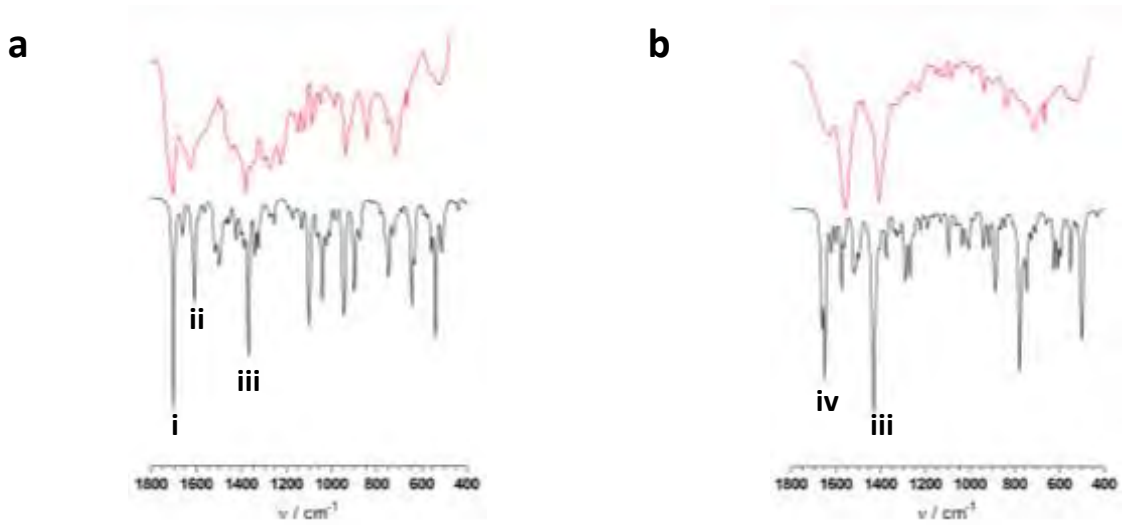


Figure 4.5. Computed (black) and experimental (red) IR spectra of monomeric (**a**) $[H_2O\text{-ferrihaem}]^0$; and (**b**) $[HO\text{-ferrihaem}]^{2-}$. See text for assignment of vibrational modes (i) – (iv).

In general, there is good agreement between calculated and experimental IR spectra despite no scaling factors being used. Characteristic features of each ferrihaem species are well reproduced from computation although, in some cases, they are shifted in wavenumber with respect to their experimental counterparts. The computed spectrum of monomeric $[H_2O\text{-ferrihaem}]^0$ displays three important bands at 1702, 1606 and 1367 cm^{-1}, labelled (i) – (iii) in Fig. 4.5a. The first and third bands, (i) and (iii), represent the carbonyl stretching frequencies, υ(C=O), of propionic acid and propionate groups, respectively. These are in good agreement with the corresponding experimental peaks at 1704 and 1382 cm^{-1}, respectively. The second intense band in the computed spectrum, (ii), is produced by the H-O-H bending motion of the axial water ligand. This mode agrees well with the observed IR peak at approx. 1630 cm^{-1} which has not previously been assigned. The experimental IR spectrum of monomeric $[HO\text{-ferrihaem}]^{2-}$ does not appear to have been previously reported (see Fig. 4.5b). Two strong peaks dominate the spectrum at higher wavenumber which were well reproduced by computation. The peak at 1434 cm^{-1} in the simulated spectrum arises from the propionate υ(C=O) stretch, thus is labelled (iii) in Fig. 4.5b, and likely corresponds to the experimentally observed band at 1405 cm^{-1}. The second peak at 1648 cm^{-1} originates from the vinyl υ(C=C) stretching mode, labelled (iv) in Fig. 4.5b, and given its intensity relative to the υ(C=O) feature can possibly be assigned to the similar experimental band at 1560 cm^{-1}.

Features present in the theoretical spectrum of dimeric $[\mu\text{-propionato ferrihaem}]^0$ agree well with the experimental spectrum recorded for β-haematin, although are shifted in some cases (see Fig. 4.6a). The three characteristic bands used to identify β-haematin occur at 1709, 1660 and 1210 cm^{-1}. The first arises from the υ(C=O) stretch of free propionic acid group while the second and third are due to the υ(C=O) and υ(C-O) stretching frequencies of iron-coordinated propionate groups, respectively. The calculated mode corresponding to the first IR band, labelled (i) in Fig. 4.6a, occurs at 1700 cm^{-1} and is in good agreement with experiment. Computed stretching frequencies of the coordinated propionate groups are somewhat shifted as compared to experimental counterparts. The υ(C=O) mode is slightly lower in wavenumber than experiment, labelled (iii) in Fig. 4.6a, and occurs at 1570 cm^{-1} while υ(C-O), labelled (v) in Fig. 4.6a, is a little higher at 1275 cm^{-1}. These minor shifts, however, are not surprising given the intramolecular bond formed between the propionic acid and propionate groups in the computed molecule which is not present in the structure of β-haematin.

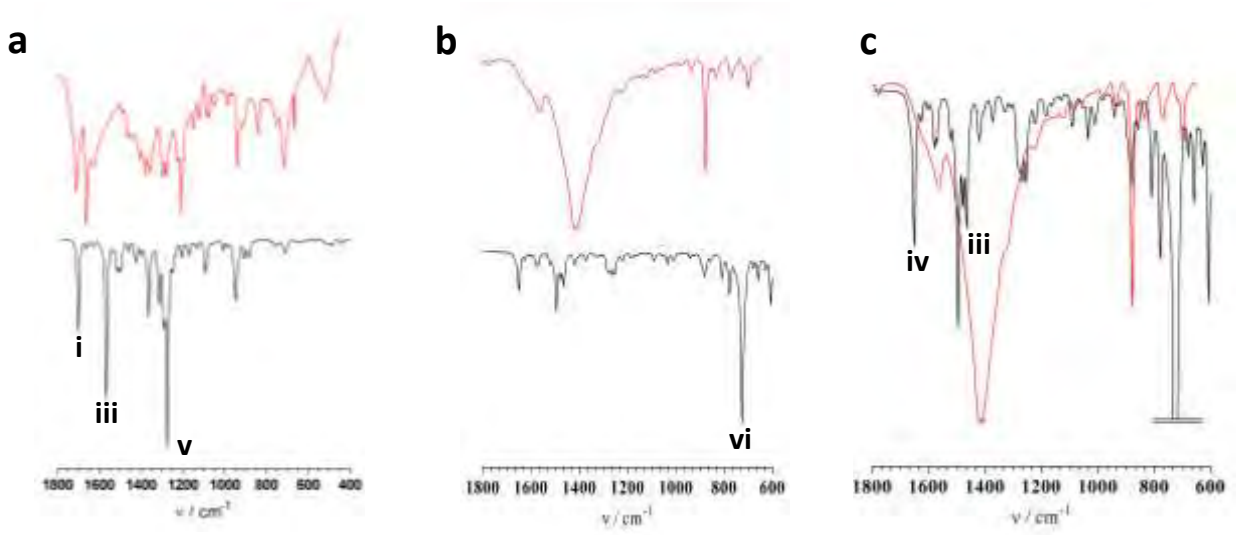


Figure 4.6. Computed (black) and experimental (red) IR spectra of dimeric (**a**) [μ-propionato ferrihaem]^{0} (S = $^{10}/_{2}$) and (**b**) [μ-oxo ferrihaem]$^{4-}$ (antiferromagnetically coupled). (**c**) Y-axis expansion of (b). See text for assignment of vibrational modes (i), (iii), (v) and (vi). No scaling factors have been applied.

The computed spectrum of dimeric [μ-oxo ferrihaem]$^{4-}$ is reported in Fig. 4.6b and is dominated by a single intense band at 724 cm^{-1}. This peak arises from the asymmetric υ(Fe-O) stretch, labelled (vi) in Fig. 4.6b, and is considerably lower in energy than the corresponding band assigned to this mode in the experimental spectrum at 880 cm^{-1}. The intense broad band at higher wavenumber in the experimental IR spectrum is not reproduced in the calculated spectrum. This feature is probably broadened owing to the presence of water in the solid which is hygroscopic. On the other hand, computed peaks near this region compare reasonably well with shoulders of the broad band (see Fig. 4.6c) and thus the υ(C=O) stretch of the propionate side chains and vinyl υ(C=C) stretches likely underlay the broad experimental band.

4.5 Analysis of Electronic Transitions of Ferrihaem Species

UV-visible spectroscopy is routinely used to determine association constants between ferrihaem and various molecules[183], including anti-malarial drugs,[163, 177] or in colourimetric assays[306-308] to identify compounds that inhibit haemozoin formation. In addition, it has proven very useful in solution-state speciation investigations owing to the marked spectroscopic differences observed between different ferrihaem species (see Section 3.3.2).[75, 76] Experimentally, assignments of UV-visible spectra of monomeric species have previously been made by analogy, using calculations on the porphine moiety of the porphyrin ring or on symmetric metalloporphyrins such as TPP or OEP and their complexes with iron(III) or zinc(II).[193, 196, 199, 206, 207, 309, 310] However, despite the extensive application of this technique to ferrihaem studies, spectra have never been directly computed for these molecules and only very recently have spectra been computed using TD-DFT for the related camphor-bound ferrihaem complex in cytochrome P450 as well as for zinc(II)protoporphyrin IX, gallium(III)protoporphyrin IX and its μ-propionato dimer.[209, 278] With the QM structures of ferrihaem species in hand, calculation of UV-visible spectra for ferrihaem species was possible using TD-DFT. Spectra were predicted for the purpose of identifying electronic transitions making up spectroscopic bands and comparing these with those assigned using simplified models. Furthermore, by calculating spectra from first-principles, computed spectra could be used as a fingerprint for ferrihaem species and show that spectroscopic differences were due to fundamental differences in structure.

4.5.1 Predicting Spectra Using TD-DFT.

As in the case of DFT structure optimisation, various functionals and basis sets can be used to compute UV-visible spectra. In order to determine the conditions which are likely to produce UV-visible spectra in good agreement with experimental spectra, TD-DFT was first performed using [H_2O-ferrihaem]$^{1+}$ as a test molecule. The LANL2DZ basis set was used to describe all atoms in combination with three levels of theory PBE0, CAM-B3LYP and LC-wPBE.[292, 311-314] All calculations were conducted using the implicit water solvent model, SMD, as it has previously been shown that calculation of UV-visible spectra in vacuum often produce poor agreement with experiment.[315-317] The PBE0 functional was specifically chosen owing to reports of good reproduction of UV-visible spectra being obtained for small

and medium-size organic radicals.[318] The remaining two levels of theory employed are long-range corrected functionals which were developed specifically to account for the inherent overestimation in energy of charge transfer electronic transitions by TD-DFT.[319, 320] Spectra computed using these three combinations are displayed in Fig. 4.7a – c. For comparison purposes, the UV-visible spectrum of [H_2O-ferrihaem]$^{1+}$ was also computed using a different approach to TD-DFT, namely the configuration interaction method, CI-Singles (CIS)[321] and is shown in Fig. 4.7d.

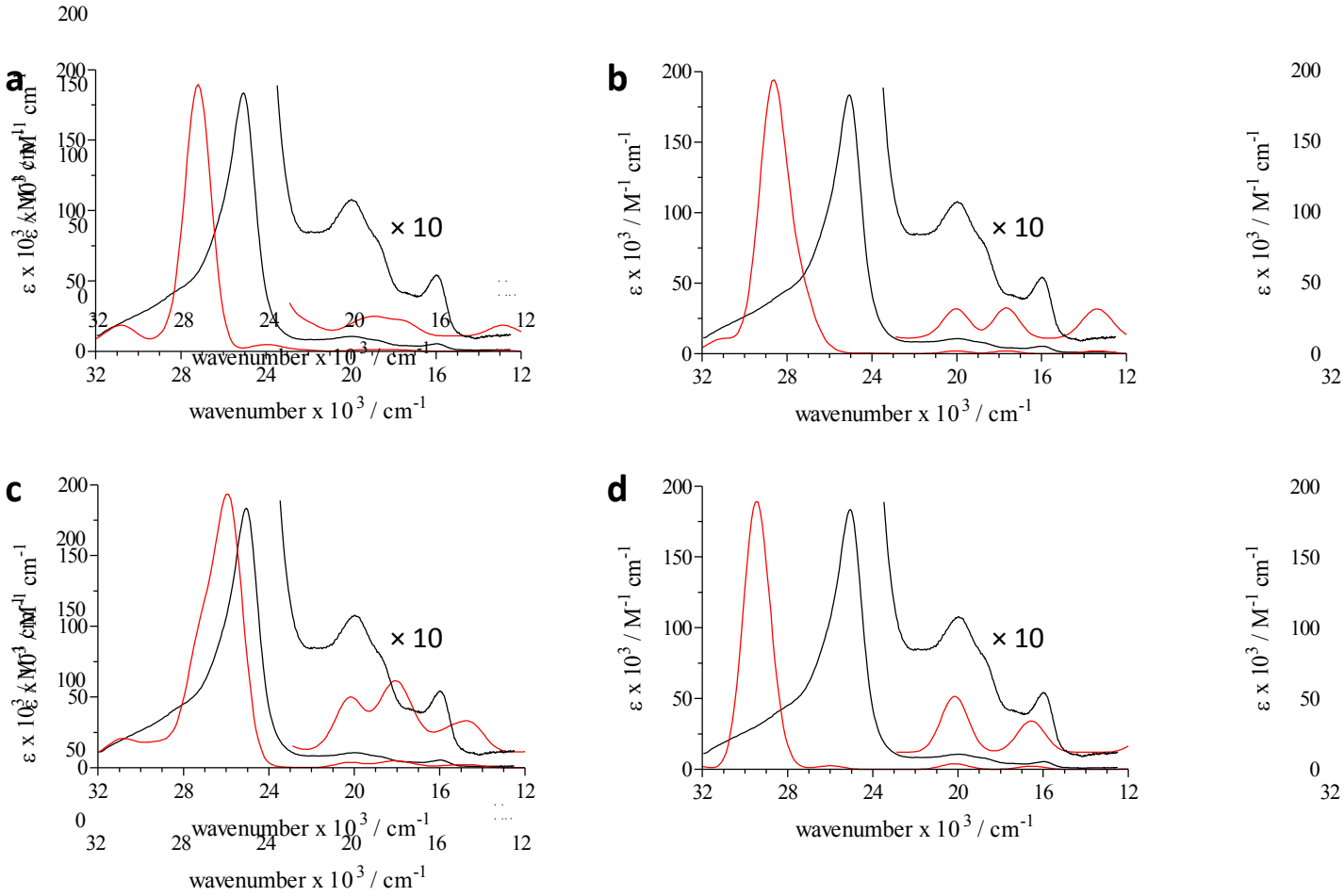


Figure 4.7. Experimental (black) and computed (red) spectra of [H_2O-ferrihaem]$^{1+}$ calculated using (**a**) TD-DFT with CAM-B3LYP functional and LANL2DZ basis set; (**b**) TD-DFT with LC-wPBE functional and LANL2DZ basis set; (**c**) TD-DFT with PBE0 functional and LANL2DZ basis set; and (**d**) the CIS method. Computed spectra were convoluted using Gaussian functions with a half-width of 700 cm^{-1} and for the sake of comparison, have been scaled in intensity by 0.45, 0.56, 0.67 and 0.22 (a to d).

The two long-range functionals specifically developed for more accurate charge-transfer descriptions by TD-DFT calculations produced UV-visible spectra in poor agreement with experiment. In the case of CAM-B3LYP (Fig. 4.7a), while the intense Soret band experimentally observed around 25 060 cm^{-1} was moderately well reproduced (occurring around 27 175 cm^{-1}), poor reproduction of the lower energy region of the spectrum was observed. Conversely, the spectrum computed using LC-wPBE (Fig. 4.7b) displayed better agreement in the low energy range and was comparatively worse at describing the Soret region. Similar findings were observed for the spectrum computed using the alternate method, CIS, although in this case the Soret peak was at much higher energy than experiment (approx. 29 410 cm^{-1}) and the low energy regions were better represented (Fig. 4.7d). By contrast, the UV-visible spectrum calculated using TD-DFT with PBE0 functional and LANL2DZ basis set was in good agreement with the corresponding experimental spectrum (Fig. 4.7c). The computed Soret band was similar in energy to experiment (approx. 26 180 cm^{-1}) and while the three peaks making up the lower energy envelope were somewhat shifted in energy as compared to the corresponding region of the experimental spectrum, the overall features were well reproduced.

Given the good agreement obtained using the PBE0 functional, the UV-visible spectrum was re-computed using the LANL2DZdp basis set which includes diffuse and polarisation functions on all non-metal atoms (see Fig. 4.8a). Indeed, the use of these additional functions better reproduced the UV-visible spectrum than when excluded in the form of the LANL2DZ basis set (see Fig. 4.7c). Further attempts to improve agreement were made by following a similar approach as that used by Paulat and Lenhert for Cl-iron(III)TPP where the overestimated Fe-Cl bond length was shortened to that observed experimentally.[207] In the case of H_2O-ferrihaem, the Fe-O bond length was shortened to the value obtained from the crystal structure of H_2O-iron(III)OEP (2.045 Å).[322] As seen in Fig. 4.8b, only minor differences were seen between the UV-visible spectrum computed using the DFT-calculated structure and that in which the axial bond length was reduced. In both cases, the calculated Soret peak almost overlapped the experimental counterpart (395 vs. 399 nm), while the putative Q-band was slightly overestimated (510 vs. 535 nm). Conversely, the band putatively assigned as the charge transfer peak was substantially overestimated in energy (581 vs. 626 nm).

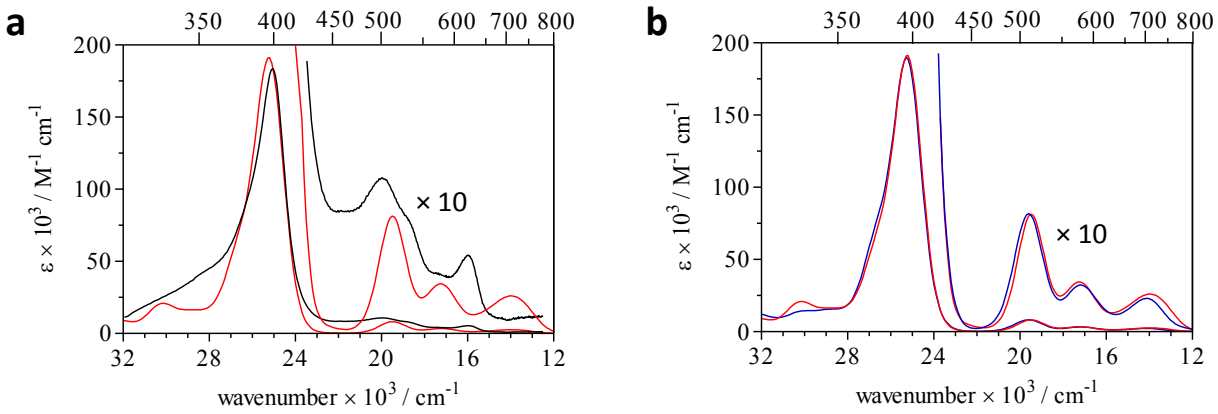


Figure 4.8. (**a**) Experimental (black) and computed (red) UV-visible spectra of $[H_2O\text{-ferrihaem}]^{1+}$. The calculated spectrum was generated using TD-DFT with PBE0 functional and LANL2DZdp basis set, and has been scaled by 0. 62 for comparison purposes. (**b**) Computed UV-visible spectra of $[H_2O\text{-ferrihaem}]^{1+}$ in which the axial bond is unmodified (2.169 Å, red) and shortened (2.045 Å, blue). Spectra were calculated using TD-DFT with PBE0 functional and LAN2DZdp basis set and convoluted using Gaussian functions with a 700 cm^{-1} half-width and intensities scaled by 0.62.

Since the best agreement of calculated UV-visible spectra of $[H_2O\text{-ferrihaem}]^{1+}$ with experiment was obtained using TD-DFT with PBE0 functional and LANL2DZdp basis set, the UV-visible spectra of the remaining ferrihaem species of interest were calculated using this method. These included $[HO\text{-ferrihaem}]^{2-}$ and $[\mu\text{-oxo ferrihaem}]^{4-}$. Owing to the insolubility of haemozoin, no experimental spectrum could be recorded for this species and thus the UV-visible spectrum for μ-propionato ferrihaem was not determined. In the case of $[\mu\text{-oxo ferrihaem}]^{4-}$, the calculated spectrum produced good agreement with that observed experimentally (see Fig. 4.9a). In particular, the Soret band with prominent high-energy shoulder was well reproduced in the computed spectrum. The lower energy envelope was somewhat overestimated in energy as compared to the corresponding feature in the experimental spectrum (535 vs. 585 nm).

By contrast, the computed UV-visible spectra of neither the high- (S = $^5/_2$) nor intermediate-spin (S = $^3/_2$) [HO-ferrihaem]$^{2-}$ species reproduced experiment well (see Fig. 4.9b and c). This is likely because of the spin-admixed character of this species which cannot be accurately described by the Gaussian09 software package. Therefore further discussion of computed UV-visible spectra is limited to [H_2O-ferrihaem]$^{1+}$ and [μ-oxo ferrihaem]$^{4-}$. To facilitate analysis, the spectrum was also calculated for porphinate in which no iron is present and thus can only partake in π→π* transitions.

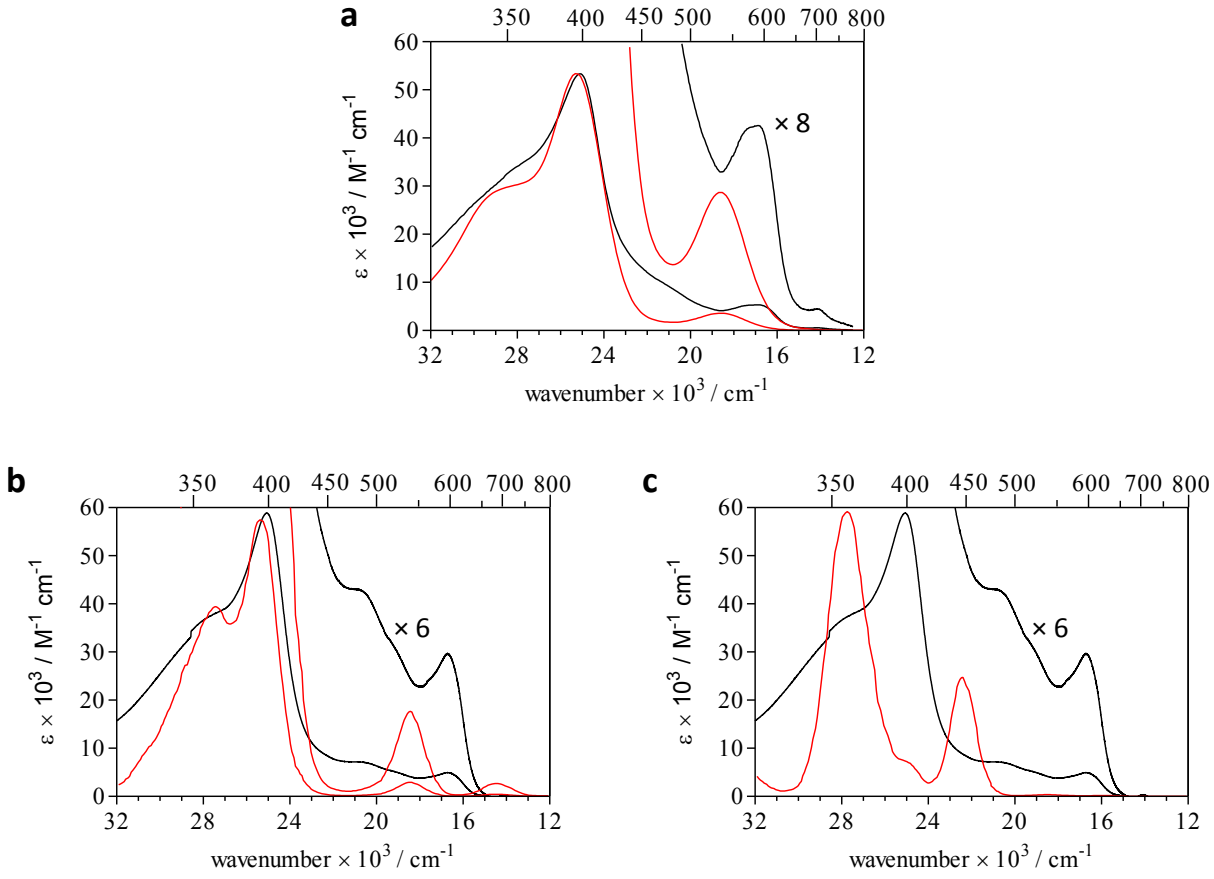


Figure 4.9. Experimental (black) and computed (red) UV-visible spectra of (**a**) [μ-oxo ferrihaem]$^{4-}$, spin paired; (**b**) [HO-ferrihaem]$^{2-}$, S = $^5/_2$; and (**c**) [HO-ferrihaem]$^{2-}$, S = $^3/_2$. The spectrum in (a) was calculated using TD-DFT with PBE0/LANL2DZ:LANL2DZdp while those in (b) and (c) were calculated using PBE0/LANL2DZ. Intensity scaling factors of 0.13, 0.30 and 0.28 have been used for computed spectra in (a) – (c) respectively.

4.5.2 Analysis of the computed spectrum of porphinate.

The calculated spectrum of porphinate is shown in Fig. 4.10a. Two prominent excitations occur at 17 900 (559 nm) and 25 070 cm^{-1} (399 nm), the latter gives rise to a more intense peak which is termed the Soret band, while the former gives rise to the Q-band. To investigate electronic transitions making up the two excitations in the calculated UV-visible spectrum, a detailed description of molecular orbitals (MOs) involved in these transitions was required. The MO diagram and selected MOs contributing to major electronic transitions calculated for porphinate is displayed in Fig. 4.11. Owing to the idealized D_{4h} symmetry of this molecule, MOs can be described by symmetry labels. Only two one-electron transitions account for more than 97% of each excitation and are labelled $\pi\rightarrow\pi^*(1)$ and $\pi\rightarrow\pi^*(2)$ in Fig 4.10b. Both transitions terminate in one of two degenerate $e_g(\pi^*)$ orbitals (<82> and <83>) but differ in the MOs from which they originate. In the case of $\pi\rightarrow\pi^*(1)$, the electron is excited from the highest occupied MO (HOMO), <81>, which has a_{2u} symmetry, while that of $\pi\rightarrow\pi^*(2)$ originates in HOMO−1, <80>, which has a_{1u} symmetry.

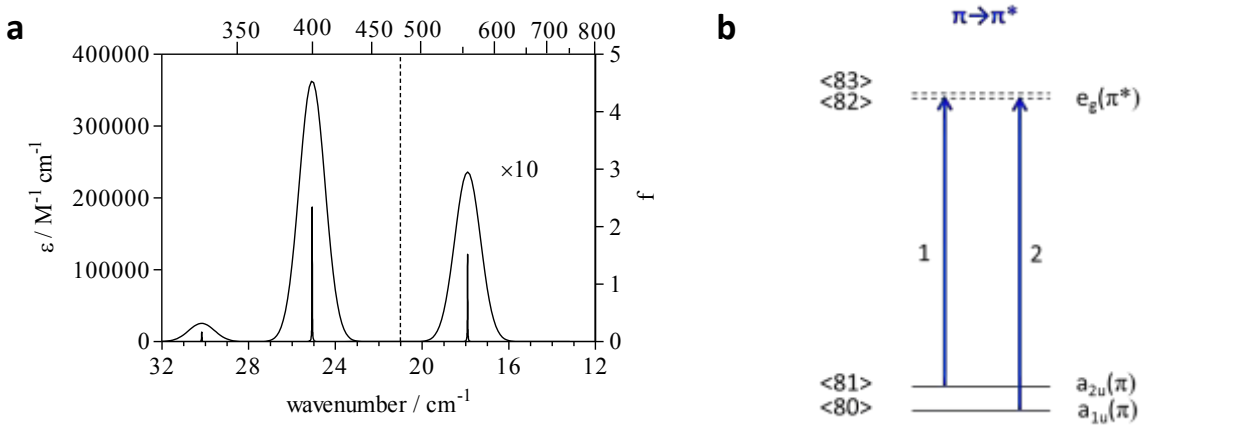


Figure 4.10. (**a**) Computed spectrum of porphinate calculated using TD-DFT and PBE0/LANL2DZ:LANL2DZdp. Vertical bars indicate underlying excitations and are plotted on the right axis in oscillator strength, *f*. (**b**) One-electron $\pi\rightarrow\pi^*$ transitions making up major contributions to calculated excitations at 17 900 and 25 070 cm^{-1} (see Fig. 4.11 for orbital numbering).

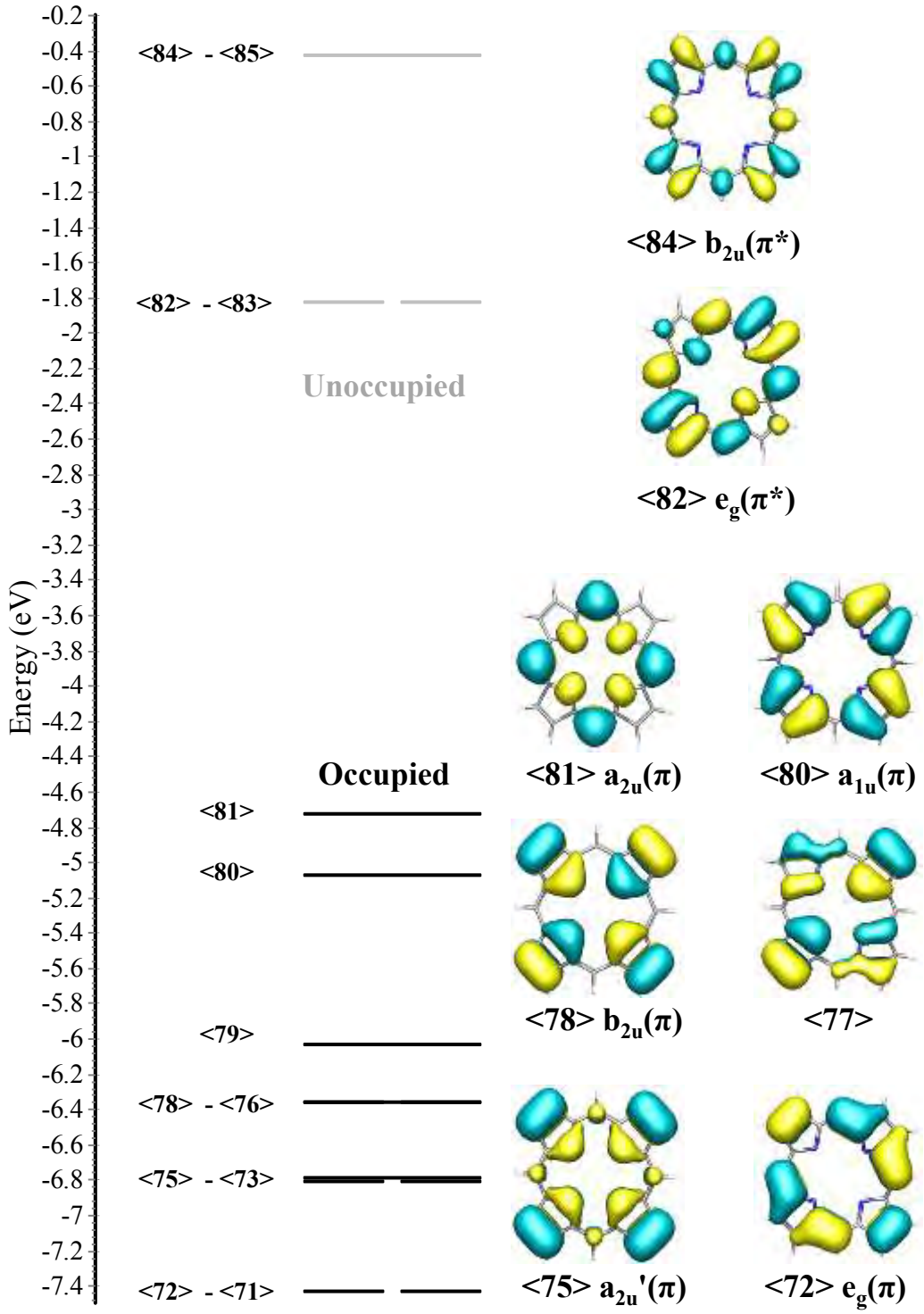


Figure 4.11. MO diagram of porphinate (left) and selected orbitals with symmetry labels (right).

The calculated excitation corresponding to the Q-band comprises 32% $\pi\rightarrow\pi^*(1)$ and 67% $\pi\rightarrow\pi^*(2)$ character, while that making up the Soret band has 66% $\pi\rightarrow\pi^*(1)$ and 31% $\pi\rightarrow\pi^*(2)$ character. Transitions comprising the Soret and Q-band identified by TD-DFT are in agreement with previous assignments made using the Gouterman four-orbital model (see Section 2.2.1).

4.5.3 Analysis of the computed UV-visible spectrum of H_2O-ferrihaem.

As in the case of porphinate, to investigate electronic transitions making up excitations giving rise to bands in the calculated UV-visible spectrum, a detailed description of MOs involved in these transitions was required. For paramagnetic species such as $[H_2O\text{-ferrihaem}]^{1+}$, both α- and β-spin orbitals contribute to $\pi\rightarrow\pi^*$ transitions, but only β-spin orbitals participate in to ligand to metal charge transfer (CT) transitions. Thus for the sake of brevity, only β-spin MOs will be discussed. The MO diagram and selected MOs contributing to major electronic transitions calculated for $[H_2O\text{-ferrihaem}]^{1+}$ is displayed in Fig. 4.12 and was determined for the structure in which the axial Fe-O bond length was shortened to 2.045 Å. Corresponding contributions to MOs by the iron atom, axial ligand, vinyl groups and propionic acid side chains are presented in Table 4.4.

All occupied MOs involved in one-electron transitions making up a significant proportion of calculated excitations are porphyrin π-orbitals (β<148> – β<158>). Two of these MOs, β<155> and β<156>, have substantial vinyl character but resemble the $a_{2u}'(\pi)$ and $b_{2u}(\pi)$ MOs of porphinate. A further two, β<149> and β<150>, have a marked contribution from the propionic acid side chains and have no similar porphinate MO counterpart. Interestingly, the $a_{1u}(\pi)$ and $a_{2u}(\pi)$ MOs (β<158> and β<157>) which correspond to HOMO and HOMO–1, respectively were inverted with respect to porphinate. The same observation was previously reported by Pinter *et al*. for Ga(III)protoporphyrin IX but not Zn(II)protoporphyrin IX.[278] The first two unoccupied MOs (β<159> and β<160>) are metal d_{xz}- and d_{zy}-orbitals, both of which contain significant porphyrin character and are termed d_π. The asymmetric substitution pattern of ferrihaem lifts the degeneracy of these MOs which in strict D_{4h} symmetry should be equivalent. The unoccupied MO β<161> contains the smallest porphyrin character and largest iron character, the latter in the form of the d_{xy}-orbital. The unoccupied MO with the next lowest energy (β<163>) corresponds to the d_z^2-orbital. Interestingly, this MO has very little axial ligand

character but a significant contribution is observed from the porphyrin, most notably on the nitrogen atoms. MOs β<164> and β<165> are porphyrin π*-orbitals which have negligible contributions from iron, ligand, vinyl or propionic acid side chain groups. These MOs correspond to the $e_g(\pi^*)$ orbitals of porphinate and despite the asymmetric substitution pattern of ferrihaem, remain almost degenerate. The unoccupied MO with largest $d_{x^2-y^2}$ iron character is observed at higher energy than the $e_g(\pi^*)$ MOs. This MO has a significant contribution from the porphyrin in the form of nitrogen σ-orbitals. Finally, the remaining MO making significant contributions to excitations was a porphyrin π*-orbital which corresponds to the $b_{2u}(\pi^*)$ MO in porphinate.

Table 4.4. Charge contributions of β-spin MOs of $[H_2O\text{-ferrihaem}]^{1+}$ calculated with PBE0/LANL2DZ:LANL2DZdp.[a]

Orbital		Label [b]	Energy (eV)	Contribution (%)				
				Fe	H_2O	Porphyrin [d]	Vinyl	Propionic [e]
β<166>		$b_{2u}(\pi^*)$	-1.209	0	0	94	5	1
β<165>		$d_{x^2-y^2}$	-1.972	22	0	77	0	0
β<164>		$e_g(\pi^*)$	-2.499	4	0	92	1	1
β<163>		$e_g(\pi^*)$	-2.520	5	1	91	2	1
β<162>		d_{z^2}	-3.031	39	4	55	2	0
β<161>		d_{xy}	-3.460	75	0	24	0	1
β<160>		d_{π}	-3.688	49	2	47	1	1
β<159>	LUMO	d_{π}	-3.897	44	0	53	1	2
β<158>	HOMO	$a_{1u}(\pi)$	-5.784	0	0	96	3	1
β<157>		$a_{2u}(\pi)$	-5.935	3	0	95	1	1
β<156>		$b_{2u}(\pi)$	-6.579	1	0	80	18	1
β<155>		$a_{2u}'(\pi)$	-6.807	3	0	84	11	3
β<154>		–[c]	-6.994	4	1	87	4	4
β<153>		–[c]	-7.139	4	0	89	1	6
β<152>		$e_g(\pi)$	-7.993	2	0	87	10	1
β<151>		$e_g(\pi)$	-8.031	1	0	90	7	2
β<150>		–[c]	-8.463	0	0	80	4	15
β<149>		–[c]	-8.481	0	0	52	0	48
β<148>		–[c]	-8.504	1	0	80	15	4

[a] All orbitals below the dotted horizontal line are occupied; [b] in the case of π-orbitals, symmetry labels refer to related orbitals in porphinate which has idealized D_{4h} symmetry. See Section 4.5.2 for corresponding porphinate MOs; [c] no corresponding porphinate orbital could be identified; [d] includes methyl substituents; [e] includes methylene and carboxylic acid groups.

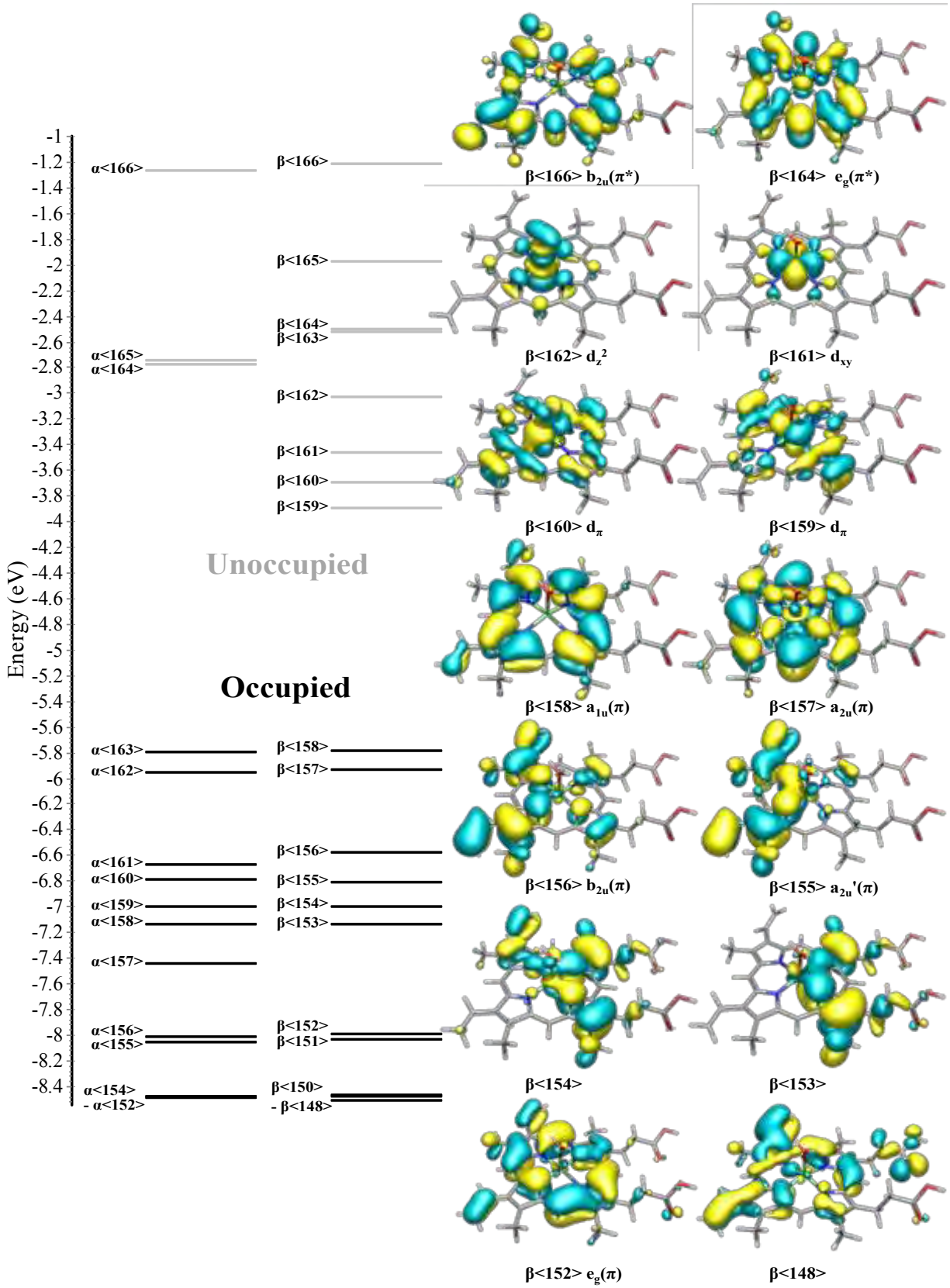


Figure 4.12. MO diagram of $[H_2O\text{-ferrihaem}]^{1+}$ (left) showing selected β-spin orbitals (right).

The calculated spectrum of $[H_2O\text{-ferrihaem}]^{1+}$ is shown in Fig. 4.13 along with its experimental counterpart. There are a limited number of one-electron transitions that contribute 5% or more to calculated excitations between 12 000 (800 nm) and 28 000 cm^{-1} (357 nm), and are represented diagrammatically in scheme 4.1. To reduce complexity, transitions originating from the same MOs and terminating in unoccupied MOs that are degenerate in D_{4h} symmetry are grouped together as a single transition. Such transitions include those terminating in the $e_g(\pi^*)$ (β<164> and β<165>) or d_π MOs (β<159> and β<160>). For example, the $\pi \rightarrow d_\pi$ charge transfer transitions β<156>→β<159> and β<156>→β<160> are given the same number and labelled as CT(1). Furthermore, all β-spin $\pi \rightarrow \pi^*$ transitions have α-spin counterparts and both are given the same numbering in scheme 4.1.

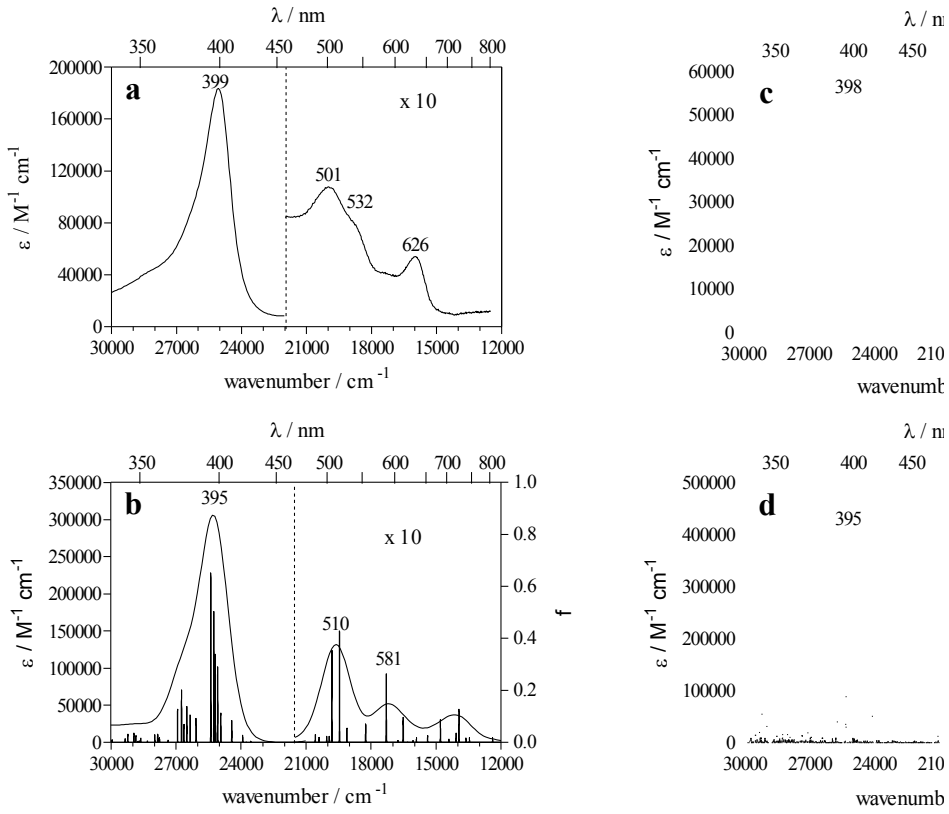
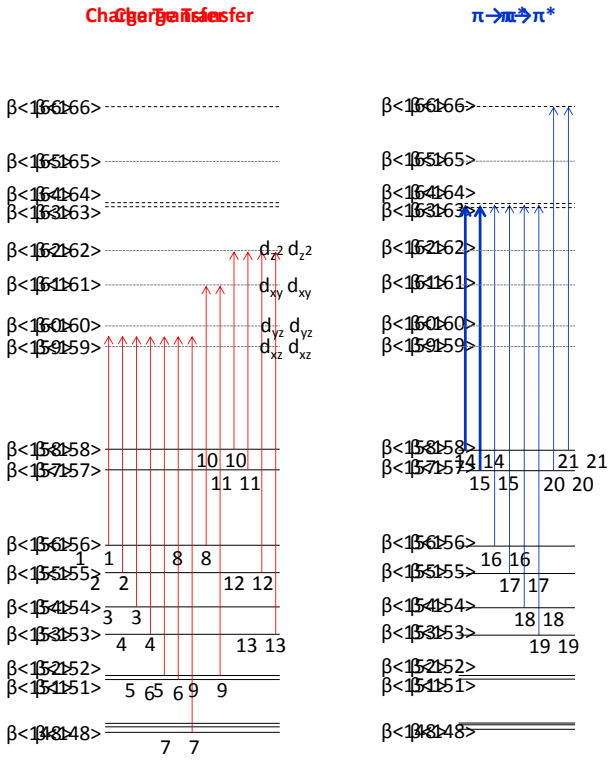


Figure 4.13. Experimental (**a**) and computed (**b**) UV-visible spectra of $[H_2O\text{-ferrihaem}]^{1+}$. The experimental spectrum was recorded in acidified aqueous acetone solution (5.64 M), pH 2. Excitations in the calculated spectrum are represent by vertical bars plotted on the right axis in oscillator strength, *f*, and have been convoluted using Gaussian functions with a half-width of 700 cm^{-1}.

Scheme 4.1



Energy cm^{-1} (nm)	f	Dominant Transitions
13933 (717.7)	0.0127	10(24%) + 1(21%) + **15(26%)** + **14(13%)** + (···)
14789 (676.2)	0.0087	3(49%) + 2(21%) + 1(14%) + (···)
16510 (605.7)	0.0096	**14(34%)** + **15(18%)** + 3(16%) + 1(10%) + 4(10%) + (···)
17283 (578.6)	0.0264	4(27%) + **15(24%)** + **14(21%)** + 2(13%) + (···)
18235 (548.4)	0.0070	4(39%) + 2(20%) + **14(16%)** + **15(12%)** + (···)
19102 (523.5)	0.0055	**14(22%)** + 3(18%) + 4(12%) + 11(11%) + 2(10%) + **15(10%)** + (···)
19444 (514.3)	0.0429	**14(46%)** + **15(35%)** + (···)
19798 (505.1)	0.0354	4(23%) + **15(28%)** + **14(22%)** + (···)
23912 (418.2)	0.0268	13(40%) + 18(13%) + 12(12%) + 5(10%) + (···)
24414 (409.6)	0.0846	5(40%) + (···)
24925 (401.2)	0.1130	6(64%) + (···)
25063 (399.0)	0.2906	17(20%) + 16(20%) + **15(14%)** + (···)
25170 (397.3)	0.3395	5(17%) + **15(16%)** + 17(15%) + **14(10%)** + (···)
25246 (396.1)	0.5026	16(29%) + 17(22%) + **15(14%)** + **14(11%)** + (···)
25381 (394.0)	0.6522	16(25%) + **15(21%)** + **14(16%)** + 17(10%) + (···)
26062 (383.7)	0.0920	5(48%) + 17(10%) + (···)
26337 (379.7)	0.1052	6(50%) + 21(11%) + (···)
26483 (377.6)	0.1391	16(45%) + 18(20%) + (···)
26624 (375.6)	0.0696	16(57%) + 18(10%) + (···)
26731 (374.1)	0.2018	19(65%) + 16(14%) + 18(12%) + (···)
26918 (371.5)	0.1279	19(48%) + 18(11%) + (···)
27360 (365.5)	0.0086	20(30%) + (···)
27755 (360.3)	0.0191	17(53%) + 20(12%) + (···)
27785 (359.9)	0.0079	9(66%) + (···)
27832 (359.3)	0.0314	21(27%) + 9(18%) + (···)
27972 (357.5)	0.0292	17(46%) + 18(15%) + (···)

There are 46 excitations over the range of interest in the calculated spectrum of [H_2O-ferrihaem]$^{1+}$ (12 000 – 28 000 cm^{-1}), all of which are made up of multiple configurationally mixed one-electron transitions. The compositions of the 26 most intense excitations are listed in the table in scheme 4.1. Nine are responsible for the major spectroscopic features observed over the range 15 000 – 25 500 cm^{-1} (392 – 667 nm). The Soret band around 25 316 cm^{-1} (395 nm) is comprised of four closely spaced excitations occurring at 25 381, 25 246, 25 170 and 25 063 cm^{-1} which are predominately made up of $\pi\rightarrow\pi^*$ transitions. As observed for porphinate, there are significant contributions by $a_{1u}(\pi)/a_{2u}(\pi)\rightarrow e_g(\pi^*)$ transitions to these excitations which are labelled $\pi\rightarrow\pi^*(14)$ and $\pi\rightarrow\pi^*(15)$ in scheme 4.1 and are shown in bold. As previously mentioned in Section 2.2, according to the Gouterman four-orbital model, these transitions have been suggested to account for the B-band. In contrast to porphinate, however, the transitions involved in the Gouterman model have a much smaller contribution to excitations making up the Soret band where only 14 to 37% is calculated for [H_2O-ferrihaem]$^{1+}$ as compared to 99% for porphinate. Furthermore, two additional transitions $\pi\rightarrow\pi^*(16)$ and $\pi\rightarrow\pi^*(17)$ make an overall larger contribution to this region. Both terminate in the same $e_g(\pi^*)$ MOs as the transitions implicated in the Gouterman four-orbital model but originate from the two lower MOs, HOMO–2 and HOMO–3, respectively. Interestingly, a small but significant charge transfer component, CT(5), is also computed as part of the excitation at 25 170 cm^{-1} (397 nm).

The spectroscopic feature centred on 19 608 cm^{-1} (510 nm) in the calculated spectrum of [H_2O-ferrihaem]$^{1+}$ is comprised of two intense excitations at 19 797 and 19 444 cm^{-1} (505 and 514 nm). The latter is dominated by $\pi\rightarrow\pi^*(14)$ and $\pi\rightarrow\pi^*(15)$ transitions which have a combined contribution of 81%. While these same transitions are also present in the former, they contribute markedly less to the overall excitation (50%) and instead the charge transfer transition, CT(4), makes a significant contribution (23%). In accordance with the Gouterman four-orbital model, the predominance of the $\pi\rightarrow\pi^*$ transitions identifies the feature in this region as the Q-band. This is similar to the lower-intensity peak calculated for porphinate at 17 900 cm^{-1} (559 cm^{-1}) although the contribution by $a_{1u}(\pi)/a_{2u}(\pi)\rightarrow e_g(\pi^*)$ transitions is more dominant in porphinate (97%). Finally, three prominent excitations at 18 235, 17 283 and 16 510 cm^{-1} underlie the band with peak maximum at 17 212 cm^{-1} (581 nm) and all consist of heavily admixed $\pi\rightarrow d_\pi$ charge transfer and $\pi\rightarrow\pi^*$ character, involving CT(1), CT(2), CT(3), CT(4), $\pi\rightarrow\pi^*(14)$ and $\pi\rightarrow\pi^*(15)$ transitions.

In an attempt to assign computed excitations to underlying bands in the experimental UV-visible spectrum of H_2O-ferrihaem, deconvolution of the spectroscopic envelope was necessary. Fitting Gaussian functions to the spectrum cannot produce a unique fit owing to the inherent broad spectroscopic features. By contrast, the MCD spectrum of H_2O-ferrihaem has better resolved features and can be used in combination with the UV-visible spectrum to identify component bands more reliably through simultaneous deconvolution. This is possible by employing the rigid shift approximation which assumes the band shape of an underlying excitation is not affected by the magnetic field present in MCD.[295, 296] Previous studies have successfully applied this approach to other metalloporphyrins and related molecules.[199, 205, 207] The UV-visible and MCD spectra of H_2O-ferrihaem were deconvoluted using the minimum number of Gaussian functions (eighteen) which gave the most parsimonious fit to spectroscopic features in both spectra (see Fig. 4.14 and Table 4.5). Particular attention was paid to producing good fits in the areas corresponding to numbered Gaussian functions in Fig. 4.14. On the other hand, the featureless regions represented by grey Gaussians are less certain owing to the lack of resolution in these areas.

Table 4.5. Band fit parameters of H_2O-ferrihaem.

Band Number [a]	E_0 (cm^{-1}) [b]	σ (cm^{-1}) [c]	ΔE (cm^{-1}) [d]	$I_{UV\text{-}vis}$ ($M^{-1}\ cm^{-1}$) [e]	I_{MCD} ($M^{-1}\ cm^{-1}\ T^{-1}$) [e]
-	30608	1015	1690	18150	2.49
-	29240	1001	1666	23371	-1.59
-	27921	1001	1666	36172	-7.30
1	26667	861	1434	47841	2.24
2	25728	764	1272	81697	8.77
3	25095	571	950	92362	29.06
4	24810	540	900	42763	-29.07
5	24447	683	1137	34973	-11.90
-	23645	1001	1666	14122	0.87
-	21854	1001	1667	7624	4.48
6	20564	780	1299	6997	4.70
7	19636	752	1252	9254	-1.84
8	18865	446	743	4071	5.13
9	18356	429	714	4000	-6.72
10	17418	802	1336	3661	-1.18
11	16070	660	1099	4539	2.71
12	15736	375	625	1049	-8.04
-	12737	1000	1665	304	-0.22

[a] See Fig. 4.14; [b] peak position; [c] bandwidth; [d] full width at half-maximum (see Section 4.2.4); [e] intensity.

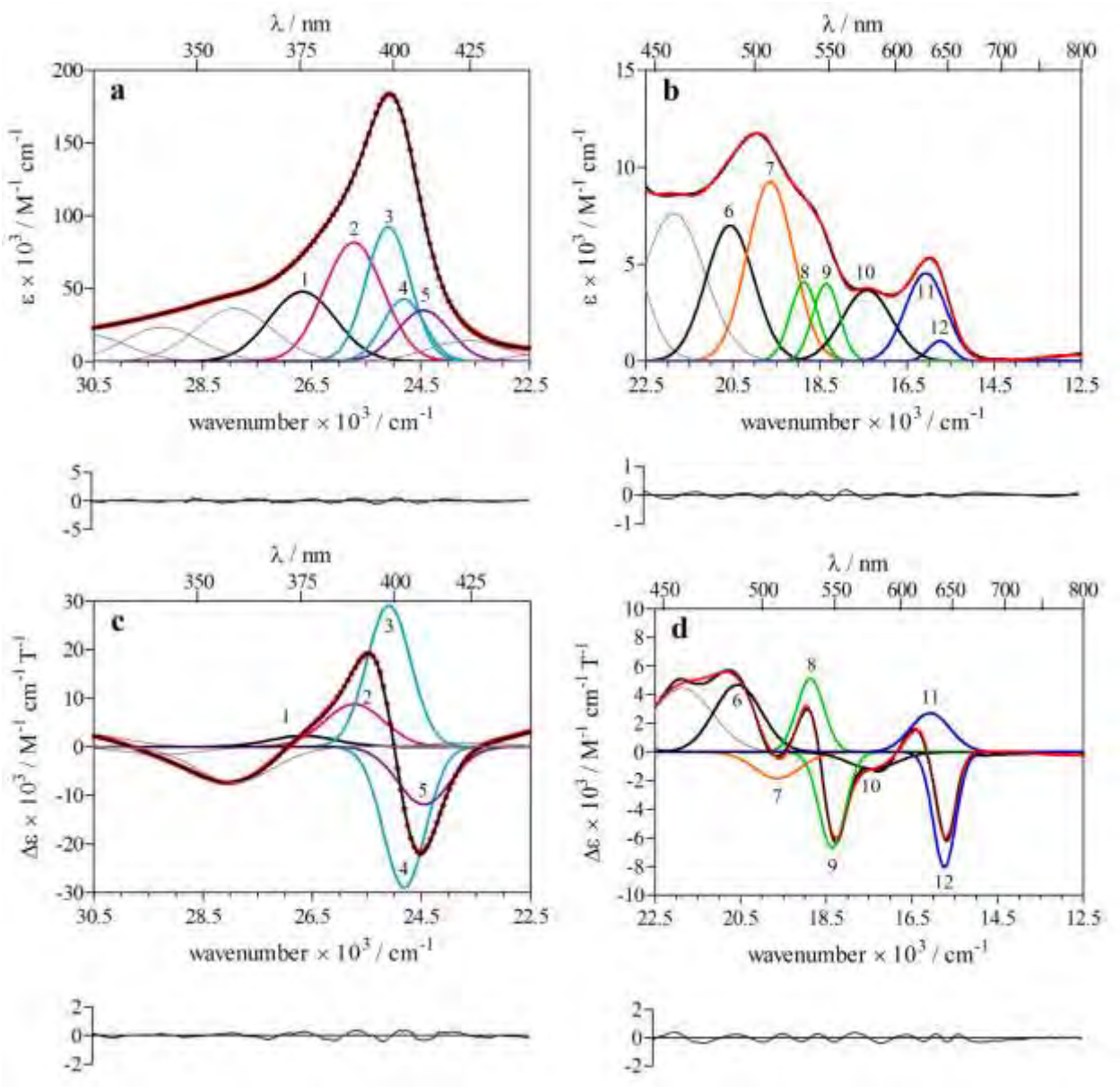


Figure 4.14. Deconvoluted UV-visible (**a** and **b**) and MCD (**c** and **d**) spectra of H_2O-ferrihaem at room temperature. Residuals are shown below each spectrum. Red dots represent the observed spectrum and black lines the sum of fitted Gaussian functions. Band numbers are referred to in the text and Table 4.5. Solvent conditions are the same as described in Section 3.2.6.

The lowest number of Gaussian functions required to fit the Soret region coincided with the four strong excitations in the computed UV-visible spectrum. On this basis, bands 2 – 5 most likely consist of $\pi\rightarrow\pi^*(14)$, $\pi\rightarrow\pi^*(15)$, $\pi\rightarrow\pi^*(16)$ and $\pi\rightarrow\pi^*(17)$ transitions with a minor CT(5) component. Furthermore, the two sets of bands of alternating sign (bands 2 and 5, and 3 and 4) in the MCD spectrum form pseudo-A terms which are expected for $\pi\rightarrow\pi^*$ transitions (see Section 2.3.3). A similar pseudo-A term is observed in the MCD spectrum around 18 500 cm^{-1} (540 nm) and is indicative of the Q-band. These bands can be assigned to computed excitations at 19 444 and 19798 cm^{-1}, which consist of $\pi\rightarrow\pi^*(14)$ and $\pi\rightarrow\pi^*(15)$ transitions with a small charge transfer component, CT(4). Band 7 likely corresponds to the Q_v-band given that the energy difference from band 9 (1280 cm^{-1}) is similar to that reported for ferrihaem in cytochrome $P450_{cam}$ (*ca.* 1300 cm^{-1}).[209] The positive MCD Q_v-component presumably underlies band 6 and cannot be resolved. Since the Q_v-band is a vibronic excitation, no such feature is expected in the computed spectrum and indeed no excitation in this region was observed. The three intense excitations computed at 16 510, 17 283 and 18 235 cm^{-1} in the calculated spectrum may correspond to bands 10 – 12, although assignment of these is less certain. If this tentative assignment is indeed correct, these bands consist of highly admixed $\pi\rightarrow\pi^*$ and $\pi\rightarrow d_\pi$ transitions. Finally, the region between the Soret and Q-bands (21 500 – 23 500 cm^{-1}) is poorly resolved but based on computed excitations, is likely to consist of largely charge transfer transitions. A summary of the dominant spectroscopic features in the experimental and computed UV-visible spectra is presented in Table 4.6.

The putative assignments of the Soret band in the UV-visible spectrum of H_2O-ferrihaem are largely in agreement with those previously reported by Makinen and Churg for the related iron(III)porphyrin, H_2O-metMb (see Section 2.2.2).[196] These authors found that $\pi\rightarrow\pi^*$ transitions dominated this region and implicated transitions involved in the Gouterman four-orbital model, $a_{1u}/a_{2u}(\pi)\rightarrow e_g(\pi^*)$, as the dominant contributors. While the same transitions are indeed observed in the computed spectrum of $[H_2O\text{-ferrihaem}]^{1+}$, a larger proportion is calculated from $a_{2u}'/b_{2u}(\pi)\rightarrow e_g(\pi^*)$ transitions. Interestingly, the latter transitions originate in MOs having substantial delocalisation onto the vinyl substituents and thus may account for the experimental differences in peak maxima of H_2O-ferrihaem (399 nm) as compared to related H_2O-iron(III)mesoporphyrin (390 nm) and H_2O-iron(III)deuteroporphyrin (386 nm), where the former porphyrin has ethyl groups in place of vinyl substituents and the latter hydrogen atoms.[163]

Table 4.6. Component band positions obtained from deconvolution of UV-visible absorbance and MCD spectra of H_2O-ferrihaem. Prominent calculated excitations obtained from TD-DFT calculations (PBE0/LANL2DZ:LANL2DZdp) making up the major spectroscopic envelope features are shown for comparison. [a, b]

Experimental		Computed	
energy range: 26 500 – 24 000 cm^{-1} (Soret)		*energy range*: 26 500 – 24 000 cm^{-1} (Soret)	
band 2	25 728 (389)	25 381 (394)	$\pi\rightarrow\pi^*(16)$, **$\pi\rightarrow\pi^*(15)$**, **$\pi\rightarrow\pi^*(14)$**, $\pi\rightarrow\pi^*(17)$
band 3	25 095 (398)	25 246 (396)	$\pi\rightarrow\pi^*(16)$, $\pi\rightarrow\pi^*(17)$, **$\pi\rightarrow\pi^*(15)$**, **$\pi\rightarrow\pi^*(14)$**
band 4	24 810 (403)	25 170 (397)	CT(5), **$\pi\rightarrow\pi^*(15)$**, $\pi\rightarrow\pi^*(17)$, **$\pi\rightarrow\pi^*(14)$**
band 5	24 447 (409)	25 063 (399)	$\pi\rightarrow\pi^*(17)$, $\pi\rightarrow\pi^*(16)$, **$\pi\rightarrow\pi^*(15)$**
energy range: 20 000 – 18 000 cm^{-1} (Q/Q_V)		*energy range*: 21 000 – 19 000 cm^{-1} (Q)	
band 7	19 636 (509)[c]	19 798 (501)	CT(4), **$\pi\rightarrow\pi^*(15)$**, **$\pi\rightarrow\pi^*(14)$**
band 8	18 865 (530)	19 444 (514)	**$\pi\rightarrow\pi^*(14)$**, **$\pi\rightarrow\pi^*(15)$**
band 9	18 356 (545)		
energy range: 18 000 – 15 000 cm^{-1}		*energy range*: 18 500 – 16 000 cm^{-1}	
band 10	17 418 (574)	18 235 (548)	CT(4), CT(2), **$\pi\rightarrow\pi^*(14)$**, **$\pi\rightarrow\pi^*(15)$**
band 11	16 070 (622)	17 283 (579)	CT(4), **$\pi\rightarrow\pi^*(15)$**, **$\pi\rightarrow\pi^*(14)$**, CT(2)
band 12	15 736 (635)	16 510 (606)	**$\pi\rightarrow\pi^*(14)$**, **$\pi\rightarrow\pi^*(15)$**, CT(3), CT(1), CT(4)

[a] Transitions shown in bold have counterparts in porphinate that correspond to those involved in Gouterman's four-orbital model; [b] band positions and excitations in wavenumbers (cm^{-1}) with wavelength in parentheses (nm); [c] Probable Q_V-band.

While transitions producing the Soret band are well established, those governing absorbance envelopes at lower energy are less definitive. Assignment of spectroscopic features of H_2O-metMb by Makinen and Churg in this region included the Q-band at 18 350 cm^{-1} (545 nm) and two charge transfer bands at 17 200 and 15 800 cm^{-1} (581 and 633 nm).[196] Pinter *et al.*, however, suggested the Q-band was the feature at 16 722 cm^{-1} (598 nm) in the spectrum of HO-ferrihaem based on the observed A-term feature in the experimental MCD spectrum and owing to similar positions of this band at 17 301 and 17 153 cm^{-1} (578 and 583 nm) in Ga(III)protoporphyrin IX and Zn(II)protoporphyrin IX, respectively.[278] Assignments made to the spectrum of H_2O-ferrihaem using TD-DFT in this study tend to support those proposed by Makinen and Churg for H_2O-metMb with the Q-band assigned to the peak at 18 797 cm^{-1} (532 nm). In addition, this peak is similar in position to the feature in the spectrum of ferrihaem in cytochrome $P450_{cam}$ assigned as the Q-band by Galinato *et al.*[209] On the other hand, assignment by Makinen and Churg of charge transfer bands as consisting of exclusively $\pi\rightarrow d_\pi$ transitions does not agree with the calculated excitations tentatively assigned to these bands for H_2O-ferrihaem. In this case, excitations in this region consist of heavily admixed $\pi\rightarrow d_\pi$ and $\pi\rightarrow\pi^*$ transitions.

4.5.4 Analysis of the computed UV-visible spectrum of μ-oxo ferrihaem.

The analysis of MOs and the computed UV-visible spectrum of μ-oxo ferrihaem are vastly more complex than that of H_2O-ferrihaem. The fact that there is almost double the number of atoms produces a correspondingly larger number of MOs, which although similar in energy are not degenerate since the structure is not centrosymmetric. Consequently a larger number of transitions are expected with the added complication that transitions from an occupied MO localised on one porphyrin to an unoccupied MO on the other, are allowed. Lastly, owing to the antiferromagnetic coupling between iron centres, charge transfer transitions to both unfilled α- and β-spin MOs are also possible for μ-oxo ferrihaem.

In the interest of brevity, the same approach as used to describe $[H_2O\text{-ferrihaem}]^{1+}$ MOs will be adopted for $[\mu\text{-oxo ferrihaem}]^{4-}$, namely only β-spin orbitals will be discussed. It should be reiterated, however, that every β-spin MO on one iron centre has an α-spin counterpart on the other. The MO diagram depicting β-spin orbitals associated with major transitions in the calculated spectrum of $[\mu\text{-oxo ferrihaem}]^{4-}$ is shown in Fig. 4.15. Table 4.7 lists the contribution of iron, oxide, porphyrin, vinyl and propionate side chain groups to selected β-spin MOs. As observed for $[H_2O\text{-ferrihaem}]^{1+}$, all occupied MOs in the range of interest are π-orbitals. Five of the thirteen occupied MOs (β<308>, β<309>, β<314> – β<316>) exhibit extensive delocalisation (> 30%) over both porphyrins, while the remaining eight MOs are predominately localised on a single porphyrin. Ordering of occupied MOs is similar to that observed in $[H_2O\text{-ferrihaem}]^{1+}$, particularly with regard to MOs with a_{1u} and a_{2u} symmetry. In the case of $[\mu\text{-oxo ferrihaem}]^{4-}$, however, each MO is duplicated. Delocalisation over two porphyrins was also observed for unoccupied π* MOs β<319> and β<321>. Interestingly, two π* MOs, β<318> and β<321>, have significant contributions from iron $3p_x$- and $3p_y$-orbitals resulting from the large out of plane doming of the iron atom from the porphyrin ring (0.56 Å) which provides better overlap with porphyrin π-orbitals. Similar overlap is not observed in $[H_2O\text{-ferrihaem}]^{1+}$ as the iron atom resides more in the plane of the porphyrin (0.27 Å) and thus is situated more in the nodal plane of the porphyrin π-system. Interestingly, there is a reordering of $[\mu\text{-oxo ferrihaem}]^{4-}$ unoccupied MOs as compared to $[H_2O\text{-ferrihaem}]^{1+}$ that stems from the increased energy of metal d-orbitals rather than a change in energy of unfilled π* orbitals. The reordering places the d_π- (β<322> and β<323>) and d_{z^2}-orbitals (β<324>) above the porphyrin $e_g(\pi^*)$ MOs (β<318> – β<321>) while the LUMO is the d_{xy}-orbital (β<317>). A notable difference in the d_{z^2}-orbital is that it includes a greater proportion of axial ligand character than the

corresponding MO in [H_2O-ferrihaem]$^{1+}$. Furthermore, all MOs involving significant metal d-orbital character have smaller overall contributions than calculated for the same MOs in [H_2O-ferrihaem]$^{1+}$.

Table 4.7 Charge contributions of β-spin MOs of [μ-oxo ferrihaem]$^{4-}$ calculated with PBE0/LANL2DZ:LANL2DZdp. The labels 1 and 2 refer to the first and second Fe(III)porphyrin in the dimer, respectively. [a]

Orbital	Label [b]	Energy (eV)	Composition (%)								
			Fe		Porphyrin [e]		Vinyl		Propionate [f]		O^{2-}
			1	2	1	2	1	2	1	2	
β<327>	$b_{2u}(\pi^*)$	-0.943	9	1	65	14	9	0	2	0	0
β<326>	$b_{2u}(\pi^*)$	-1.014	0	4	16	69	0	8	0	2	0
β<325>	$d_{x^2-y^2}$	-1.559	16	0	75	7	1	0	1	0	0
β<324>	d_{z^2}	-1.735	31	8	36	16	0	0	1	0	7
β<323>	d_{π}	-1.938	20	2	51	24	1	0	0	0	2
β<322>	d_{π}	-1.951	27	1	58	9	1	0	1	0	2
β<321>	$e_g(\pi^*)$	-2.392	3	17 [d]	34	45	0	1	0	0	0
β<320>	$e_g(\pi^*)$	-2.427	5	12 [d]	21	59	0	1	0	1	0
β<319>	$e_g(\pi^*)$	-2.505	17 [d]	8	35	38	1	1	0	0	1
β<318>	$e_g(\pi^*)$	-2.529	24 [d]	3	48	21	1	0	1	0	1
β<317> LUMO	d_{xy}	-2.561	35	1	34	27	1	0	1	1	0
β<316> HOMO	$a_{1u}(\pi)$	-5.409	0	0	43	54	1	1	0	1	0
β<315>	$a_{1u}(\pi)$	-5.463	0	0	45	51	1	1	1	1	0
β<314>	$a_{2u}(\pi)$	-5.549	2	5	32	58	0	0	1	1	0
β<313>	$a_{2u}(\pi)$	-5.637	5	2	65	26	1	0	1	0	0
β<312>	– [c]	-6.312	1	1	67	13	16	0	1	0	1
β<311>	– [c]	-6.339	0	2	14	68	0	14	0	1	0
β<310>	– [c]	-6.381	1	0	60	24	8	2	3	0	2
β<309>	– [c]	-6.431	3	1	55	27	1	3	8	0	1
β<308>	– [c]	-6.435	4	5	30	46	1	9	3	1	1
β<307>	– [c]	-6.506	1	4	6	74	0	1	0	14	0
β<306>	– [c]	-6.570	3	5	18	59	0	1	1	14	1
β<305>	$a_{2u}'(\pi)$	-6.674	3	4	55	20	4	0	12	1	0

[a] All orbitals below the dotted horizontal line are filled; [b] in the case of π-orbitals, symmetry labels refer to related orbitals in monomeric porphinate which has idealized D_{4h} symmetry; [c] no corresponding porphinate orbital can be identified; [d] Fe p_x and p_y; [e] includes methyl substituents; [f] includes methylene and carboxylate groups.

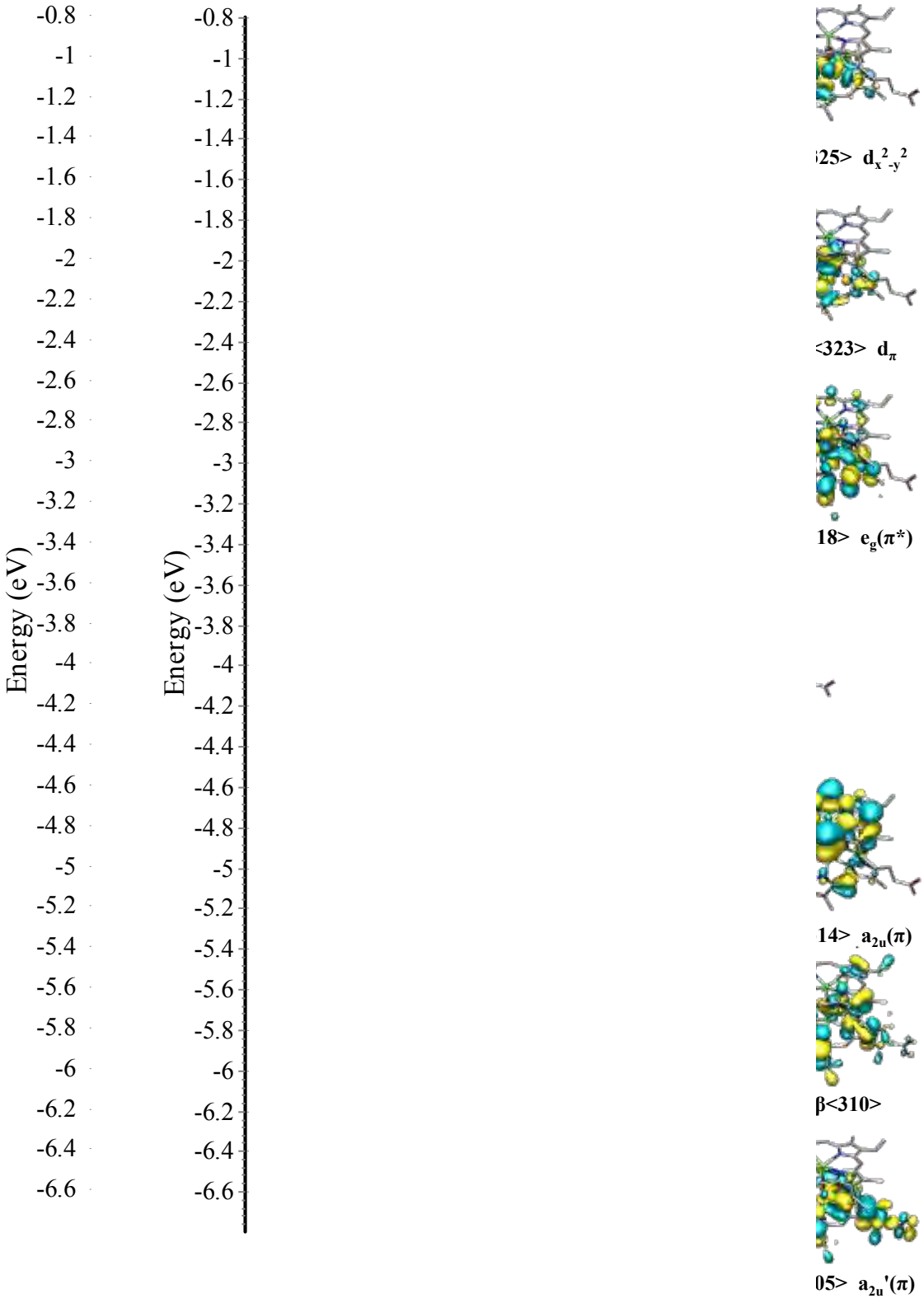


Figure 4.15. β-Spin MO diagram of [μ-oxo ferrihaem]$^{4-}$ (left) showing selected MOs (right).

The computed UV-visible spectrum of [μ-oxo ferrihaem]$^{4-}$ modelled using broken symmetry to describe antiferromagnetic coupling between iron centres is displayed in Fig 4.16 along with the experimental spectrum for comparison. The characteristic spectroscopic features used to identify this species, namely the blunt Soret peak with prominent high energy shoulder, absorbance envelope at lower energy and broadened region between the two, is well reproduced by the calculated spectrum. The Soret band peak in the computed spectrum at 25 316 cm^{-1} (395 nm) is slightly blue-shifted as compared the corresponding experimental peak at 21 126 cm^{-1} (398 nm), while the lower energy spectroscopic feature is significantly overestimated (535 vs. 593 nm).

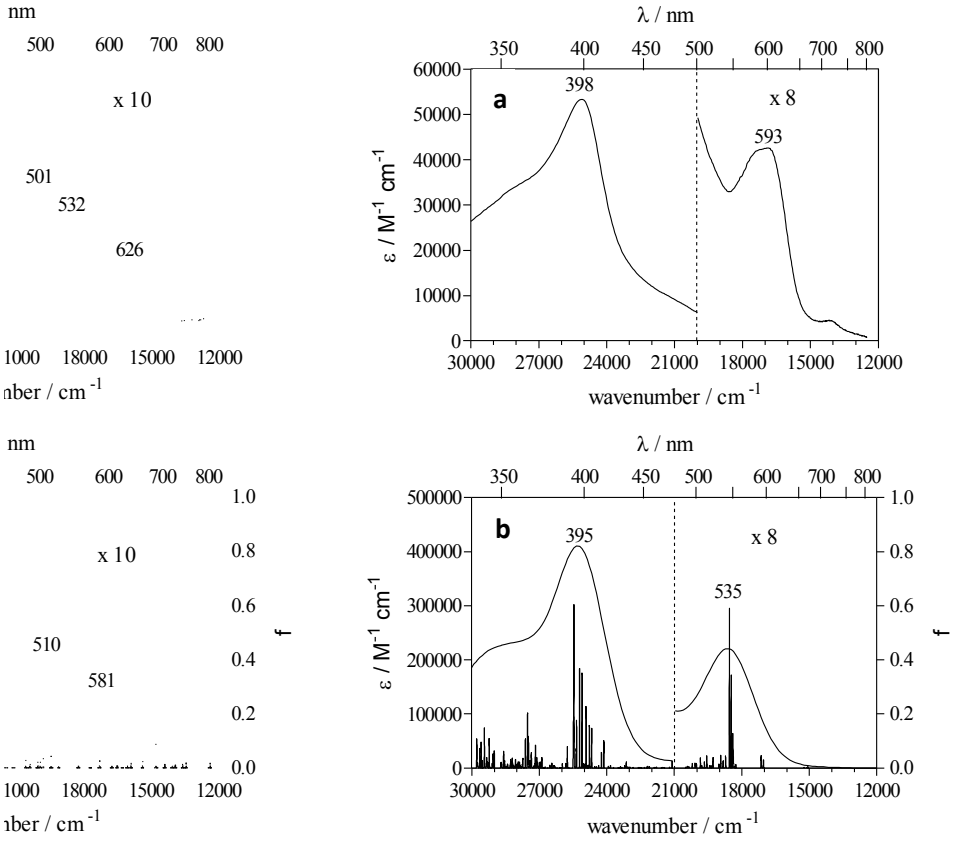
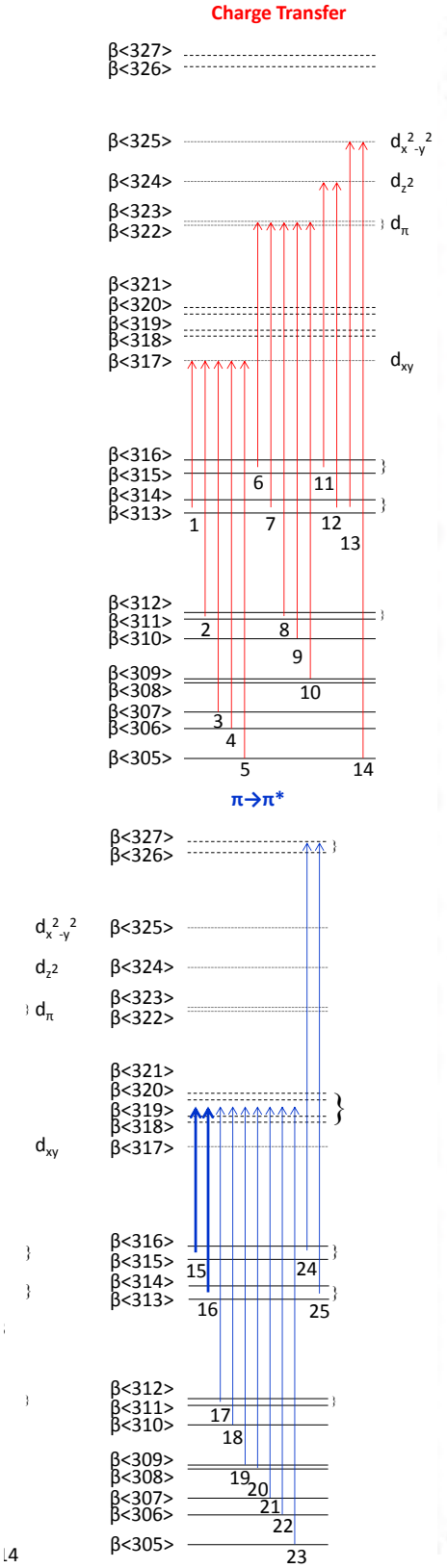


Figure 4.16. Experimental (**a**) and computed (**b**) UV-visible spectra of [μ-oxo ferrihaem]$^{4-}$. The experimental spectrum was recorded in aqueous DMSO solution (5.64 M), pH 10. Excitations in the calculated spectrum are represent by vertical bars plotted on the right axis in oscillator strength, *f*, and have been convoluted using Gaussian functions with a half-width of 700 cm^{-1}.

As expected, the larger number of MOs of [μ-oxo ferrihaem]$^{4-}$ produces many more one-electron transitions in the computed spectrum. Transitions contributing 5% or more to calculated excitations for [μ-oxo ferrihaem]$^{4-}$ are labelled and presented in Scheme 4.2. To facilitate analysis, MOs associated with major transitions have been grouped together in a similar approach as used to describe [H_2O-ferrihaem]$^{1+}$. In this case, however, [μ-oxo ferrihaem]$^{4-}$ MOs that are degenerate or nearly degenerate in [μ-oxo ferriporphine]^{0} are treated as equivalent. These include grouping of the two porphyrin β-spin $a_{1u}(\pi)$, two $a_{2u}(\pi)$, four $e_g(\pi^*)$ and two $b_{2u}(\pi^*)$ MOs as well as the two occupied MOs β<311> and β<312> which have no porphinate counterpart. In the case of the four $e_g(\pi^*)$ MOs, in- and out-of-phase mixing of the iron $3p_x$- and $3p_y$-orbitals lifts the degeneracy even in [μ-oxo ferriporphine]^{0} but can still be thought of as equivalent. Furthermore, the unoccupied metal d_π-orbitals have also been grouped as one. As for [H_2O-ferrihaem]$^{1+}$, $\pi\rightarrow\pi^*$ transitions labelled in Scheme 4.2 describe transitions involving both α- and β-spin MOs. In contrast to [H_2O-ferrihaem]$^{1+}$ where only β-spin charge transfer transitions were possible, all ligand to metal d-orbital transitions labelled in Scheme 4.2 for [μ-oxo ferrihaem]$^{4-}$ represent both α- and β-spin MOs.

In the computed spectrum of [μ-oxo ferrihaem]$^{4-}$ over the range 12 500 – 28 000 cm^{-1} there are 147 excitations, all of which consist of varying degrees of admixture of multiple one-electron transitions. The Soret region is comprised of eight relatively strong excitations at 24 911 (401 nm), 25 088 (399 nm), 25 197 (397 nm), 25 331 (395 nm), 25 352 (394 nm), 25 448 (393 nm) 25 475 (392 nm) and 25 739 cm^{-1} (388 nm). These excitations are dominated by $\pi\rightarrow\pi^*$ transitions, all of which terminate in $e_g(\pi^*)$ MOs. While transitions implicated in Gouterman's four-orbital model, labelled in bold as $\pi\rightarrow\pi^*$(15) and $\pi\rightarrow\pi^*$(16) in Scheme 4.2, do contribute to the calculated excitations in this region, they are not dominant. Instead, transitions from lower-lying MOs (with no counterpart in porphinate) have a larger overall contribution and, similar to that observed for [H_2O-ferrihaem]$^{1+}$, involve MOs with substantial delocalisation onto the vinyl groups. In addition, there is a larger proportion of charge transfer transitions calculated for [μ-oxo ferrihaem]$^{4-}$ than for [H_2O-ferrihaem]$^{1+}$. These involve $\pi\rightarrow d_\pi$ transitions, CT(7) and CT(8) as well as the $\pi\rightarrow d_z^2$ transition, CT(12). The high energy shoulder is comprised of a large number of excitations consisting of various admixed charge transfer and $\pi\rightarrow\pi^*$ transitions and thus assignment of this region is too complex to discuss in further detail. The broad low-energy shoulder is also comprised of many excitations, although the majority involve $\pi\rightarrow d_z^2$ character.

Scheme 4.2



Energy (nm)	f	Dominant Transitions
17129 (583.8)	0.0058	**16(37%) + 15(31%)** +7(13%) + 1(6%) + (···)
18362 (544.6)	0.0052	1(62%) + **16(9%) + 15(7%)** + 7(5%) + (···)
18430 (543.8)	0.0160	1(59%) + **16(9%)** + 7(8%) + (···)
18458 (541.8)	0.0430	**15(28%)** + 7(20%) + 1(17%) + 6(6%) + (···)
18539.5	0.0739	7(33%) + **15(30%)** + 6(9%) + **16(6%)** + (···)
18734.6	0.0055	**15(80%)** + (···)
18923 (528.5)	0.0059	**15(38%)** + 17(18%) + 7(5%) + (···)
18911.7	0.0056	**16(38%)** + 6(6%) + (···)
19174.3	0.0263	12(70%) + **16(5%)** + (···)
19252.7	0.0053	13(80%) + 14(6%) + (···)
19351.0	0.0065	13(78%) + 14(6%) + (···)
23122 (432.5)	0.0225	7(22%) + 18(19%) + 19(6%) + (···)
23192 (431.2)	0.0097	7(50%) + (···)
23380 (427.7)	0.0053	7(21%) + 18(18%) + 19(11%) + 17(9%) + (···)
23824 (419.7)	0.0097	7(23%) + **16(14%)** + 18(5%) + (···)
23918.1	0.0072	7(17%) + **16(20%) + 15(14%)** + (···)
24011.6	0.1024	17(33%) + 18(11%) + (···)
24012.7	0.0582	17(24%) + 18(8%) + 22(8%) + 2(7%) + (···)
24350 (410.8)	0.0085	17(26%) + 22(7%) + (···)
24456 (408.9)	0.0058	17(18%) + 11(8%) + 22(7%) + 18(7%) + (···)
24661 (405.5)	0.1467	12(29%) + 19(7%) + 22(7%) + (···)
24759 (403.9)	0.1577	12(33%) + 17(13%) + (···)
24802.9	0.0111	12(34%) + 22(6%) + 19(6%) + (···)
24902.4	0.0090	12(48%) + 19(7%) + (···)
24911 (401.4)	0.2280	12(20%) + 22(9%) + 19(6%) + (···)
25000.7	0.0146	19(15%) + 22(8%) + 23(8%) + 5(8%) + (···)
25099.9	0.0105	5(55%) + 23(11%) + (···)
25099.7	0.1159	5(47%) + 23(8%) + (···)
25088 (398.6)	0.0223	17(7%) + 18(6%) + 22(6%) + 7(5%) + (···)
25196.9	0.3490	17(22%) + 7(7%) + 8(6%) + 23(5%) + **16(5%)** + (···)
25331 (394.8)	0.1767	17(17%) + 20(15%) + 23(5%) + (···)
25352 (394.4)	0.0702	17(29%) + 23(6%) + 20(5%) + (···)
25448 (393.0)	0.6045	17(20%) + **16(7%) + 15(5%)** + 7(5%) + (···)
25475 (392.5)	0.1896	17(34%) + 23(12%) + (···)
25739 (388.5)	0.0793	21(25%) + 23(9%) + 19(7%) + (···)
25816 (387.4)	0.0181	21(34%) + 23(6%) + 20(6%) + 10(6%) + (···)
25965 (385.1)	0.0153	21(31%) + 10(6%) + 23(6%) + 20(5%) + (···)
26303 (380.2)	0.0080	17(38%) + 24(17%) + (···)
26333 (379.8)	0.0099	24(29%) + 17(28%) + (···)
26429 (378.4)	0.0184	20(21%) + 17(13%) + 21(5%) + (···)
26516 (377.1)	0.0088	24(26%) + 17(11%) + 25(8%) + 21(8%) + (···)
26872 (372.1)	0.0399	17(22%) + 20(12%) + (···)
26948 (371.1)	0.0218	17(49%) + 20(16%) + (···)
26987 (370.6)	0.0191	17(35%) + 20(11%) + 4(8%) + (···)
27000 (370.4)	0.0056	3(52%) + 21(7%) + (···)
27038 (369.9)	0.0056	3(61%) + 21(9%) + (···)
27070 (369.4)	0.0209	3(20%) + 20(16%) + 19(6%) + (···)
27100 (369.0)	0.0384	23(17%) + 3(14%) + 8(8%) + 17(6%) + 22(5%) + (···)
27160 (368.2)	0.0848	20(22%) + 22(11%) + 23(10%) + (···)
27183 (367.9)	0.0172	23(52%) + 4(10%) + 17(6%) + (···)
27232 (367.2)	0.0076	22(21%) + 23(14%) + 4(10%) + 17(7%) + (···)
27345 (365.7)	0.0561	17(16%) + 18(15%) + 20(7%) + 8(7%) + 4(7%) + (···)
27367 (365.4)	0.0279	17(27%) + 20(10%) + 8(9%) + 18(7%) + (···)
27409 (364.8)	0.0259	17(26%) + 9(12%) + 23(8%) + 22(7%) + (···)
27473 (364.0)	0.1173	18(22%) + 23(12%) + 20(10%) + 22(7%) + (···)
27507 (363.5)	0.2031	18(19%) + 8(14%) + 20(14%) + (···)
27608 (362.2)	0.1070	8(14%) + 22(14%) + 18(11%) + 23(8%) + 17(6%) + (···)
27725 (360.7)	0.0359	18(24%) + 8(20%) + 21(13%) + 9(6%) + (···)
27748 (360.4)	0.0127	21(26%) + 18(15%) + 20(14%) + 9(9%) + 8(6%) + (···)
27797 (359.8)	0.0055	18(27%) + 8(17%) + 17(6%) + 20(6%) + (···)
27837 (359.2)	0.0110	18(45%) + 21(13%) + 8(10%) + 20(10%) + (···)
27871 (358.8)	0.0224	21(42%) + 8(18%) + 18(5%) + (···)
27931 (358.0)	0.0232	21(41%) + 18(17%) + 10(6%) + (···)
27976 (357.5)	0.0147	21(42%) + 18(12%) + 8(8%) + (···)

The low-energy spectroscopic envelope in the computed spectrum of [μ-oxo ferrihaem]$^{4-}$ is markedly different from that calculated for [H_2O-ferrihaem]$^{1+}$. It is predominately comprised of two moderately intense, closely-spaced excitations at 18 458 and 18 536 cm^{-1} (542 and 540 nm). These excitations consist of heavily admixed $\pi\rightarrow\pi^*$ and $\pi\rightarrow d_\pi$ character, the former of which involves the transitions implicated in the Gouterman four-orbital model, $\pi\rightarrow\pi^*(15)$ and $\pi\rightarrow\pi^*(16)$. There is also a notable contribution from $\pi\rightarrow d_{xy}$ transitions in this region. The excitation at 18 707 cm^{-1} corresponds to the Q-band and has 80% contribution from $\pi\rightarrow\pi^*(15)$ and $\pi\rightarrow\pi^*(16)$ transitions. In contrast to [H_2O-ferrihaem]$^{1+}$, this excitation is very weak in [μ-oxo ferrihaem]$^{4-}$. Owing to the complex charge transfer and $\pi\rightarrow\pi^*$ admixture, the low-energy spectroscopic envelope cannot be thought of as solely a Q- or charge transfer band.

As described for H_2O-ferrihaem, assignment of underlying bands in the experimental UV-visible spectrum of μ-oxo ferrihaem was attempted using simultaneous peak deconvolution. Fitted band parameters are listed in Table 4.8 and UV-visible and MCD spectra are displayed in Fig. 4.17.

Table 4.8. Band fit parameters of μ-oxo ferrihaem.

Band Number	E_0 (cm^{-1})	σ (cm^{-1})	ΔE (cm^{-1})	$a_{UV\text{-}vis}$ ($M^{-1}\,cm^{-1}$)	a_{MCD} ($M^{-1}\,cm^{-1}\,T^{-1}$)
-	33254	970	1616	12647	-1.03
-	31811	954	1589	13813	0.23
-	30530	952	1585	17865	7.90
-	29271	951	1584	22685	8.78
-	28287	770	1282	16442	-4.11
-	27264	950	1582	28602	-15.25
1	25913	951	1583	36844	-3.59
2	25017	708	1179	19835	25.86
3	24363	950	1582	29255	-5.27
-	20636	781	1301	6150	-2.69
-	21457	681	1134	5208	0.60
-	22395	847	1410	10676	-3.52
-	23252	704	1173	7738	1.77
-	19548	793	1321	4160	1.04
-	18831	517	861	1333	-2.77
4	17984	730	1216	4102	6.49
5	17373	450	749	1333	-3.22
6	16729	684	1139	4575	22.15
7	16065	545	908	1255	-26.81
-	15281	633	1054	582	0.05
-	14160	688	1145	519	0.99
-	13001	704	1173	162	1.17

[a] See Fig. 4.17; [b] peak position; [c] bandwidth; [d] full width at half-maximum (see Section 4.2.4); [e] intensity.

The number of Gaussian functions required to produce the most parsimonious fit to both UV-visible and MCD spectra of μ-oxo ferrihaem was twenty-two. Particular attention was given to the Soret and low-energy envelope features to provide reliable fits of underlying bands (numbered in Fig 4.17). Owing to the poorly resolved features of high-energy shoulder of the Soret peak as well as the broad low-energy shoulder in the UV-visible spectrum, underlying bands fitted to these regions are less certain (bands shown in grey in Fig. 4.17).

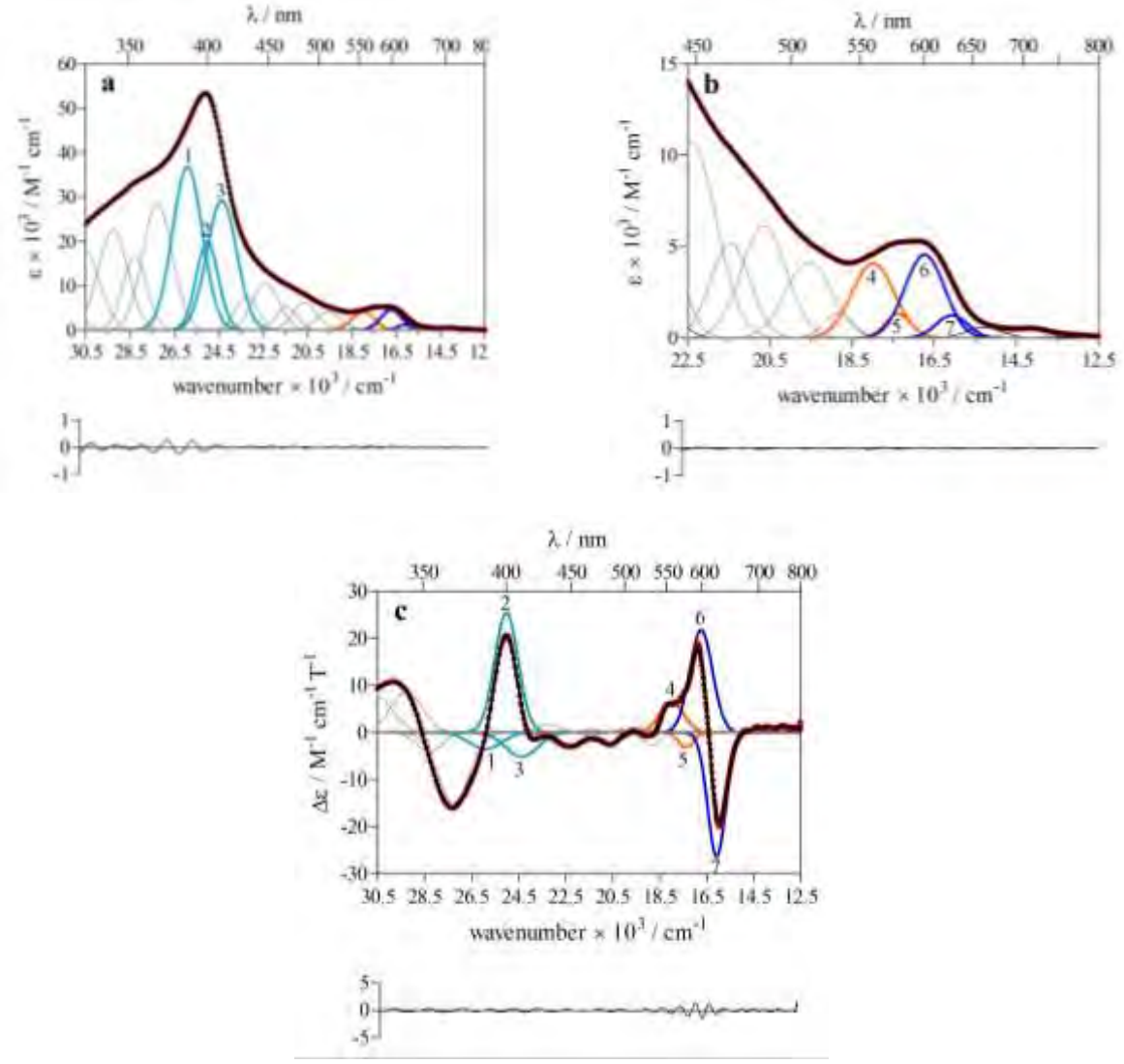


Figure 4.17. Deconvoluted UV-visible (**a** and **b**) and MCD (**c**) spectra of μ-oxo ferrihaem at room temperature. The spectrum in (b) is an enlargement of (a). Red dots represent the observed spectrum and black lines the sum of fitted Gaussian functions. Band numbers are referred to in the text. Solvent conditions are the same as described in Section 3.2.6.

As noted in Section 3.3.2, the MCD spectrum of μ-oxo ferrihaem is distinctly different from that of H_2O-ferrihaem, particularly in the Soret region where an apparent reverse A-term feature is observed. Deconvolution of this region, however, reveals the Soret peak is instead likely comprised of three Gaussian functions (bands 1, 2 and 3 in Fig. 4.17) of alternating sign, similar to that observed for Cl-iron(III)TPP and ferrihaem in cytochrome $P450_{cam}$.[209] By contrast, the A-term feature of the low-energy envelope in the MCD spectrum is indeed produced by two Gaussian functions (bands 6 and 7 in Fig. 4.17) of opposite sign. In addition, two Gaussian functions (bands 4 and 5 in Fig. 4.17) at higher energy appear to produce a weak pseudo A-term feature. These may correspond to a vibronic band given that the relative energy difference from bands 6 and 7 is similar to that observed for the Q- and Q_v-bands in H_2O-ferrihaem (approx. 1280 cm^{-1}).

Assignment of deconvoluted bands to computed excitations for [μ-oxo ferrihaem]$^{4-}$ is considerably more complex than for [H_2O-ferrihaem]$^{1+}$ owing to the much larger number of excitations calculated for the former species. In addition, there is an added complication that the UV-visible spectrum is calculated for [μ-oxo ferrihaem]$^{4-}$ in which the unpaired electrons are spin-paired to give S = 0. In reality, this only occurs at 0 K and as the temperature is increased to the ambient experimental conditions thermal uncoupling occurs.[220] Thus exact comparison of the computed and experimental UV-visible spectrum is not feasible. To determine the effects of uncoupling electrons would have on the calculated excitations, the UV-visible spectrum was computed for the model compound, [μ-oxo ferriporphine]^0, in the high-spin (S = $^{10}/_2$) and spin-paired (S = 0) states (see Fig. 4.18). β-Spin MOs associated with major transitions are presented for the high-spin and antiferromagnetically coupled molecules in Appendix A Fig. A1 and A2, respectively. Additional MOs are also included for comparison with [μ-oxo ferrihaem]$^{4-}$ (see Fig. 4.15). Compositions of [μ-oxo ferriporphine]^0 β-spin MOs are detailed in Appendix A Tables A1 and A2 for the high-spin and antiferromagnetically coupled states, respectively.

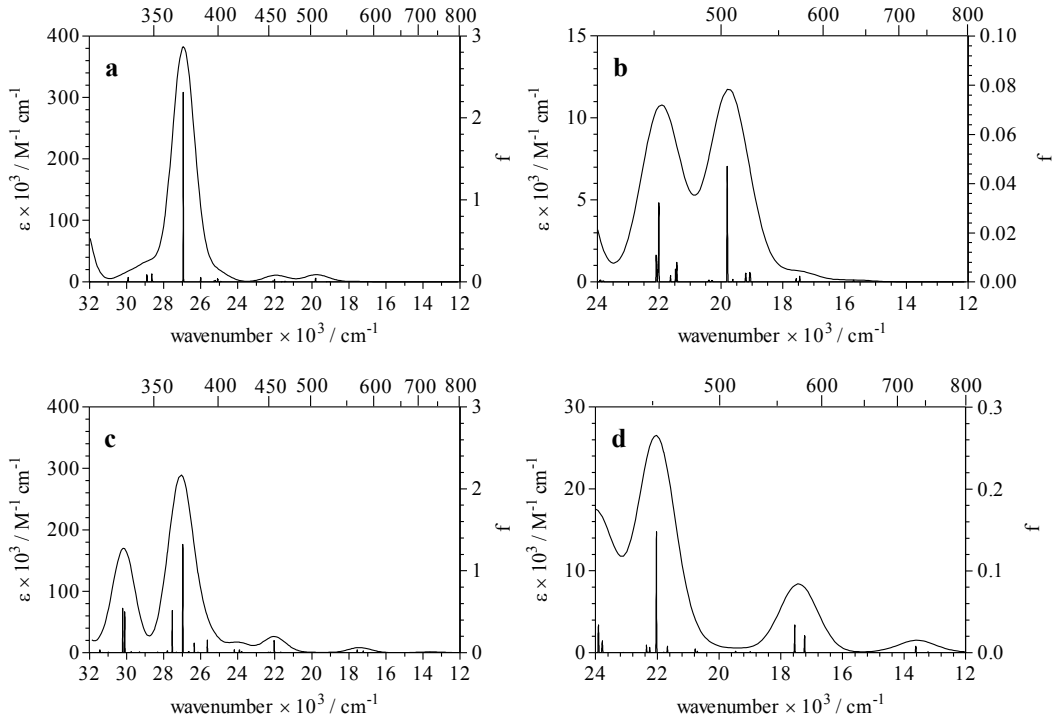


Figure 4.18. Computed UV-visible spectra of [μ-oxo ferriporphine]^0 in the (**a** and **b**) antiferromagnetically coupled (S = 0) and (**c** and **d**) high-spin (S = $^{10}/_2$) state. Excitations in the calculated spectrum are represent by vertical bars plotted on the right axis in oscillator strength, *f*, and have been convoluted using Gaussian functions with a half-width of 700 cm^{-1}. Spectra were calculated using TD-DFT with PBE0/LANL2DZ:LANL2DZdp.

Major features of the antiferromagnetically coupled UV-visible spectrum for [μ-oxo ferriporphine]^0 compare well with those observed in the corresponding spectrum of [μ-oxo ferrihaem]$^{4-}$. These include the Soret peak at 26 954 cm^{-1} (371 nm) and two lower-energy bands at 21 978 and 19 802 cm^{-1} (445 and 505 nm). Two intense, closely-spaced excitations underlie the Soret feature and predominately involve the $a_{1u}/a_{2u}(\pi)\rightarrow e_g(\pi^*)$ transitions implicated in the Gouterman four-orbital model. A small $\pi\rightarrow d_\pi$ charge transfer contribution is also present, however. A single excitation produces the peak at 21 978 cm^{-1} which is almost exclusively comprised of $\pi\rightarrow d_z^2$ transitions, while the envelope at 19 802 cm^{-1} is composed of two closely-spaced excitations with an almost equal $a_{1u}/a_{2u}(\pi)\rightarrow e_g(\pi^*)$ and $\pi\rightarrow d_\pi$ admixture.

The UV-visible spectrum calculated for high-spin [μ-oxo ferriporphine]^{0} displays some similarity to the antiferromagnetically coupled species. The Soret peak, while less intense, is shifted to slightly higher energy at 27 100 cm^{-1} (369 nm) but is still comprised of two closely-spaced excitations. These excitations involve mainly $\pi \rightarrow \pi^*$ transitions originating in the four highest energy occupied MOs and terminating in the four unfilled $e_g(\pi^*)$ MOs. In addition, a significant charge transfer component is also present and has an approx. 10% greater contribution than that observed for the antiferromagnetic species. A prominent higher-energy band is also observed which has varying degrees of $\pi \rightarrow d_\pi$ and $\pi \rightarrow \pi^*$ admixture, the latter of which originates in MOs below the $a_{1u}/a_{2u}(\pi)$ pair. The low-energy feature is markedly red-shifted to 17 422 cm^{-1} (574 nm). Two closely-spaced excitations underlie this envelope and are predominately comprised of $\pi \rightarrow \pi^*$ transitions. A minor $\pi \rightarrow d_\pi$ charge transfer contribution is also observed, the proportion of which is smaller than that calculated for the spin-paired [μ-oxo ferrihaem]$^{4-}$. The excitation corresponding to $\pi \rightarrow d_z^2$ charge transfer transitions in the antiferromagnetic species shifted and instead contributes to the low-energy feature in the paramagnetic species. In addition, a new feature was observed between the Soret and low-energy envelope which has predominately $\pi \rightarrow d_\pi$ character. These findings suggest that despite the differences in spectroscopic appearance and MO energies, the composition of the major features in the UV-visible spectrum remain generally similar, with the exception of the $\pi \rightarrow d_z^2$ excitation which shifts to higher energy in the high-spin species.

Given the general overall similarities in the spectra of high-spin and antiferromagnetically coupled [μ-oxo ferriporphine]^{0}, it is reasonable to conclude that the spectrum of the high-spin [μ-oxo ferrihaem]$^{4-}$ would not be greatly different from that calculated for the spin-paired species. Thus, even if there is some contribution from the uncoupled [μ-oxo ferrihaem]$^{4-}$, the spectrum is unlikely to be drastically altered. For this reason, the Soret peak of μ-oxo ferrihaem presumably is comprised of mainly $\pi \rightarrow \pi^*$ transitions with a significant proportion of $\pi \rightarrow d_\pi$ charge transfer character also present. In the case of the $\pi \rightarrow \pi^*$ character, the dominant transitions originate in MOs with substantial delocalisation onto vinyl substituents. As noted for H_2O-ferrihaem, this may account for similar experimentally observed shifts in the Soret band of a number of 2,4-disubstituted μ-oxo ferriporphyrins.[74] Furthermore, of the reported μ-oxo ferriporphyrins, μ-oxo ferrihaem displays the smallest difference in energy between the high-energy shoulder and Soret peak. This is understandable given that μ-oxo ferrihaem is the only molecule in the data set that has substituents conjugated with the porphyrin ring system and

the fact that the high-energy shoulder is computed to have many transitions involving MOs with prominent delocalisation onto the vinyl groups. Larger splitting is observed in the computed spectrum of [μ-oxo ferriporphine]^0.

Excitations making up the low-energy feature most likely consist of π→π*(15), π→π*(16), CT(6) and CT(7) transitions as these are conserved in both sets of spectra computed for [μ-oxo ferriporphine]^0 molecules as well as for the antiferromagnetic [μ-oxo ferrihaem]$^{4-}$ species. In contrast to the Soret peak, this band is relatively insensitive to porphyrin substitution, presumably owing to the fact that no transitions involve MOs with major delocalisation onto the vinyl groups. The broad low-energy shoulder of the Soret band may consist of π→d_π or π→d_z^2 charge transfer transitions or both, while the latter may also contribute to the low-energy feature. The composition of these regions may be better resolved using variable-field variable-temperature MCD which is able to identify the polarisation of electronic transitions because calculated transition dipole moments for π→d_z^2 transitions lie 70 ° off the *x*,*y*-plane. A summary of the dominant spectroscopic features in the experimental and computed UV-visible spectra is presented in Table 4.10.

Table 4.10. Component band positions obtained from deconvolution of UV-visible absorbance and MCD spectra of [μ-oxo ferrihaem]$^{4-}$. Prominent calculated excitations obtained from TD-DFT calculations (PBE0/LANL2DZ:LANL2DZdp) making up the major spectroscopic envelope features are shown for comparison. [a, b]

experimental		computed	
energy range: 26500 – 24000 cm^{-1} (Soret)		*energy range*: 26500 – 24800 cm^{-1} (Soret)	
		25739 (389)	π→π*(21), π→π*(23), π→π*(19)
		25475 (393)	π→π*(17), **π→π*(15)**, **π→π*(16)**, CT(7)
band 1	25913 (386)	25448 (393)	π→π*(17), π→π*(23), π→π*(20)
band 2	25017 (400)	25352 (394)	π→π*(17), π→π*(20), π→π*(23)
band 3	24363 (410)	25331 (395)	π→π*(17), CT(7), CT(8), π→π*(23), **π→π*(16)**
		25197 (397)	π→π*(17), π→π*(23), π→π*(20)
		25088 (399)	π→π*(17), π→π*(18), π→π*(22), CT(7)
		24911 (401)	CT(12), π→π*(22), π→π*(19)
energy range: 15500 – 19000 cm^{-1}		*energy range*: 15500 – 19000 cm^{-1}	
band 4	17984 (556) [c]		
band 5	17373 (576) [c]	18536 (540)	CT(7), **π→π*(15)**, CT(6), **π→π*(16)**
band 6	16729 (598)	18458 (542)	**π→π*(15)**, CT(7), CT(1), CT(6)
band 7	16065 (622)		

[a] Transitions shown in bold have counterparts in porphinate that correspond to those involved in Gouterman's four-orbital model. [b] Band positions and excitations in wavenumbers (cm^{-1}) with wavelength in parentheses (nm). [c] Possible vibronic bands.

4.5.5 Account of overall spectroscopic differences between H_2O-ferrihaem and μ-oxo ferrihaem.

The characteristic spectroscopic differences between $[H_2O\text{-ferrihaem}]^{1+}$ and $[\mu\text{-oxo ferrihaem}]^{4-}$ can be attributed to the changes in MO energies, in particular, the filled π MOs and selected metal d-orbitals, but not π* MOs. In the case of π MOs, energies are somewhat higher in $[\mu\text{-oxo ferrihaem}]^{4-}$ and result from the increased delocalisation possible between the two ring systems. The unfilled metal d_π- and d_z^2-orbitals, however, are substantially higher in energy owing to increased porphyrin π* character. In the case of d_π-orbitals, increased π* MO character likely arises from better π overlap of the lone pair orbitals of the O^{2-} ligand consistent with it being a weaker-field ligand. This raises these MOs closer in energy to porphyrin $e_g(\pi^*)$ MOs. Both lone pairs of the oxide ligand overlap with d_π-orbitals in $[\mu\text{-oxo ferrihaem}]^{4-}$ whereas overlap is only observed with one of the d_π-orbitals in $[H_2O\text{-ferrihaem}]^{1+}$ (see Fig. 4.19). Improved overlap of the d_z^2-ortibals is likely a result of the porphyrin doming.

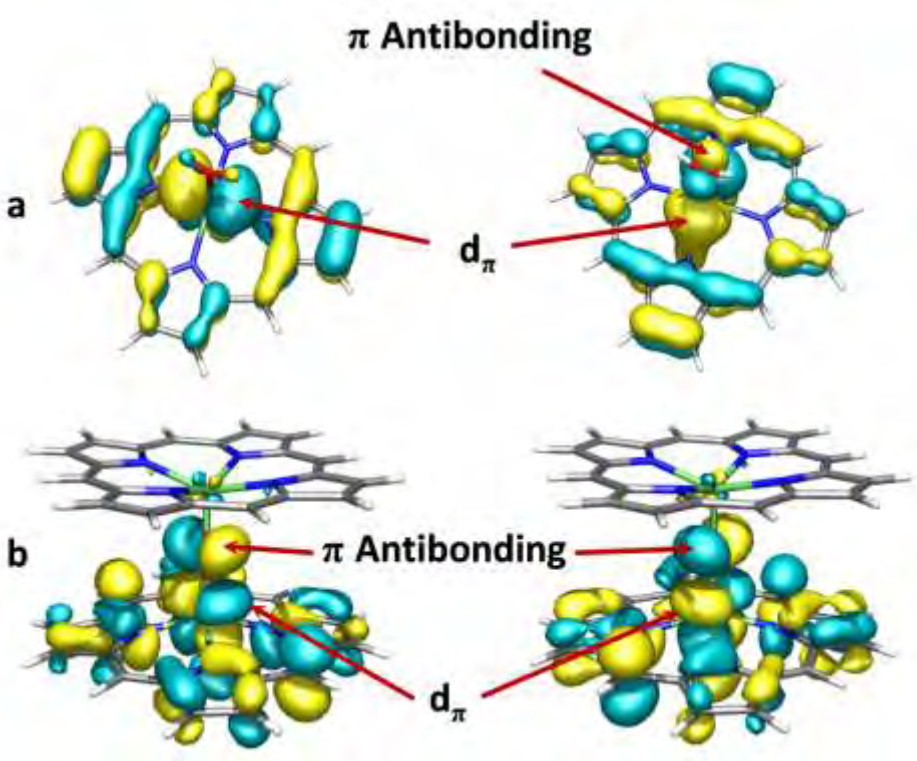


Figure 4.19. Metal d_π-orbitals of (**a**) $[H_2O\text{-ferriporphine}]^{1+}$ ($S = {}^5/_2$) and (**b**) antiferromagnetically coupled $[\mu\text{-oxo ferriporphine}]^0$ showing axial ligand π-overlap. MOs are similar to those observed in ferrihaem species but the simpler porphine system has been used for the sake of clarity.

These changes in MO energies cause $\pi \rightarrow \pi^*$ transitions to occur at lower energies in the calculated spectrum of [μ-oxo ferrihaem]$^{4-}$ and $\pi \rightarrow d_\pi$ and $\pi \rightarrow d_z^2$ transitions to occur at higher energies. Thus the broadening of the Soret band in μ-oxo ferrihaem arises in part from $\pi \rightarrow \pi^*$ transitions originally at higher energy in [H_2O-ferrihaem]$^{1+}$, while the low-energy broadening likely results from $\pi \rightarrow d_\pi$ and $\pi \rightarrow d_z^2$ transitions shifting to higher energy. The lower energy envelope of [μ-oxo ferrihaem]$^{4-}$ centred at 593 nm consists of heavily admixed $\pi \rightarrow \pi^*$ and charge transfer transitions. It can be thought to arise from the shifting of the Q-band in [H_2O-ferrihaem]$^{1+}$ to lower energy where it overlaps with the charge transfer peaks which are now shifted to higher energy and are comprised of $\pi \rightarrow d_{xy}$ transitions.

4.6 Conclusions

Structures of ferrihaem species calculated using DFT with the OPBE functional and LANL2DZ basis set were in good agreement with experiment. IR spectra calculated for these structures, in general, reproduced experimental counterparts reasonably well, with the exception of the υ(Fe-O-Fe) stretch of [μ-oxo ferrihaem]$^{4-}$ which was underestimated in energy. On the other hand, carbonyl υ(C=O) and vinyl υ(C=C) stretches mostly agreed with experiment. UV-visible spectra of [H_2O-ferrihaem]$^{1+}$ and [μ-oxo ferrihaem]$^{4-}$ calculated using TD-DFT with PBE0 functional and the LANL2DZ:LANL2DZdp basis set combination in implicit water (SMD) were found to reproduce the overall features of the corresponding experimental spectra remarkably well. By contrast however, the calculated spectrum of [HO-ferrihaem]$^{2-}$ in either the high- ($S = {}^5/_2$) or intermediate-spin ($S = {}^3/_2$) state did not agree with the experimental spectrum. This was attributed to the spin-admixed character of this species as determined from magnetic susceptibility measurements which cannot be reproduced by the Gaussian software package used for DFT calculations.

In the case of H_2O-ferrihaem, previous assignment of the experimental spectrum based on the simplified porphinate model is generally in agreement with that determined using TD-DFT. Specifically, the Q-band arises from $\pi \rightarrow \pi^*$ transitions originating in MOs displaying a_{1u} and a_{2u} symmetry in D_{4h} porphinate and terminating in unfilled $e_g(\pi^*)$ MOs. The same transitions make substantial contributions to the Soret peak, as previously implied by the Gouterman four-orbital model, however, there are also significant contributions from $\pi \rightarrow e_g(\pi^*)$ transitions originating in lower energy MOs. By contrast, the

spectroscopic envelope at lower energy than the Q-band, previously assigned as solely a $\pi \rightarrow d_\pi$ charge transfer band, was found in this study to rather have heavily admixed $\pi \rightarrow \pi^*$ and $\pi \rightarrow d_\pi$ character. In the case of μ-oxo ferrihaem, both the Soret band and low-energy spectroscopic feature are made up of multiple overlapping excitations with both $\pi \rightarrow \pi^*$ and charge transfer character. In particular, the overlap observed in the lower energy envelope prevents this from being referred to as solely a Q- or charge transfer band and rather should be thought as a superposition of the two. The features of the μ-oxo ferrihaem spectrum arise from the antiferromagnetic coupling between iron centres which allow both α- and β-spin transitions; the increased delocalisation over the two porphyrin rings; doming of the porphyrin ring; and, finally, the decreased ligand-field strength of the O^{2-} bridging ligand as compared to water. The ability of TD-DFT to reproduce the experimental spectroscopic features of H_2O- and μ-oxo ferrihaem, both of which can be rationalised in terms of structure and MOs, indicates that such spectra must be characteristic of these species.

Given the capability of DFT and TD-DFT functionals and basis set combinations to reproduce structural, vibrational and electronic properties of ferrihaem species, it was concluded that the resulting QM models were a sufficiently accurate approximation of the real molecules for use in MM parameterisation.

CHAPTER 5

MOLECULAR DYNAMICS SIMULATION OF FERRIHAEM SPECIES

5.1 Introduction

While several recent experimental studies have provided a clearer understanding of ferrihaem speciation in solution,[75, 76, 163] little is known regarding the structure and solvation of the various species. Probing these solution-phase properties is difficult using experimental methods, particularly in the case of ferrihaem where the paramagnetic iron centre precludes techniques such as NMR spectroscopy. On the other hand, computational simulations such as MD are better equipped to investigate these properties owing to the atomistic-level detail they provide. These methods, however, are heavily reliant on accurate force fields.

In this study, the CHARMM software package[246] was used to perform MD simulations on ferrihaem species with the aim of better understanding the solution behaviour of these molecules. Previous CHARMM MD simulations of systems involving the haem moiety have largely been confined to haemoproteins in which the iron is in the ferrous state.[323-327] Indeed, while the CHARMM software package contains parameters for ferrous haem, none are provided for the ferric form. Two relatively recent ferrihaem models have been reported that are compatible with the CHARMM force field.[328, 329] Both models, however, were developed for six-coordinate ferrihaem species and for use in haemoproteins. Consequently, these reported models are unsuitable for simulating five-coordinate ferrihaem species in aqueous solution (i.e. not protein-bound). While a previous study adapted parameters to model the neutral π-π dimer, $[(H_2O\text{-ferrihaem})_2]^0$, in vacuum and aqueous solution,[89] no force field compatible with CHARMM has been specifically developed to study solvation and structure of free ferrihaem species.

This chapter details the development of ferrihaem CHARMM force field parameters created specifically to model five-coordinate ferrihaem species in aqueous solution. Molecules parameterised included monomeric $[H_2O\text{-ferrihaem}]^0$, $[H_2O\text{-ferrihaem}]^{1-}$ and $[HO\text{-ferrihaem}]^{2-}$ as well as dimeric [μ-oxo ferrihaem]$^{4-}$. Diffusion coefficients were calculated and compared to available experimental data, solvation was investigated using spatial distribution functions (SDFs) and MD solution structures used to fit experimental EXAFS spectra. The newly-parameterised models provide a more accurate description of five-coordinate ferrihaem species simulated in aqueous solution using MD.

5.2 Methods

5.2.1 Parameterisation procedure for ferrihaem species.

A new residue describing ferrihaem was created. To ensure consistency with the ferrous haem residue in CHARMM, the protoporphyrin IX moiety of ferrihaem was fashioned according to the same topology. Owing to the structural differences between the five-coordinate ferrihaem species of interest in this work and six-coordinate haem species previously parameterised, the creation of a new iron atom type was required which was labelled "FE3" (see figure 5.1). Four further atom types were created, three describing axial ligands labelled "OLC" for HO^-, "OLW" for H_2O and "OLH" for O^{2-}, and one to describe hydrogen atoms attached to the axial ligands, "HLW". In a similar approach to Autenrieth *et al.*,[328] Lennard-Jones parameters for Fe^{2+} were used for the new Fe^{3+} atom type, although in this case values reported by de Hatten *et al.* for ferrocene were employed.[330] Oxygen and hydrogen Lennard-Jones parameters were taken from corresponding atoms of TIP3P water. Ligands were coordinated axially to the iron centre, on the under face of ferrihaem species with orientation displayed in Fig. 5.1.

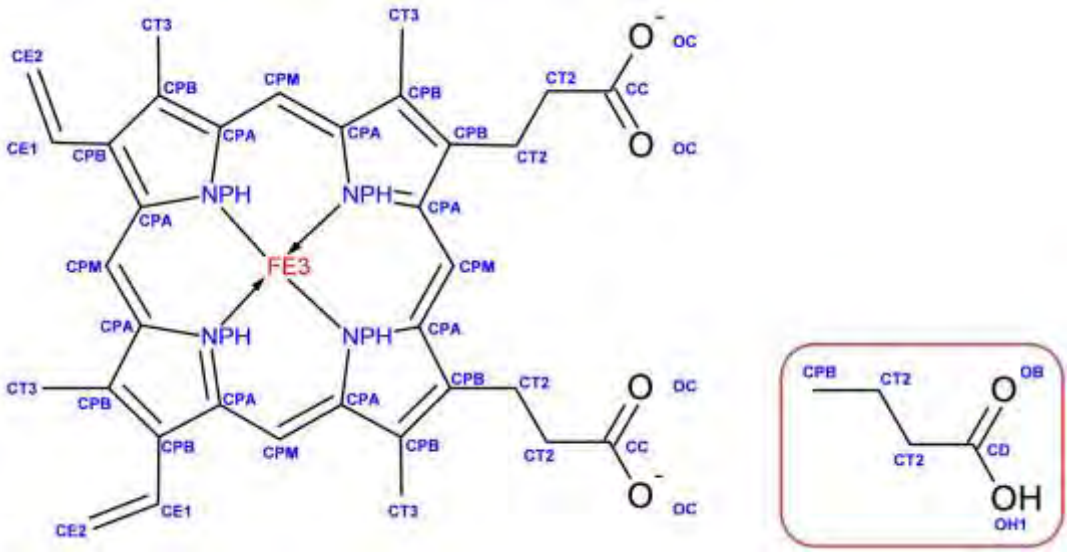


Figure 5.1. Topology and atom typing of the ferrihaem residue, HEME, employed in CHARMM. A new atom type, FE3, was created. Hydrogen atoms and axial ligands have been excluded for the sake of clarity. **Inset.** $[H_2O\text{-ferrihaem}]^0$ has one propionic acid with atom typing shown.

Initial atomic charges for the monomeric H_2O- and HO-ferrihaem species were determined by averaging the Mulliken charges for the equivalent atom types shown in Fig. 5.1. Initial atom charges for the μ-oxo dimer were calculated employing eq. 5.2.1 for equivalent or near equivalent atom types of [HO-ferrihaem]$^{2-}$. In this expression $q_{init}(\mu)$ is the initial atomic charge for [μ-oxo ferrihaem]$^{4-}$, $q_{opt}(HO)$ is the optimised atomic charge for [HO-ferrihaem]$^{2-}$ and $q_{Mull}(\mu)$ and $q_{Mull}(HO)$ are the Mulliken atomic charges for [μ-oxo ferrihaem]$^{4-}$ and [HO-ferrihaem]$^{2-}$ respectively. Excess charge was summed into the oxide ligand. In accordance with CHARMM methodology,[246, 248] for both monomeric and dimeric ferrihaem species, aliphatic hydrogen atoms were assigned values of 0.09 e and excess charge summed into the attached carbon atom.

$$q_{init}(\mu) = \frac{q_{opt}(HO) \times q_{Mull}(\mu)}{q_{Mull}(HO)} \qquad (5.2.1)$$

Interaction energies between a water molecule and selected hydrogen, iron, nitrogen and oxygen atoms were determined for monomeric species using OPBE/6-31G* and scaled by a factor of 1.12 (see Section 5.3.1 for details). As suggested by the CHARMM parameterisation protocol for neutral molecules, interaction energies for [H_2O-ferrihaem]^0 were scaled by a further 1.16. Corresponding water interaction energies (WIEs) were calculated in CHARMM and charges adjusted to reproduce DFT values. Furthermore, MM atomic charges were made to reproduce QM ESPs by using the FitCharge module in CHARMM. This module minimises χ^2 in eq. 5.2.2, where V_{MM} and V_{QM} are the MM and QM potentials respectively, *i* the grid point and *N* the total number of grid points. The added constraint of reproducing the QM dipole was also enforced. WIEs were then recalculated using ESP-modified charges, and where markedly different from DFT values, the two procedures performed iteratively until satisfactory agreement between both WIEs and ESPs was obtained.

$$\chi^2 = \sum_{i}^{N} \left(V_{MM,i} - V_{QM,i}\right)^2 \qquad (5.2.2)$$

Optimisation of structural parameters relating to Fe-O_{ax} and Fe-N bonds; Fe-N-C_{α}, Fe-O_{ax}-H_{ax} and Fe-O_{ax}-Fe angles; and the N*-Fe-C_{α}-C_{α} improper dihedral followed. Minimised MM structures were compared to DFT structures (OPBE/LANL2DZ) as well as to average bond lengths and angles of related iron(III)porphyrin crystal structures in the CSD. A schematic of the parameterisation procedure used is shown in Fig. 5.2.

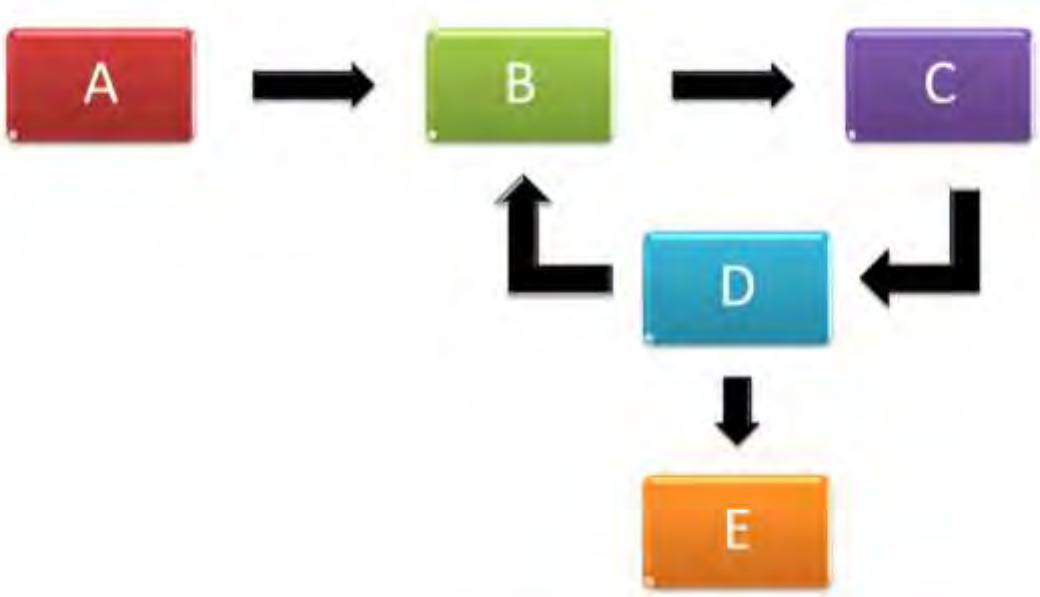


Figure 5.2. Diagrammatic representation of the procedure used to optimise atomic charges and structural parameters. (**A**) Initial atomic charges were optimised to reproduce QM WIEs using QM-structural geometry. (**B**) Atomic charges were modified iteratively to fit QM ESPs and reproduce QM WIEs. (**C**) MM-structure with optimised atomic charges was allowed to minimise and the structure compared to QM and crystal structure data. Structural parameters were modified iteratively until satisfactory agreement with target data was obtained. (**D**) WIEs and ESPs were generated from MM-minimised structure. If values were markedly different from QM, re-optimisation of atomic charges was conducted (**B**) and the entire process performed iteratively until structural and electronic properties were in agreement. (**E**) Parameterisation was deemed complete when WIEs, ESPs and structural data produced satisfactory agreement with target data.

5.2.2 MD simulation and analysis.

MD simulations of ferrihaem species were conducted using the CHARMM academic code.[246] All ferrihaem molecules were placed in the centre of a box of length 55 Å which contained 5832 TIP4P-Ew water molecules, pre-equilibrated for 2 ns at 298 K under canonical ensemble (NVT). Sodium ions were used to neutralise the overall charge when required and were placed in the corners of the water box in order to minimise interactions with the solute. The system was minimised by steepest decent (100 steps), followed by conjugate gradient (200 steps) methods. Equilibration dynamics were performed under isobaric-isothermal ensemble (NPT) for 2 ns from which the average box length (over the final 200 ps of equilibration) was determined. This was used in subsequent NVT simulations which included 2 ns equilibration and 10 ns production dynamics. All simulations employed periodic boundary conditions and were performed using the Leap-frog Verlet integrator with 1 fs time-step at 298 K. Lennard-Jones interactions were smoothly decreased to zero between 12.0 and 14.0 Å using an atom-based switching function, while electrostatic interactions were calculated using particle-mesh Ewald with 60 Å grid size and 16 Å real-space cutoff. Simulations were conducted on the TSESSEBE cluster at the CHPC across 8 Intel Nehalem or Westmere processors (2.93 GHz) and required approx. 7 weeks of continuous calculation to produce the necessary data.

Diffusion coefficients, *D*, were calculated using the mean squared displacement (MSD) method in a similar procedure to that previously reported for metal chloride molecules.[331] Briefly, this involved separating the 10 ns simulation into 1 ns blocks, calculating the average MSD for each iron atom over 250 ps (using eq. 5.2.3) and plotting the resultant values against time (1 ps interval). An expression describing the linear regression fitted over the range 30 to 80 ps was then used to determine diffusion constants according to eq. 5.2.4. Diffusion coefficients of monomeric ferrihaem species are thus reported as the average and standard deviation of 10 values. For dimeric species, the diffusion coefficient of the complex is reported as the average of the two diffusion coefficients calculated for each iron atom.

$$MSD = \langle |r(t) - r(0)|^2 \rangle \quad (5.2.3)$$

$$MSD = 6D\Delta t \quad (5.2.4)$$

SDFs were calculated using previously reported methodology.[331-334] This method is similar to the commonly used radial distribution function (RDF) which sorts the distances between pairs of atoms, r_{ij}, into a histogram where each bin has a width δr.[335] This is done for every frame in a MD trajectory to improve sampling. The histogram is then normalised by dividing the number in each bin by the total number of pairs, followed by normalisation using the average number of atoms in the same δr for an ideal gas or liquid with the same density, ρ.[335] Because RDFs are calculated as a function of distance only, directionality is lost. SDFs on the other hand, produce the same information as RDFs except orientation of molecules are retained. In the case of SDFs, a grid (subdivided into cubic segments) was placed around the solute of interest and a histogram of solvent atoms within each cubic segment over the course of the MD simulation produced.[334] Importantly, the rotational and translational diffusion of the solute was removed to ensure orientation-dependent solvent-binning was correct. This was done by reorienting each frame from the MD simulation to a reference structure and excluding structures with root mean squared deviations greater than a specified tolerance.[331, 334] Histograms were normalised to bulk water using eq. 5.2.5, where $\rho_{norm}(x_i, y_j, z_k)$ is the normalised probability density function for grid point (x_i, y_j, z_k), ρ the non-normalised probability density, n_{xbin}, n_{ybin} and n_{zbin} the number of grid divisions in the x-, y- and z-directions, n_{atom} the number of solute atoms in each frame and n_{frames} the number of frames.[332] Thus the normalised density of 1 corresponds to a bulk solvent, while that of 1.5 represents 50% greater density than bulk, etc.[331] Gaussian distribution functions were centred on each solvent site within the cubic segments to represent the distribution of atoms across the grid.[331]

$$\rho_{norm}(x_i, y_j, z_k) = \rho(x_i, y_j, z_k)\frac{n_{xbin}n_{ybin}n_{zbin}}{n_{atom}n_{frames}} \qquad (5.2.5)$$

Reference geometries used for monomeric and [μ-oxo ferrihaem]$^{4-}$ species were selected by determining the frame with lowest root mean squared deviation from the average structure calculated over the course of the 10 ns simulation. In the case of π-π dimeric species, a cluster analysis was performed in CHARMM using the torsion angle defined between the C_{meso} atom adjacent to propionate side chains and iron centre of both porphyrins, i.e. $C_{meso}(1)$-Fe(1)-Fe(2)-$C_{meso}(2)$. The frame nearest the centre of the cluster with largest population was used as the reference structure in SDF calculations.

Analysis of porphyrin orientation for dimeric ferrihaem species was performed by calculating the least-squares plane (LSP) through iron and porphine moieties of the two porphyrins rings, A and B. The entire dimer was then transposed so the LSP through porphyrin A (LSP_A) lay in the x,y-plane. Vertical interplanar distance between LSP_A and LSP_B was given by the z-component of LSP_B. Horizontal (lateral) shifts were determined by measuring the distance between the centre of mass (COM) of LSP_B and the COM of LSP_A projected onto LSP_B. Interplanar angles, α, were calculated from the expression in eq. 5.2.6, where the numerator involves the dot product of normal vectors of each LSP (N_A and N_B) and denominator is the product of normal vector magnitudes (n_A and n_B). Interplanar torsion was measured from the dihedral specified using meso carbon atoms between propionates and the LSP COM on each porphyrin, $C_{meso,A}$-COM_A-COM_B-$C_{meso,B}$.

$$\alpha = \cos^{-1} \frac{N_A \cdot N_B}{n_A \times n_B} \qquad (5.2.6)$$

5.2.3 Magnetic susceptibility and EXAFS measurements.

The magnetic susceptibility of ferrihaem in two solutions was determined using the Evans NMR method (see Section 2.5.1).[222] The first involved adding 0.4 mL polyethyleneglycol 400 (PEG400) to 1.35 mL of a ferrihaem solution which consisted of 3 mg haematin dissolved in NaOD (26 mM). The resulting solution was pD adjusted to 7.4 by adding 0.05 mL H_3PO_4 (0.1 M). The second solution involved adding 0.8 mL PEG400 to 1.2 mL of a ferrihaem solution which comprised 3 mg haematin dissolved in NaOD (1.6 mM). Reference solutions were prepared in the same manner with the exclusion of haematin.

The EXAFS spectrum of solid μ-oxo ferrihaem was recorded on a powdered sample prepared as described in Section 3.2.3 which was placed between two Kapton tapes attached to Tefon cells (DuPont, Wilmington, Delaware). The EXAFS spectrum of μ-oxo ferrihaem in solution was recorded on a sample prepared by diluting 0.25 mL of a haemin stock solution (20 mM in 0.1 M NaOH) in 0.2 mL PEG400, 0.4 mL DMSO and 0.1 mL N-cyclohexyl-2-aminoethanesulfonic acid (CHES) buffer (0.2 M, pH 10). The

resulting solution was pH adjusted to 10 using 1 M perchloric acid and diluted to 1 mL with water. Stable solutions were only obtained when using PEG400 as a cryoprotectant, whereas using glycerol caused ferrihaem to precipitate from solution. The sample solution was rapidly frozen in liquid nitrogen following injection into 0.1 mL Teflon cells consisting of two Kapton windows (Goodfellow Cambridge, Cambridge, UK). All EXAFS data were acquired at the Australian Synchrotron XAS beamline (1.9 T Wiggler). Four iron K_α edge (E_0 = 7112 eV) X-ray absorption scans up to k = 12 Å^{-1} were recorded for each sample in fluorescence mode employing a 100-element liquid nitrogen Ge detector. Each scan was collected at different spots on the sample in order to monitor radiation damage. This was tested by quick XANES measurements at 30 min exposure intervals. Solid samples were recorded at room temperature, while liquid samples were acquired at approx. 15 K using a helium displex cryostat. An ionization chamber was used to monitor incident X-ray intensity and the stability of the monochromator energy was monitored for all spectra by simultaneous accumulation of an iron foil spectrum in transmittance mode.

Processing of EXAFS spectra was conducted by extracting χ(k) oscillations from experimentally measured absorption coefficients. This was achieved using an automated background subtraction algorithm (AUTOBK) implemented in the ATHENA software package.[336] EXAFS χ(k) oscillations were analysed employing the ARTEMIS software package[336] using *ab initio* theoretical amplitude, phase and mean-free path factors calculated by the FEFF code[337] provided in the ARTEMIS interface. Least-squares fits to the experimental spectra were conducted using constrained and restrained refinement procedures in order to reduce the number of free parameters and increase determinacy of the model. Fitting was approached by successive refinements in k-space (2.4 ≤ k ≤ 9.8 Å^{-1}) with k^1, k^2 and k^3 weights of χ(k) data. Final refinements with k^3 in R-space (1 ≤ R ≤ 4.3 Å) subsequently followed. Refinement involved minimising the values of R-factor and χ^2 (reduced goodness-of-fit) which are described by equations 5.2.7 – 5.2.9:[338, 339]

$$R = \frac{\sum_i^N f(r_i)^2}{\sum_i^N \chi_{data}(r_i)^2} \tag{5.2.7}$$

$$\chi^2 = \frac{N_{ind}}{\varepsilon_k^2 N(N_{ind} - N_{var})} \sum_{i=1}^{N} f(r_i)^2 \tag{5.2.8}$$

$$f(r_i)^2 = [\chi_{data}(r_i) - \chi_{model}(r_i)]^2 + \sum \left(w \frac{\Delta_{restraint}}{\sigma_{restraint}} \right)^2 \tag{5.2.9}$$

where $f(r_i)$ is the minimised function consisting of the real and imaginary components of the differences between data and model XAFS, $\chi_{data}(r_i)$ and $\chi_{model}(r_i)$ respectively, over the fit range r_i; $\Delta_{restraint}$ is the penalty function and is 0 if the restraint expression is satisfied; $\sigma_{restraint}$ is the uncertainty value for the restraint; w is an amplifier which determines the magnitude of the penalty (w = 100 and 1000 for solid-state and solution fits respectively). ε_k is the measurement uncertainty, N_{var} is the number of variables in the fit, N is the number of data points in R-space and N_{ind} is the number of independent points.

All multiple scattering paths having amplitudes of 10% or greater were considered in the fit as well as all triple scattering paths up to four legs (l = 4). The EXAFS spectrum of the solid was fitted using an input structure generated from the crystal structure of μ-oxo ferrihaem dimethyl ester.[72] Six first- and six second-order cumulants, ΔR and σ^2 respectively, were used as variables in the fitting procedure. These described the oxide bridging ligand (O_{oxo}), pyrrole nitrogen atoms, C_α, C_β, C_{meso} and Fe atoms. Second order cumulants for collinear multiple scattering paths were described by $q \times \sigma^2$ as suggested by Hudson *et al*.[340], where q can range between 0 and 4. In the case of Fe-O_{oxo} and O_{oxo}-Fe-O_{oxo} collinear paths, a value of 4 was used for q. The EXAFS spectrum recorded on the solution was fitted employing the solvated structure of [μ-oxo ferrihaem]$^{4-}$ generated from SDF calculations as the initial input geometry. MD simulation data were used to estimate ΔR and σ^2 values from eq. 5.2.10 and 5.2.11:[341, 342]

$$\Delta R = \alpha \times R_{i,ave} \tag{5.2.10}$$

$$\sigma_i^2 = \beta \times \left(R_{i,ave} - R_i\right)^2 \tag{5.2.11}$$

where i refers to the scattering atom, R_i the distance of the atom from Fe, $R_{i,ave}$ the average R_i and α and β adjustable scale factors. R_i and $R_{i,ave}$ were computed using the 200 000 structures generated over 10 ns of MD simulation. Amplitude reduction factors and photoelectron energy threshold adjustment parameters, S_0^2 and ΔE_0 respectively, were fitted for both solid- and solution-state EXAFS spectra.

5.3 Force Field Parameterisation

5.3.1 Atomic charge parameterisation.

To parameterise ferrihaem species, target data from QM calculations were required. Given the size of these molecules, DFT and not higher levels of theory, was used to generate the necessary data. This included geometry-optimised structures, Mulliken charges, ESPs and WIEs. The DFT-optimised structures of $[H_2O\text{-ferrihaem}]^0$, $[H_2O\text{-ferrihaem}]^{1-}$, $[HO\text{-ferrihaem}]^{2-}$ and [μ-oxo ferrihaem]$^{4-}$ modelled with iron centres in the high-spin ($S = {}^5/_2$) state (broken symmetry in the case of the last) were used as starting points for force field parameterisation (see Chapter 4 details of DFT calculations). Mulliken charges were calculated for these structures employing the same functional and basis set combination as used in Chapter 4 (OPBE/LANL2DZ). A similar approach to that employed by Autenrieth *et al*. for six-coordinate haem species was adopted for five-coordinate ferrihaem species,[328] and involved reducing the complexity imposed by the asymmetric ferrihaem moiety by averaging Mulliken charges for equivalent atom types (see Fig. 5.1 in Section 5.2.1, above). Furthermore, in compliance with the CHARMM parameterisation protocol,[246, 248] all non-polar hydrogen atoms were given charges of 0.09e, with any residual charge being summed into the adjacent carbon atom.

Parameterisation of MM atomic charges (see Fig. 5.2B) was conducted using two criteria, (i) that MM atomic charges reproduced QM-derived ESPs; and (ii) that MM atomic charges reproduced WIEs between selected atoms and a water molecule (see Fig. 5.3). In the case of (i), the FitCharge module in CHARMM was used to fit QM ESPs generated using the Merz-Kollman scheme with the added constraint to reproduce the QM dipole. In the case of (ii), CHARMM parameterisation methodology suggests HF/6-31G* should be used to calculate energies, however, when this was employed for ferrihaem species, substantial spin-contamination was observed. As an alternative, the use of OPBE/6-31G* was explored. This functional was chosen because it successfully modelled ferrihaem species in Chapter 4. The LANL2DZ basis set employed in Chapter 4 was not used since the CHARMM parameterisation protocol recommended 6-31G* as the preferred basis set. While the latter basis set is larger than the former, it could be used because only hydrogen-bonding distances were optimised in these calculations and thus did not require long computing times. Usable interaction energies were obtained only in the case of

monomeric species. In an attempt to retain compatibility with previously developed CHARMM force fields, WIEs calculated using OPBE/6-31G* were scaled by a factor of 1.12. This was the value calculated to scale the interaction energy of water dimer computed using OPBE/6-31G* to the energy determined using HF/6-31G*. WIEs of $[H_2O\text{-ferrihaem}]^0$ were scaled by a further 1.16, as suggested by the CHARMM parameterisation protocol for neutral species. Despite the use of OPBE/6-31G* to calculate WIEs, spin-contamination was still obtained for [μ-oxo ferrihaem]$^{4-}$. Consequently, atomic charge parameterisation of this species could only be performed by fitting to the QM ESP. To aid this process, initial atomic charges were calculated from the ratio between optimised and Mulliken atomic charges for each atom type in $[HO\text{-ferrihaem}]^{2-}$ and its product with corresponding equivalent (or near equivalent) Mulliken charge determined for the μ-oxo dimer (see eq. 5.2.1 in Section 5.2.1). MM ESPs and WIEs calculated using optimised atomic charges are presented in Fig. 5.4 and Table 5.1, respectively. QM ESPs and WIEs are included for comparison.

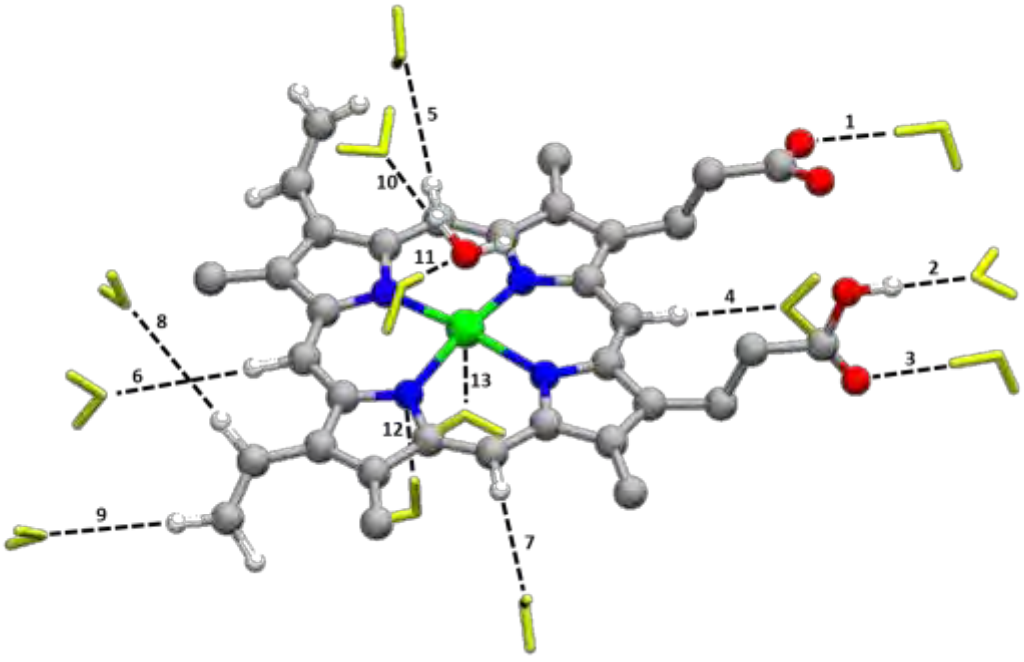


Figure 5.3. Water interactions with selected H_2O-ferrihaem atoms determined from DFT calculation (OPBE/6-31G*). Water molecules have been coloured yellow, interactions numbered and methyl and methylene hydrogen atoms omitted for clarity.

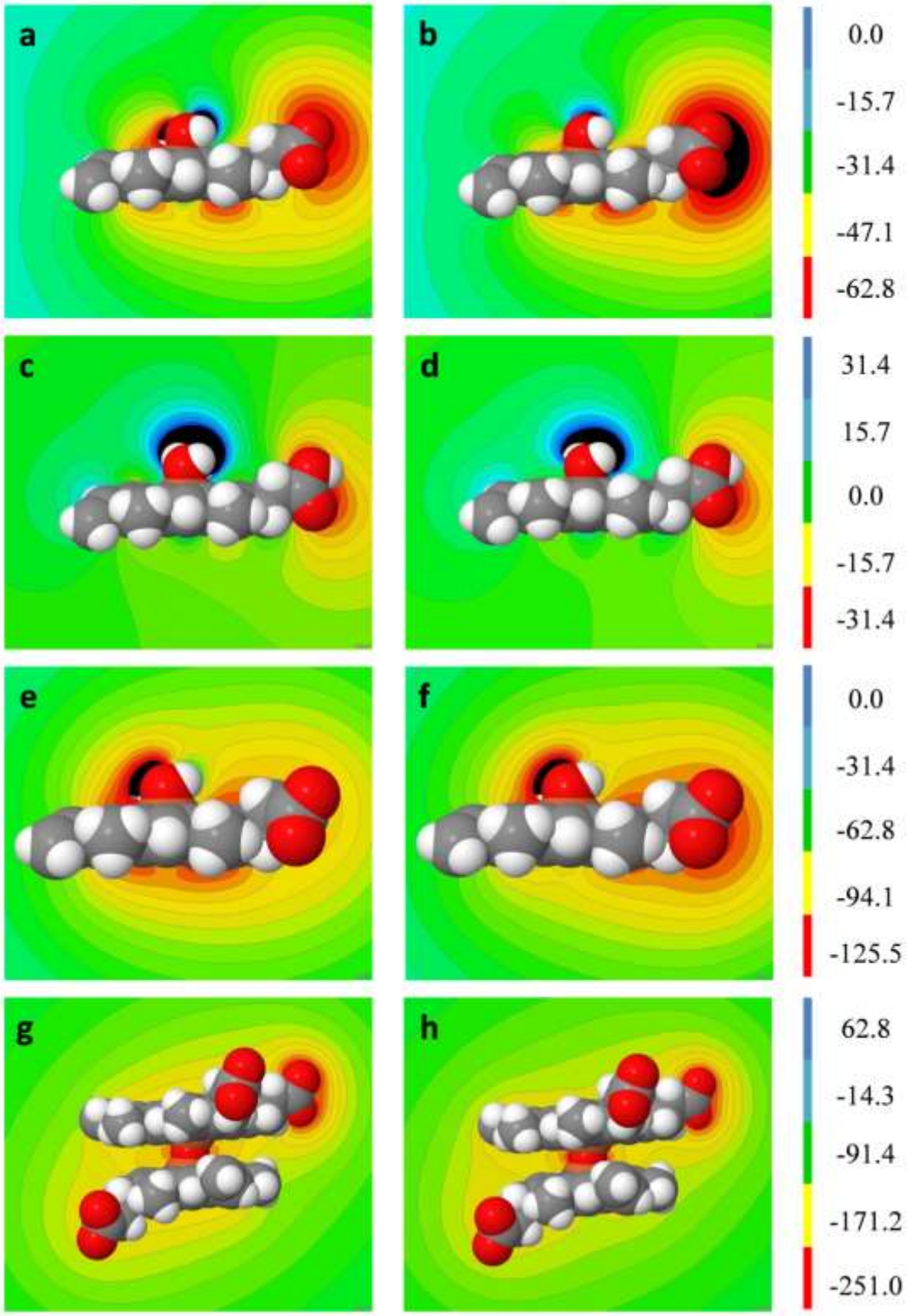


Figure 5.4. QM- (left) and MM-generated (right) ESPs of ferrihaem species. (**a, b**) $[H_2O\text{-ferrihaem}]^{1-}$; (**c, d**) $[H_2O\text{-ferrihaem}]^{0}$; (**e, f**) $[HO\text{-ferrihaem}]^{2-}$; (**g, h**) [μ-oxo ferrihaem]$^{4-}$. ESPs were generated using the Jmol software package. Energies displayed are in kcal mol^{-1}.

In general, ESPs generated using MM atomic charges were found to reproduce QM counterparts well. Good agreement was particularly observed for $[H_2O\text{-ferrihaem}]^0$ and [μ-oxo ferrihaem]$^{4-}$ species. In the case of $[H_2O\text{-ferrihaem}]^{1-}$, only minor deviation around the axial water oxygen was observed, while slightly poorer agreement was obtained for $[HO\text{-ferrihaem}]^{2-}$, specifically around the unligated face and propionate side chains. In the case of the latter ferrihaem species, no further adjustment was possible to these areas without causing significant deviation to between QM and MM WIEs. Decreasing the negative potential surrounding the propionate side chains to better fit the QM ESP resulted in weaker WIEs with the propionate groups, while increasing the negative potential of the unligated face produced interaction energies with nitrogen atoms that were too strong. Consequently, the MM ESP presented in Fig. 5.4f was produced from atomic charges that gave the best agreement to both QM ESP WIEs.

Table 5.1. WIEs (kcal mol^{-1}) for monomeric ferrihaem species calculated using optimised atomic charges in CHARMM and using DFT (OPBE/6-31G*).

	Interaction Type [a]	**Water Interaction Energy** $[H_2O\text{-Ferrihaem}]^{0}$ [b]		$[H_2O\text{-Ferrihaem}]^{1-}$ [c]		$[HO\text{-Ferrihaem}]^{2-}$ [c]	
		OPBE/6-31G*	**CHARMM**	**OPBE/6-31G***	**CHARMM**	**OPBE/6-31G***	**CHARMM**
1	COO^----H-OH	−7.41	−6.30	−10.42	−9.59	−16.30	−13.05
2	COOH --- OH_2	−7.64	−6.78	-		-	
3	C=O --- H-OH	−4.85	−5.00	-		-	
4	H_{meta} ---OH_2	*Spin Contaminated*		−3.00	−3.84	*Spin Contaminated*	
5	H_{meta} ---OH_2	−2.95	−2.61	−0.71	−1.03	*Spin Contaminated*	
6	H_{meta} ---OH_2	−2.98	−3.07	−0.77	−1.45	0.94	0.83
7	H_{meta} ---OH_2	−1.30	−0.97	0.74	0.61	2.11	2.49
8	H_{vinyl} --- OH_2	−2.34	−2.79	−0.75	−1.74	−0.75	0.32
9	H_{vinyl} --- OH_2	−1.81	−1.44	−0.17	−0.14	0.97	1.35
10	O-H_{ax} --- OH_2	−11.45	−10.22	−6.87	−7.29	−1.95	−1.61
11	O_{ax} --- H-OH	−2.48	−1.88	−2.55	−3.70	−8.47	−8.88
12	N --- H-OH	−2.41	−3.44	−2.87	−3.89	−3.25	−4.51
13	Fe --- OH_2	−3.58	−4.17	−2.25	−3.30	2.77	1.41
	MUE	*0.60*		*0.68*		*0.91*	

[a] See figure 5.2 for numbering; [b] QM interaction energy scaled by 1.12 × 1.16 = 1.30; [c] QM interaction energy scaled by 1.12.

QM WIEs were generally well reproduced by MM atomic charges. This is particularly evident from the plot of QM vs. MM interaction energies (Fig. 5.5) which produced a statistically significant linear correlation. Mean unsigned errors (MUE) of less than 1 kcal mol^{-1} were obtained and, with the exception of the propionate-water interaction of $[HO\text{-ferrihaem}]^{2-}$, all deviations were below 1.4 kcal mol^{-1}. However, despite underestimation of the MM propionate-water energy by approx. 3.3 kcal mol^{-1}, it remains the strongest interaction energy for this species and follows the same trend as that observed for the QM interaction energies.

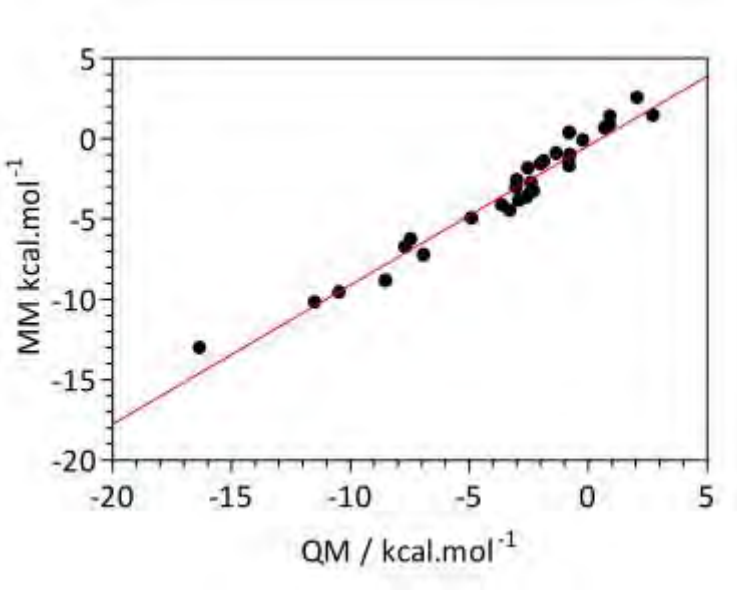


Figure 5.5. WIEs with selected atoms of monomeric ferrihaem species calculated using QM (OPBE/6-31G*) and CHARMM from optimised atomic charges. The linear regression is described by the equation y = 0.86(3)x – 0.4(2), r^2 = 0.96.

Final optimised atomic charges for all species are listed in Table 5.2. Dipole moments calculated from DFT and CHARMM have also been included. Atom labels are illustrated in Fig. 5.6.

Table 5.2. Optimised atomic charges and calculated dipole moments using QM and CHARMM for ferrihaem species. [a]

Atom type	$[HO\text{-ferrihaem}]^{2-}$	$[H_2O\text{-ferrihaem}]^{1-}$	$[H_2O\text{-ferrihaem}]^{0}$	$[\mu\text{-oxo ferrihaem}]^{4-}$
H_{ax}	0.302	0.344	0.413	–
O_{ax}	−0.806	−0.656	−0.707	−1.528
Fe	1.260	1.056	0.987	1.176
N	−0.518	−0.638	−0.568	−0.414
C_{α}	0.274	0.331	0.303	0.232
C_{β}	−0.017	−0.009	−0.003	−0.012
C_{meta}	−0.339	−0.352	−0.315	−0.266
H_{meta}	0.132	0.162	0.148	0.114
C_{methyl}	−0.332	−0.318	−0.293	−0.329
$C_{vinyl\text{-}1}$	−0.163	−0.184	−0.180	−0.133
$H_{vinyl\text{-}1}$	0.142	0.176	0.177	0.129
$C_{vinyl\text{-}2}$	−0.404	−0.446	−0.406	−0.343
$H_{vinyl\text{-}2}$	0.145	0.177	0.178	0.110
$C_{prop\text{-}1\text{-}unprot}$	−0.205	−0.200	−0.187	−0.202
$C_{prop\text{-}2\text{-}unprot}$	−0.248	−0.252	−0.213	−0.266
$C{=}O_{unprot}$	−0.591	−0.550	−0.493	−0.480
CO_{unprot}	0.578	0.712	0.672	0.359
$C_{prop\text{-}1\text{-}prot}$	–	–	−0.184	–
$C_{prop\text{-}2\text{-}prot}$	–	–	−0.173	–
$C{=}O_{prot}$	–	–	−0.425	–
CO_{prot}	–	–	0.566	–
$C\text{-}O_{prot}$	–	–	−0.536	–
$C\text{-}OH_{prot}$	–	–	0.378	–
Dipole (QM)	*42.6*	*40.5*	*15.0*	*16.6*
Dipole (CHARMM)	*43.0*	*40.0*	*15.1*	*16.6*

[a] Non-polar hydrogen atoms of methyl and propionate/propionic acid groups have a charge of 0.09 e.

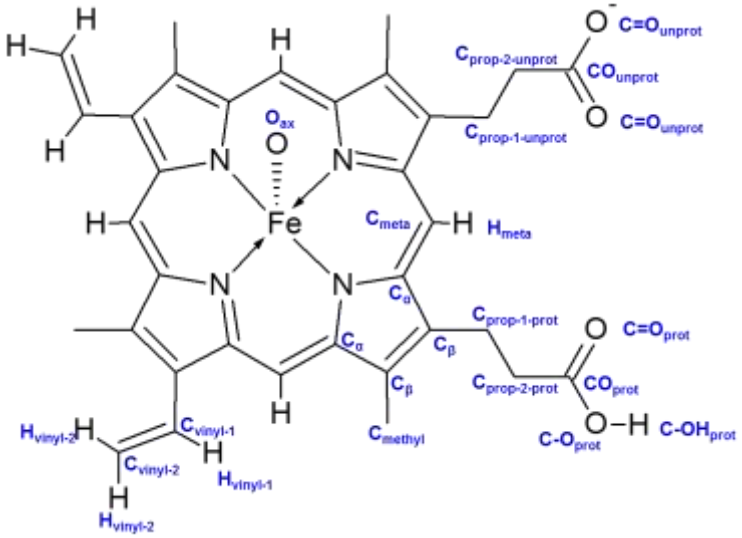


Figure 5.6. Atom labelling for ferrihaem species. Non-polar hydrogen atoms omitted for clarity.

5.3.2 Structural parameterisation.

The new atom types created for iron and ligand atoms required optimisation of bond, angle and improper dihedral parameters involving these atoms (see Section 2.7.1 for details regarding parameters). This included Fe-N and Fe-O_{ax} bonds; N-Fe-C_α, N-Fe-N, N-Fe-O_{ax}, Fe-O_{ax}-Fe ([μ-oxo ferrihaem]$^{4-}$) and Fe-O_{ax}-H_{ax} angles (H_2O- and HO-ferrihaem species); and the N*-C_α-C_α-Fe improper dihedral. Initial guess values for these parameters were taken from the closest related parameters in the CHARMM27 force field or from Autenrich *et al*. (see Table 5.3).[328] Parameters describing remaining atoms in the protoporphyrin IX moiety were retained as described in the CHARMM27 force field. To optimise values according to the CHARMM parameterisation methodology (see Fig. 5.2C), QM geometries and frequencies (MP2/6-31G*) were required as target data. Owing to the size of ferrihaem and the fact that porphyrin stretching frequencies couple to give complex vibrational modes, this approach could not be directly followed. Instead, target data were obtained from OPBE/6-31G* geometries and experimental crystallographic data of related Fe(III)porphyrins. In the case of [μ-oxo ferrihaem]$^{4-}$, the asymmetric υ(Fe-O-Fe) stretching frequency has been previously reported and thus also included as target data for this molecule.[64] Acceptable tolerances of bond and angle deviations between MM-minimised and QM or averaged crystal structure values were set to 0.03Å and 3° or less, respectively. Where possible, force constants were not adjusted.

Table 5.3. Initial guess and optimised structural parameters used for ferrihaem species.

	Initial Guess		Optimised	
Bonds:	K_b (kcal mol^{-1} Å^{-2})	b_0 (Å)	K_b (kcal mol^{-1} Å^{-2})	b_0 (Å)
Fe-N	270.2	2.00 [d]	270.2	2.18
Fe-OLW [a]	300.0	2.45 [d]	300.0	2.25
Fe-OLC [b]	250.0	1.80 [e]	250.0	1.93
Fe-OLH [c]	250.0	1.80 [e]	250.0	1.93
Angles:	K_θ (kcal mol^{-1} rad^{-2})	θ_0 (°)	K_θ (kcal mol^{-1} rad^{-2})	θ_0 (°)
N-Fe-OLW [a]	139.30	103.90 [e]	200.00	98.50
N-Fe-OLC [b]	139.30	103.90 [e]	155.00	107.00
N-Fe-OLH [c]	139.30	103.90 [e]	139.30	103.00
Fe-OLW-HLW [a]	33.00	133.00 [e]	33.00	130.00
Fe-OLC-HLW [b]	33.00	133.00 [e]	33.00	128.00
Fe-OLH-Fe [c]	160.00	130.60 [e]	200.00	180.00
Improper:	K_ψ (kcal mol^{-1} rad^{-2})	ψ_0 (°)	K_ψ (kcal mol^{-1} rad^{-2})	ψ_0 (°)
N*-C_α-C_α-Fe	137.40	0.00 [e]	18.30	0.00

[a] [H_2O-ferrihaem]$^{0/1-}$; [b] [HO-ferrihaem]$^{2-}$; [c] [μ-oxo ferrihaem]$^{4-}$; [d] From Autenrieth *et al*.[328]; [e] From CHARMM.

MM minimisation of ferrihaem molecules using initial guess parameters indicated that values used for N-Fe-N and N-Fe-C_α angles (taken from the HEME residue in CHARMM) were compatible with ferrihaem and thus no modification was required. The remaining bond lengths, angles and improper dihedral displayed significant deviation from QM and experimental values and thus indicated further optimisation of these parameters was necessary. In the case of Fe-N, Fe-OLC and Fe-OLH, only minor changes to the reference bond length, b_0, of no more than 0.2 Å produced values in good accord with both QM and experiment (see Table 5.3 and 5.4). Optimised parameters for Fe-OLH were also able to yield a MM υ(Fe-O-Fe) stretching frequency for [μ-oxo ferrihaem]$^{4-}$ in good agreement with experiment (884 vs. 880 cm^{-1}). Furthermore, the single Fe-N parameter used for the four ferrihaem molecules reproduced all target bond distances well. Satisfactory agreement of the Fe-O_{ax} bond with QM value for [H_2O-ferrihaem]^0 was also obtained by a similar decrease in b_0, but was unable to reproduce the QM value for the [H_2O-ferrihaem]$^{1-}$ species where a deviation of 0.1 Å was observed. This finding is not unexpected, however, given the large difference observed between QM and crystallographic values and suggested that the QM value was likely overestimated. Indeed, this bond length was also shown to decrease substantially upon protonation of propionate side chains (see Chapter 4). However, while the Fe-O_{ax} bond lengths for [H_2O-ferrihaem]$^{0/1-}$ were somewhat longer than the average values of five-coordinate iron(III)porphyrins in the CSD, they fell within the reported range (1.98 – 2.20 Å).

Similar to that found for Fe-O_{ax} parameters, only minor changes to the Fe-O_{ax}-H_{ax} reference angle, θ_0, were required to reproduce target values for [H_2O-ferrihaem]^0 and [HO-ferrihaem]$^{2-}$. Optimisation of N-Fe-O_{ax} angle parameters, however, posed more of a challenge because the previously optimised parameters in CHARMM described haem species in which the iron atom lay in the plane of the porphyrin ring. The planarity of the iron centre in these systems is routinely enforced by strong N*-Fe-C_α-C_α improper dihedral force constants and consequently N-Fe-X_{ax} angle force constants, K_θ, are weak. By contrast, iron atoms of the five-coordinate ferrihaem species used in this study deviated from the plane of porphyrin nitrogen atoms by between 0.2 and 0.6 Å. In order to replicate this out of plane deviation, a weaker N*-Fe-C_α-C_α improper force constant and stronger N-Fe-O_{ax} angle constants were required. Parameters relating to the former were taken from the C_α*-C_α-C_{meta}-N improper dihedral. In the case of the latter, new parameters were required for the three N-Fe-O_{ax} angles. None of the N-Fe-X parameters in the CHARMM force field used as initial guesses could reproduce target angles and thus parameters for the C_α-N-C_α angle were used instead as the initial guess since these contained the largest

K_θ value of all porphyrin parameters. In the case of [μ-oxo ferrihaem]$^{4-}$, only a minor change to θ_0 was required to produce N-Fe-O_{ax} angles close to the average crystal structure value and to nearly within one standard deviation of the QM value. Reproduction of target angles for [HO-ferrihaem]$^{2-}$ and [H_2O-ferrihaem]$^{0/1-}$ species required significant modification to K_θ and θ_0 values. Owing to the lack of vibrational data, the reliability of these values is uncertain; however, they are able to reproduce QM target values well. Similar difficulties were encountered when optimising the Fe-O-Fe angle parameters for [μ-oxo ferrihaem]$^{4-}$. A large K_θ value was necessary to ensure reproduction of the near-linear angle obtained from QM calculations and experiment. Selected bond lengths, angles and iron out of plane deviation for all ferrihaem MM-minimised structures obtained using optimised parameters are shown in Table 5.4.

Table 5.4. QM, MM-minimised and, where available, experimental, bond lengths, angles and iron out of plane deviation (OOP) for monomeric and μ-oxo dimeric ferrihaem species.

	Fe-O_{ax}	Fe-N	Fe-O_{ax}-X	N-Fe-O_{ax}	N-Fe-N	Fe-N-C_α	OOP [d]
[HO-ferrihaem]$^{2-}$							
DFT	1.88	2.13(2)	119.9 [a]	103.7(4)	86.8(6)	126.0(7)	0.5
CHARMM	1.88	2.097(3)	119.8 [a]	103.1(3)	87.1(4)	126.9(3)	0.5
Exp.	*1.85(2)*	*2.07(1)*	- [b]	*102.6(8)*	*87.3(4)*	*126.3(2)*	*0.45(3)*
[H_2O-ferrihaem]$^{1-}$							
DFT	2.28	2.077(6)	110(4) [a]	95(5)	89.6(5)	128.8(3)	0.2
CHARMM	2.18	2.069(2)	120.3(3) [a]	98.3(4)	88.8(3)	126.6(3)	0.3
[H_2O-ferrihaem]^{0}							
DFT	2.20	2.074(8)	121.6(2) [a]	96(2)	89.4(3)	126.6(3)	0.2
CHARMM	2.18	2.077(1)	121.1(2) [a]	98.2(2)	88.8(3)	126.8(2)	0.3
Exp.	*2.09(6)*	*2.00(4)*	- [b]	*93(3)*	*89.7(4)*	*126(1)*	*0.23(4)* [e]
[μ-Oxo ferrihaem]$^{4-}$							
DFT	1.80	2.13(1)	177.8 [c]	105(5)	86.1(4)	126(1)	0.6
CHARMM	1.78	2.085(2)	179.1 [c]	99.4(8)	88.5(3)	125.6(7)	0.3
Exp.	*1.76(1)*	*2.08(2)*	*176(3)* [c]	*103(2)*	*87(1)*	*125(3)*	*0.48(4)*

[a] X = H; [b] not available; [c] X = Fe; [d] Least squares plane through nitrogen atoms; [e] average of five-coordinate iron(III)porphyrins only.

5.4 MD Simulation

Eight ferrihaem species were modelled in aqueous solution using the TIP4P-Ew water model at 298 K. These included the three monomeric species $[HO\text{-ferrihaem}]^{2-}$, $[H_2O\text{-ferrihaem}]^{1-}$, and $[H_2O\text{-ferrihaem}]^{0}$; [μ-oxo ferrihaem]$^{4-}$ dimer and four π-π dimeric species $[(HO\text{-ferrihaem})_2]^{4-}$; $[(H_2O\text{-ferrihaem})(HO\text{-ferrihaem})]^{3-}$; $[(H_2O\text{-ferrihaem})_2]^{2-}$ and $[(H_2O\text{-ferrihaem})_2]^{0}$. Production dynamics were recorded for 10 ns for each species from which diffusion, solvation and structural analyses were performed.

5.4.1 Monomeric ferrihaem species.

Aqueous solvation of monomeric ferrihaem species was probed using SDFs. Only solvation of the porphyrin ring is presented since the flexibility of the propionate side-chains caused smearing of solvent atoms to occur in the SDF calculation. Three regions of high water oxygen atom probability (50% greater probability than bulk) were observed around (i) the axial water and hydroxide ligands, (ii) hydrogen atoms at the meso positions of the porphine moiety, and (iii) the unligated face (see Fig. 5.7). These findings were similar to SDFs reported by Cuya Guizado *et al*. for ferrous haem.[343] Given the strong hydrogen-bonding capabilities of the axial water and hydroxide ligands, solvation of this area is unsurprising. The cage-like structure produced by solvation of H_{meso} atoms was similar to that previously reported for aromatic molecules such as benzene,[344] but was less complete between propionate side chains. This likely resulted from the flexibility of the side chains which may have reduced the capacity of water to interact with the H_{meso} atom. Interestingly, the unligated face, which has been thought to be hydrophobic in nature, was considerably solvated by water molecules. This solvation presented in two forms. The first was observed for $[H_2O\text{-ferrihaem}]^{0}$ and $[HO\text{-ferrihaem}]^{2-}$ where a single water molecule was orientated with its oxygen atom near the sixth coordination site of the iron centre. The points of highest probability at these positions were 3.1 and 3.0 Å from the iron, respectively. While MD simulation is unable to describe possible bond formation, the high probability of finding a water molecule near the sixth coordination site may indicate there is an equilibrium involving a six-coordinate species in aqueous solution.

The second solvation type was observed in $[H_2O\text{-ferrihaem}]^{1-}$ where hydrogen atoms of water molecules orientated towards the negative potential of the pyrrole nitrogen atoms. Given the resemblance in QM ESPs for $[H_2O\text{-ferrihaem}]^{1-}$ and $[HO\text{-ferrihaem}]^{2-}$, solvation of the unligated face of these two molecules was expected to be similar. The difference likely occurred owing to the difficulties experienced in reproducing the ESP of the unligated face of $[HO\text{-ferrihaem}]^{2-}$. Thus the solvation model of this species is probably less definitive than the other two.

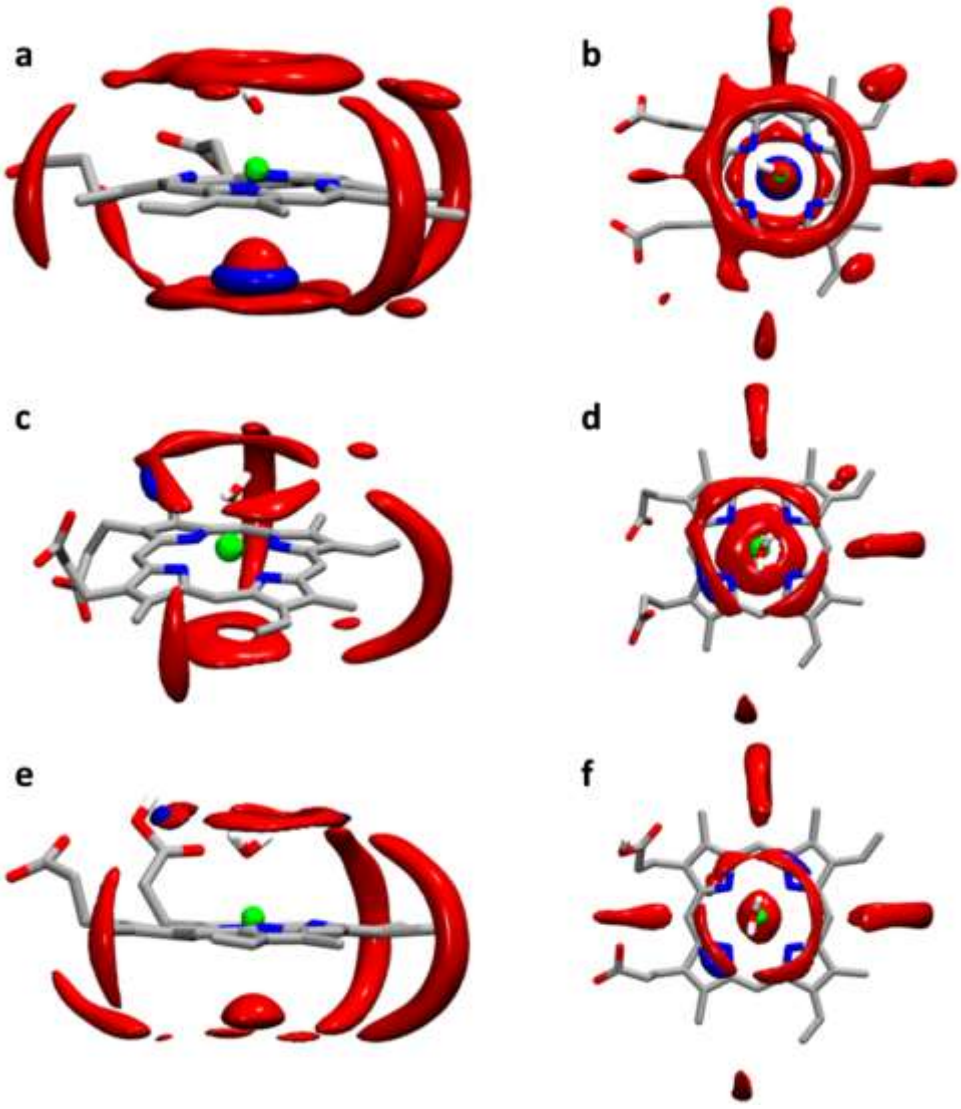


Figure 5.7. Spatial distribution functions of monomeric ferrihaem species showing sites of water oxygen (red) and hydrogen (blue) solvation at 50% greater density than bulk. (**a, b**) $[HO\text{-ferrihaem}]^{2-}$; (**c, d**) $[H_2O\text{-ferrihaem}]^{1-}$; (**e, f**) $[H_2O\text{-ferrihaem}]^{0}$.

5.4.2 π-π Dimeric ferrihaem.

Since the dominant form of ferrihaem in purely aqueous solution has been identified as the π-π dimer,[75] probing the solution behaviour of these species was of particular interest. The proposed formation of π-π dimeric ferrihaem species in aqueous solution was supported by the results of this study where dimers were found to remain associated over the entire course of the MD simulation. This finding included π-π dimeric $[(H_2O\text{-ferrihaem})_2]^0$ over which there has been some controversy in the literature. Crespo *et al*. previously suggested that this species was unlikely to exist in aqueous solution because the formal positive charge on the iron(III)porphine core would cause repulsive interactions between the two monomers and thus encourage dissociation.[163] The simulations in this work do not support the conclusion by these authors. While they are correct that the formal charge on the porphyrin core is indeed +1, the real charge differs because charge delocalisation is possible. This was demonstrated by DFT calculations where summation of Mulliken charges of the iron(III)porphine core of neutral H_2O-ferrihaem produced a charge close to +0.5 rather than +1. Furthermore, summation of Mulliken charges on the propionate side chain was approx. −0.3 rather than −1, while the charge on the propionic acid group was nearly 0. These calculations suggested delocalisation of charge from propionate side chain to porphyrin core occurs. It is the reduction of charge in this region which likely permitted the observed π-π dimer association in MD simulation.

Previously, MD simulation of $[(H_2O\text{-ferrihaem})_2]^0$ was conducted by de Villiers *et al*.[75] in implicit solvent using the HYPERCHEM software package using force fields based on the work of Allinger and extended by Marques *et al*. for porphyrin systems.[79, 80] In this study, co-planarity was enforced between the two porphyrin rings by means of dummy atoms and large force constants. By contrast, the current study used explicit water molecules, the newly developed CHARMM ferrihaem force field and no structural constraints. This approach had the advantage of incorporating explicit water molecules and thus allowed intermolecular hydrogen bond formation between solute and solvent molecules, which is not possible with implicit solvent. This provided a more realistic model of aqueous media. In addition, the force field employed in the HYPERCHEM simulation did not include atomic charges and thus electrostatic interactions were better modelled with the CHARMM force field. Lastly, by allowing free movement of porphyrin rings, an unbiased description of the dynamic structure was obtained.

Analysis of the dynamic structure of ferrihaem π-π dimers was approached by determining vertical interplanar distance, lateral porphyrin shifts, interplanar angle and relative torsion of porphyrin rings with respect to each other. The average interplanar distance between porphyrin rings (3.4 ± 0.1 Å) agreed well with the average separation reported by Scheidt *et al*. of 3.46 ±0.07 Å, determined from crystal structures of π-stacked iron(III)porphyrin systems.[345] Interplanar angles of all species were essentially co-planar with approx. 80% of the 200 000 structures recorded over the course of the simulation having angles of less than 5° (see Fig. 5.8a, c, e and g). Lateral shifts between porphyrin rings were calculated from the COM of the iron(III)porphine core, which closely corresponded to the iron atom. Consequently, this can be thought of as the lateral displacement between iron atoms. In all cases, ferrihaem species had lateral shifts around 2 Å (see Fig. 5.8b, d, f and h). Average values for $[(HO\text{-ferrihaem})_2]^{4-}$, $[(H_2O\text{-ferrihaem})(HO\text{-ferrihaem})]^{3-}$, $[(H_2O\text{-ferrihaem})_2]^{2-}$ and $[(H_2O\text{-ferrihaem})_2]^0$ were 1.8 ± 0.7, 1.7 ±0.7, 2.1 ± 0.8 and 1.7 ± 0.6 Å respectively. These values are somewhat smaller than that reported by de Villiers *et al*. for the $[(H_2O\text{-ferrihaem})_2]^0$ species modelled using HYPERCHEM where the majority of lateral shifts were instead found between 2.5 and 3.0 Å.[75]

Dihedral angles between porphyrin rings were measured by defining a dihedral involving meso carbon atoms positioned between propionate side chains and the COM of the porphine moiety, $C_{meso,A}$-COM_A-COM_B-$C_{meso,B}$. Conformations in which porphyrins are orientated with propionate groups anti to each other had a dihedral value of 180°, while eclipsed conformations had a value of 0°. Interestingly, the anti conformation was seldom observed in any of the π-π dimeric ferrihaem species studied, while dihedral angles of ±60° and to a lesser extent ±120° were instead favoured (see Fig. 5.9 a, c, e and g). As indicated from the dihedral angle time series in Fig. 5.9b, d, f and h, conformations remained in the particular orientations for lengthy periods before suddenly switching to new ones. By contrast, dominant conformations identified by de Villiers *et al*. involved porphyrins rings orientated with angles of ±135 or 180°.[75] In vacuum, the conformations observed by de Villiers *et al*. are expected since the negative repulsion of the propionate side chains are minimised. In solution, however, solvent molecules dampen the negative potential on these groups and thus allow for closer the orientations observed in the current study to occur. In addition, since the propionate side chains are strongly solvated, the possibility of solvent sharing between propionate groups of the two porphyrin rings is possible. This phenomenon would likely minimise the decrease in the entropy associated with the ordering of water molecules around the propionate and propionic acid groups. Since the entropy gain can only be

modelled using explicit water molecules, the anti conformations observed by de Villiers *et al*. using an implicit water model are understandable.

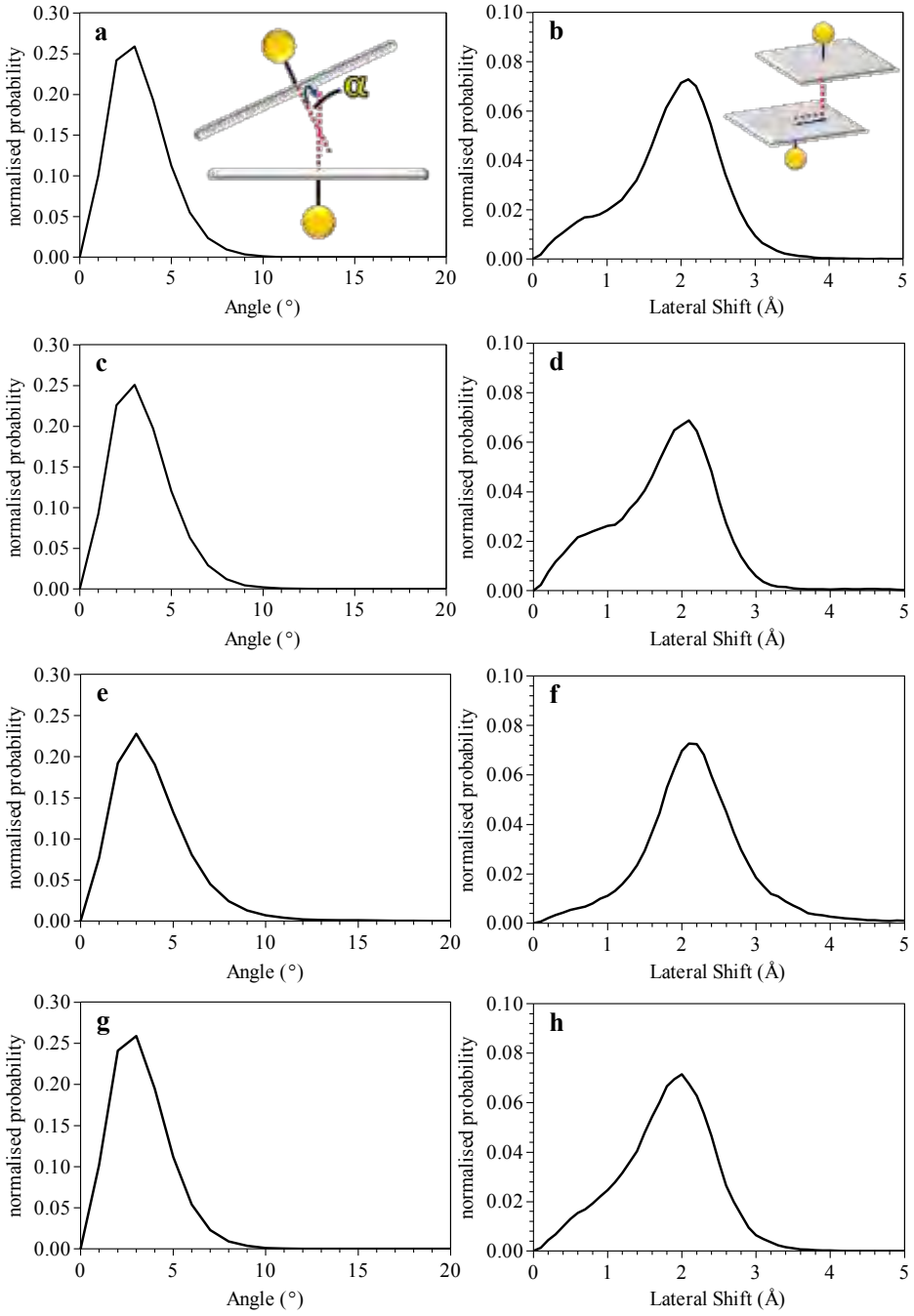


Figure 5.8. Interplanar angle (left) and lateral shift (right) histograms calculated from MD simulation of π-π dimeric (**a, b**) $[(HO\text{-ferrihaem})_2]^{4-}$; (**c, d**) $[(H_2O\text{-ferrihaem})(HO\text{-ferrihaem})]^{3-}$; (**e, f**) $[(H_2O\text{-ferrihaem})_2]^{2-}$ and (**g, h**) $[(H_2O\text{-ferrihaem})_2]^{0}$.

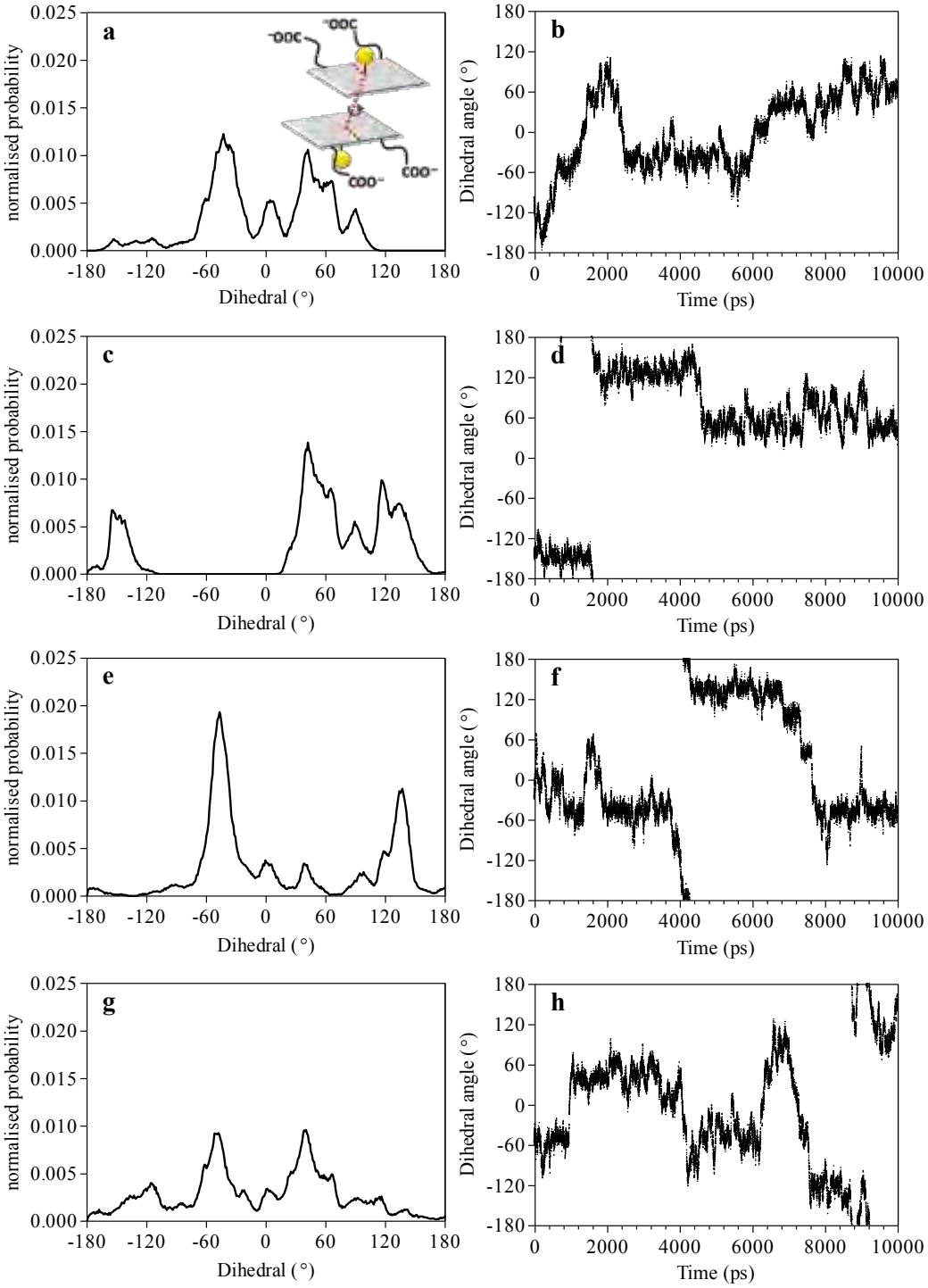


Figure 5.9. Dihedral angle histograms (left) and time series (right) for π-π dimeric (**a, b**) $[(HO\text{-ferrihaem})_2]^{4-}$; (**c, d**) $[(H_2O\text{-ferrihaem})(HO\text{-ferrihaem})]^{3-}$; (**e, f**) $[(H_2O\text{-ferrihaem})_2]^{2-}$ and (**g, h**) $[(H_2O\text{-ferrihaem})_2]^{0}$. Angles of 0 and 180 °C represent conformations where C_{meso} atoms positioned between propionate side chains on each porphyrin are eclipsed and anti to one another respectively.

Solvation of π-π dimeric ferrihaem species was similar to that observed for monomeric species (see Section 5.4.1) with the obvious exception of the unligated face (see Fig. 5.10). Formation of the π-π dimeric complex from ferrihaem monomers necessarily required desolvation of this face and thus suggested that the dimerisation process is entropy driven since water molecules were released into the bulk. Future experimental studies to investigate this observation are certainly warranted. Similar to monomeric species, axial ligands showed strong solvation but the cage-like structure produced around the H_{meso} positions were generally less well defined.

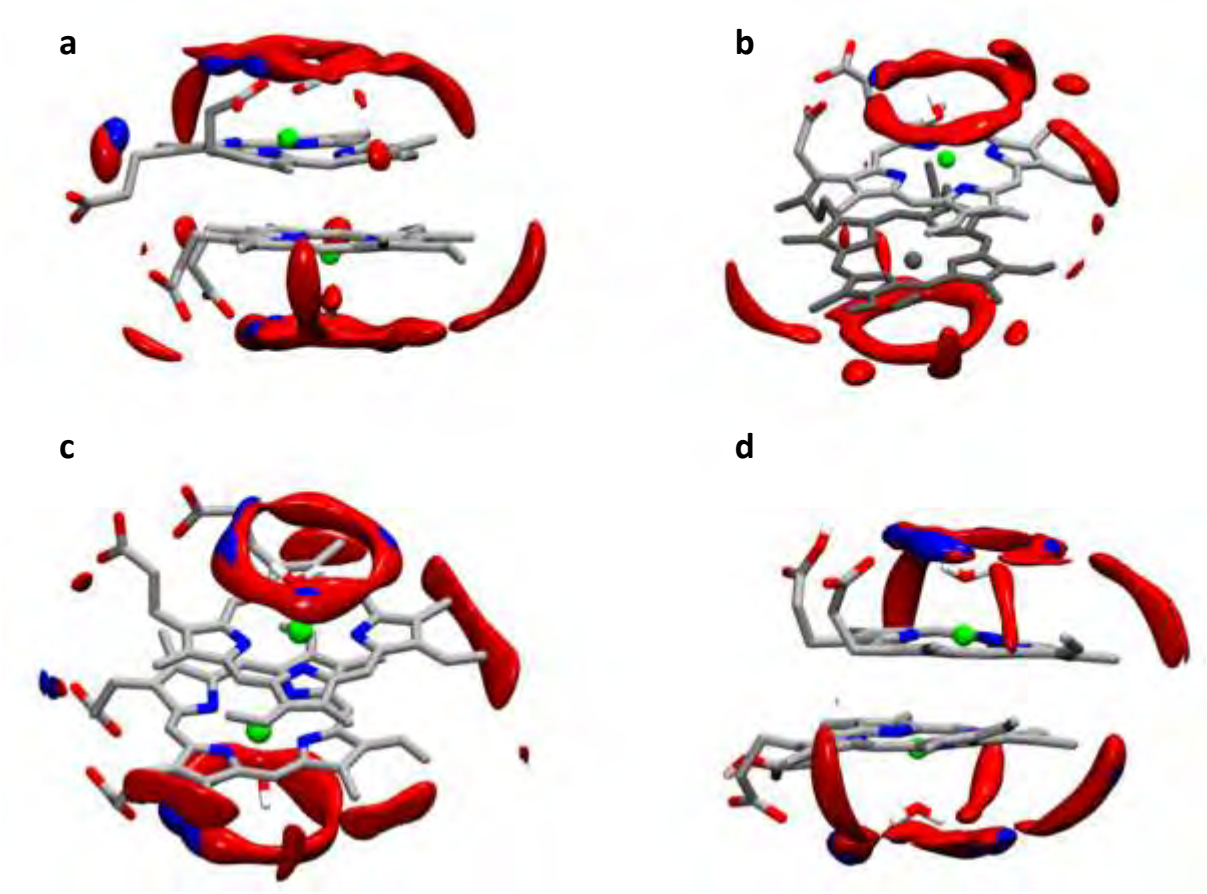


Figure 5.10. Spatial distribution functions of π-π dimeric ferrihaem species showing sites of water oxygen and hydrogen solvation (red and blue isosurfaces respectively) at 50% greater density than bulk. (**a**) $[(HO\text{-ferrihaem})_2]^{4-}$; (**b**) $[(H_2O\text{-ferrihaem})(HO\text{-ferrihaem})]^{3-}$; (**c**) $[(H_2O\text{-ferrihaem})_2]^{2-}$; (**d**) $[(H_2O\text{-ferrihaem})_2]^{0}$.

5.4.3 μ-Oxo ferrihaem.

The structure of [μ-oxo ferrihaem]$^{4-}$ in aqueous solution was markedly different to that observed for π-π dimeric species. The lateral shift for the μ-oxo dimer was smaller with average value of around 1.1 Å and had a narrower distribution of between about 0.6 and 1.6 Å (see Fig. 5.11a). This was not unexpected, however, given the constraining effect of the Fe-O-Fe linkage. The most striking differences between μ-oxo and π-π dimers were the larger interplanar angle of the former which averaged approx. 23° (Fig. 5.11b) and lack of evidence of rotation about the Fe-O_{oxo} bond (Fig. 5.11c and d). Both features appeared to arise from the presence of water between the porphyrin rings which caused the increase in interplanar angle and seemed to lock the porphyrin in the anti conformation (see Fig. 5.12a and b). The average Fe-O-Fe bond angle was approx. 168° and spanned a range of about 10° but was in good agreement with the observed angle in the μ-oxo ferrihaem dimethyl ester crystal structure of 170°.

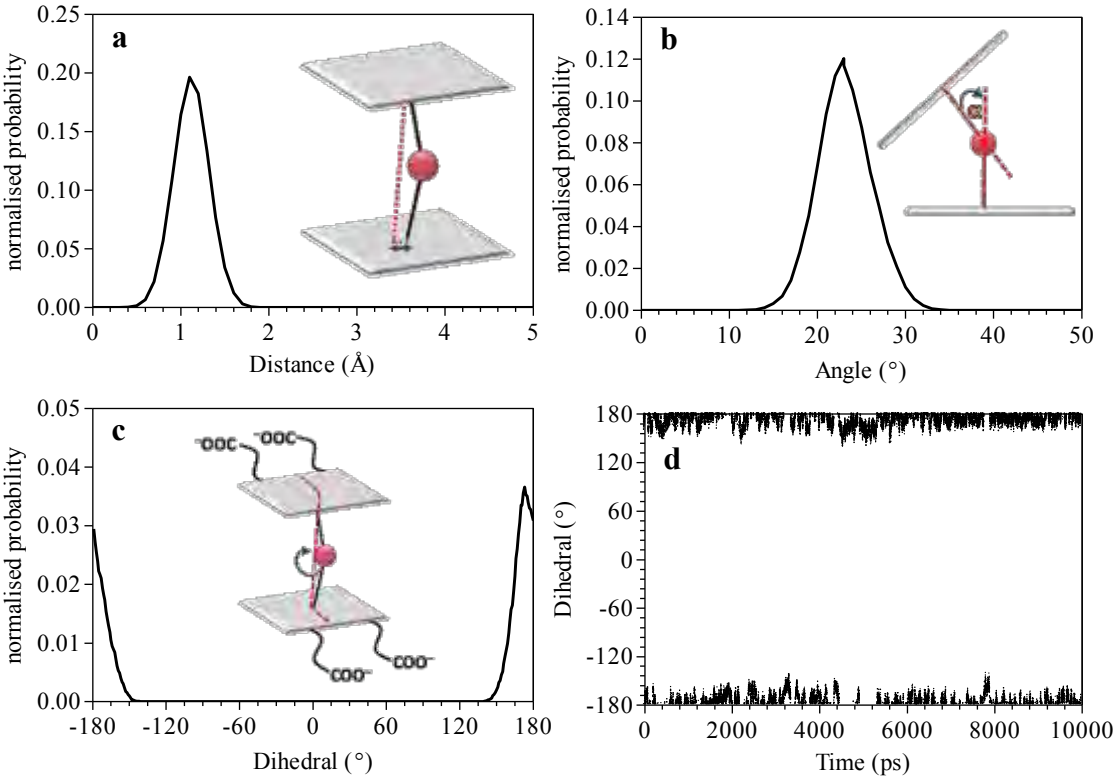


Figure 5.11. (**a**) Lateral shift; (**b**) interplanar angle; and (**c**) torsion angle histograms and (**d**) time series for [μ-oxo ferrihaem]$^{4-}$.

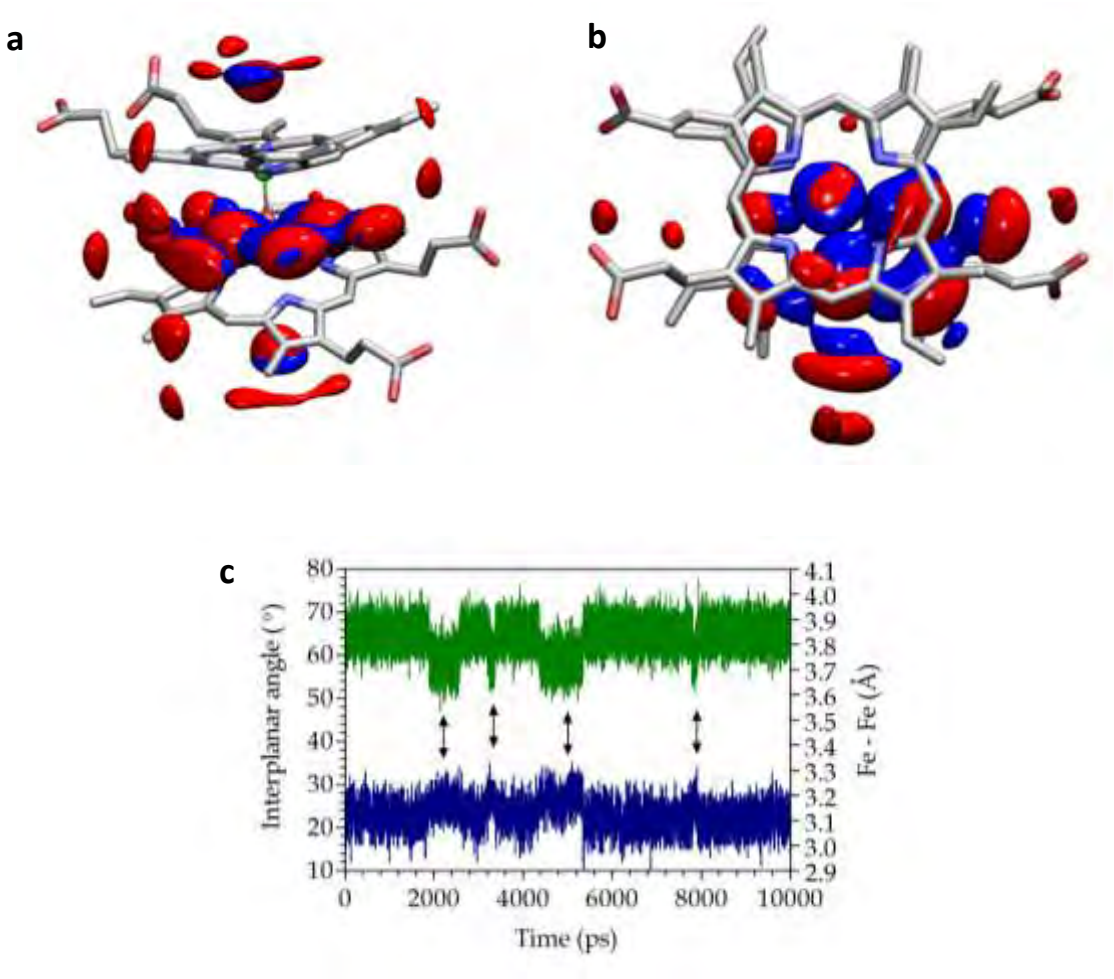


Figure 5.12. Spatial distribution functions of [μ-oxo ferrihaem]$^{4-}$ from (**a**) side and (**b**) top views, showing sites of water oxygen and hydrogen solvation (red and blue isosurfaces respectively) at 50% greater density than bulk. (**c**) Fe-Fe distance (green, plotted on right axis) and interplanar angle (blue, plotted on left axis) over the course of MD simulation. Arrows indicate structural changes which occur when only two water molecules interacted with the μ-oxo bridge.

As alluded to above, solvation of [μ-oxo ferrihaem]$^{4-}$ was considerably different to that observed for monomers and π-π dimers (see Fig. 5.12a and b). Marked solvation of the axial μ-oxo ligand was found where two water molecules strongly interact with the bridging ligand and a third was present at slightly lower probability. The loss of the third water molecule had a notable effect on interplanar angle and Fe-Fe bond distance. This can be clearly seen in Fig. 5.12c where a smaller angle and shorter Fe-Fe bond occurred. A second hydration shell surrounding that solvating the μ-oxo ligand was also found. While the solvation of this region was at first glance unexpected, it was unsurprising given the substantial negative potential around the μ-oxo bridge indicated by the ESP in Section 5.3.1. Further precedence for this type of structure can be seen from crystal structures of μ-oxo iron(III)porphyrin dimers which have acetonitrile and benzene molecules included in similar positions (see Fig. 5.13a and b).[71, 346] Minor solvation of the unligated face by a single water molecule, similar to that presented for monomeric [HO-ferrihaem]$^{2-}$ and [H_2O-ferrihaem]^{0}, was also observed. The hydration of this face thus does not support the idea that these surfaces are entirely hydrophobic.

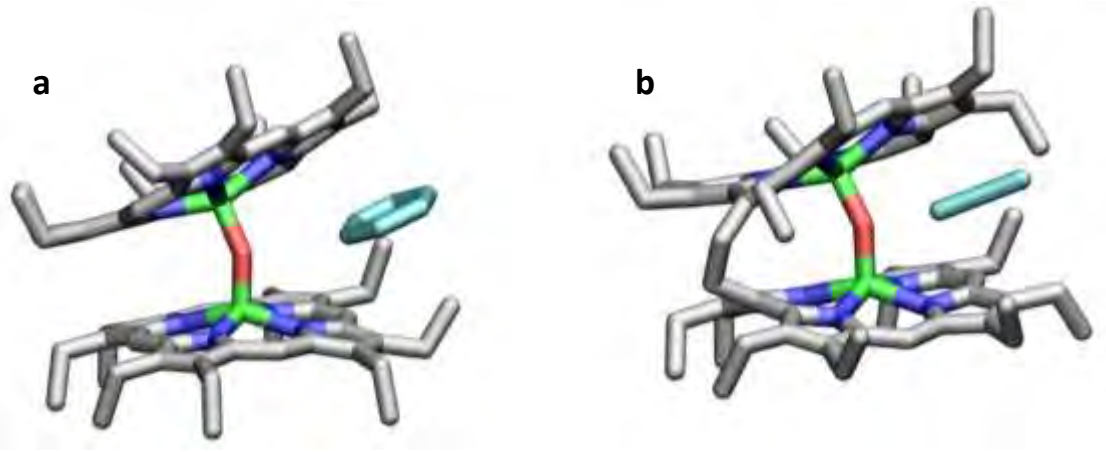


Figure 5.13. Crystal structures of (**a**) μ-oxo iron(III)OEP and (**b**) its tethered analogue reported by Lee *et al.*[71] and Ghosh *et al.*[346] respectively obtained from the CSD. Benzene and acetonitrile solvents (in a and b respectively) included between porphyrins rings are shown in cyan. The latter orientates with methyl end towards the μ-oxo bridge. Hydrogen atoms removed for clarity.

5.4.4 Effects of solvation on ferrihaem speciation.

Previously, it has been shown that solvation plays a crucial role in determining the speciation of ferrihaem.[76] Aqueous mixtures of water-miscible protic solvents were found to favour formation of π-π dimeric species, while aprotic solvents resulted in monomeric species at low pH but induced the μ-oxo dimer at higher pH. The results obtained from this MD study, specifically in the form of SDFs of monomeric and dimeric ferrihaem species, have provided a clearer indication as to why this occurs.

SDFs of monomers and π-π dimers showed prominent solvation by water oxygen atoms in Fig. 5.7 and 5.10, respectively. However, there was also considerable solvation by water hydrogen atoms which is visually obscured by the oxygen isosurfaces. Consequently, the marked solvation by both oxygen and hydrogen atoms of water indicated by the SDFs suggested that stabilisation of the π-π dimer occurs in protic solvents, particularly through strong solvation of the axial ligands. Changes to the protonation state of the axial ligand by varying pH can also be accommodated by such a solvent owing to the dual hydrogen bonding character. By contrast, aprotic solvents are not as versatile. At low to neutral pH, monomers most probably have water as the axial ligand and are likely solvated in a similar manner as shown for water oxygen atom isosurfaces in monomeric ferrihaem SDFs. As the pH becomes more alkaline, hydroxide ligands begin to dominate which have a reduced capacity to act as hydrogen bond donors. Under such conditions, as suggested by Asher *et al*.[76], the water concentration around the axial ligand will be increased relative to the bulk in order to satisfy the hydrogen bonding requirements of the hydroxide ligand. Consequently, this will cause a decrease in the overall entropy. The formation of the μ-oxo dimer is thus favoured presumably through the release of one water ligand, two hydrogen ions and the solvation shell surrounding the axial ligand into the bulk. While the strong solvation surrounding the μ-oxo bridge indicated by the SDF of [μ-oxo ferrihaem]$^{4-}$ may at first glance appear to contradict the previous statement, the fact that this solvation is shared between the oxide ligand of two porphyrins causes a drastic reduction in the overall degree of hydration of the ligated faces and thus would be entropically favoured. Indeed, induction of the μ-oxo dimer under these conditions has been shown experimentally to cause an increase in entropy.[76]

5.5 Experimental Support for MD Results

5.5.1 EXAFS.

EXAFS was used to provide experimental support for the proposed solution structure of [μ-oxo ferrihaem]$^{4-}$. This technique could only be applied to the μ-oxo dimer since experimental conditions required concentrations of approx. 5 mM and thus precluded measurement of monomeric species. Furthermore, the cryoprotectant PEG400 employed to prevent glassing at the low temperature (15 K) required to limit thermal motion was found to induce μ-oxo dimer formation. This was uncovered using magnetic susceptibility measurements which produced a magnetic moment of 4.2 μ_B under the experimental conditions used for EXAFS measurement (20% aq. PEG400, pH 7.4). Using the values determined in Chapter 3 for π-π dimeric and μ-oxo dimeric ferrihaem (4.8 and 2.0 μ_B, respectively), approx. 20% μ-oxo ferrihaem was expected to be present. Performing magnetic susceptibility measurements using experimental conditions similar to those employed by Asher *et al*. (40% alkaline aq. PEG400 solution) more clearly demonstrated that PEG400 induced μ-oxo dimer formation where a magnetic moment of 2.13 μ_B was obtained. The use of an alternative cryoprotectant in the form of glycerol was investigated, however, this solvent caused precipitation of ferrihaem from solution. For the sake of comparison, the EXAFS spectrum of [μ-oxo ferrihaem]$^{4-}$ in the solid phase was also recorded.

The EXAFS spectrum of the frozen [μ-oxo ferrihaem]$^{4-}$ solution was considerably different to that of the dried solid (see Fig. 5.14a). The spectrum of the latter was well fitted using only the reported crystal structure of μ-oxo ferrihaem dimethyl ester,[72] in which no solvent atoms were present (Fig. 5.14b). By contrast, a good fit to the spectrum of the frozen solution using the structure determined from MD simulation was only possible when five hydrating water molecules were included (see Fig. 5.14c and d). These five water molecules consisted of four surrounding the oxide bridge and one on the unligated face.

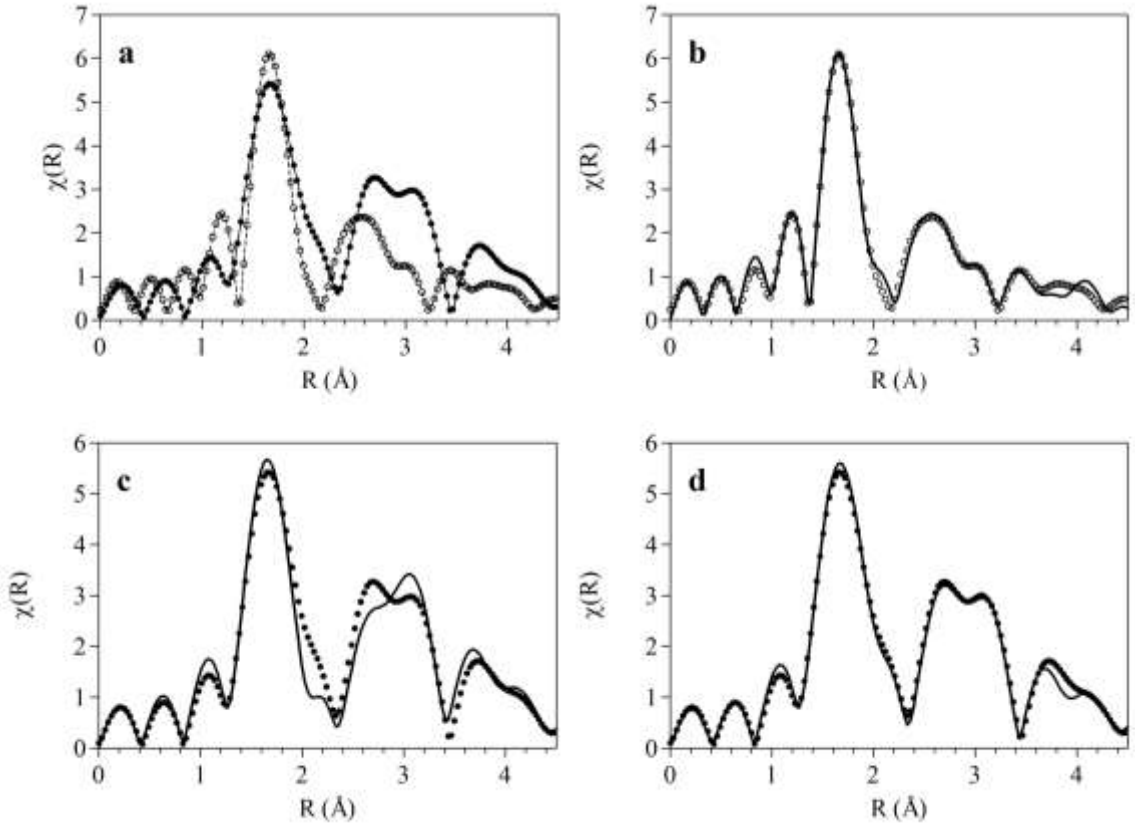


Figure 5.14. Experimental and fitted EXAFS spectra of [μ-oxo ferrihaem]$^{4-}$. (**a**) Comparison of spectra obtained from frozen solution containing 20% (v/v) aq. PEG400, 40% (v/v) DMSO, pH 10 (filled circles) and from the solid precipitated from aqueous acetone solution (unfilled circles). (**b**) Spectrum recorded on the solid (unfilled circles) fitted using the crystal structure of μ-oxo ferrihaem dimethyl ester (solid line).[72] (**c**) Spectrum recorded on the frozen solid (filled circles) fitted using the structure that corresponded most closely to the average structure in MD simulation but excluding water molecules. (**d**) Spectrum recorded on the frozen solid (filled circles) fitted with the same structure in (c) but including water molecules identified from the SDF (solid line).

The EXAFS spectrum in Fig. 5.14d was fitted using the structure corresponding closest to the average over 10 ns of MD simulation as the input geometry. Furthermore, water molecules were included from two regions, surrounding the μ-oxo bridge and the unligated face. Positions of these were determined by finding the point of highest probability in the SDF of [μ-oxo ferrihaem]$^{4-}$. The EXAFS spectrum was fitted in a similar approach to that described by Karolewski *et al.*[341] and Provost *et al.*[342]

This involved using MD simulation data to approximate the first- and second-order cumulants, ΔR and σ^2, respectively (see Section 5.2.3 for details). Scaling factors, α and β, were included as part of the fitting process in order to account for inaccuracies in experimental measurement, intrinsic errors in approximating cumulants and systematic differences between MD and EXAFS cumulants.[341] The same scaling factors were applied regardless of the path wherever possible. In this case, five α and two β scaling factors were necessary to produce good fits to the EXAFS spectrum. The five α parameters included two describing water molecules around the μ-oxo ligand and unligated face, $\alpha(1)_{wat}$ and $\alpha(2)_{wat}$ respectively; two describing the axial oxide ligand and iron centre of the adjacent porphyrin, α_{oxo} and α_{Fe} respectively; and one describing all other atoms. The two β parameters included one describing the oxide ligand, β_{oxo}, and the other accounted for all other atoms. Optimised parameters from this fit and that obtained without water molecules are listed in Table 5.5. Those determined for the solid EXAFS fit are also included for comparison. Additional statistics have been collected in the Appendix B Table B2 and B3.

Table 5.5. *R*-factor and scaling parameters for EXAFS spectra of [μ-oxo ferrihaem]$^{4-}$.[a]

Parameter	Fitted value		
	Solution [b]	Solution [c]	Solid [d]
R-factor	0.005	0.069	0.009
α	–0.012(2)	–0.004(4)	– [e]
α_{oxo}	0.029(7)	0.05(1)	– [e]
α_{Fe}	–0.05(1)	–0.02(5)	– [e]
$\alpha(1)_{wat}$	–0.174(3)	– [e]	– [e]
$\alpha(2)_{wat}$	0.072(4)	– [e]	– [e]
β	1.3(4)	0.7(6)	– [e]
β_{oxo}	2.5(8)	2(1)	– [e]

[a] Error given in parenthesis; [b] Fitted using the MD data including five water molecules at positions identified from the SDF; [c] Fitted using the MD data without solvent water molecules; [d] Fitted using the crystal structure of μ-oxo ferrihaem dimethyl ester[72]; [e] not applicable.

In general, interatomic distances determined from the fits in Fig. 5.14b and d compared well with the corresponding distances determined from the crystal structure of μ-oxo ferrihaem dimethyl ester[72] and average distances calculated from MD simulation, respectively (see Table 5.6). With the

exception of the Fe-Fe distance fitted for the frozen [μ-oxo ferrihaem]$^{4-}$ solution, all distances agreed within two standard deviations. While the fitted Fe-Fe distance was considerably shorter than the average value from MD simulation (3.56 vs. 3.83 Å), structures with this distance were indeed observed in the simulation and did not show significant changes in the other distances. The differences in temperature and solvent composition between MD simulation and experiment (298 K vs. 15 K and purely aqueous vs. aq. PEG400/DMSO) may have favoured this conformation.

Table 5.6. Comparison of fitted EXAFS interatomic distances and those obtained from the reported crystal structure of μ-oxo ferrihaem dimethyl ester[72] and from MD simulation of [μ-oxo ferrihaem]$^{4-}$ in aqueous solution. [a]

Distance	N	Solid		Frozen solution	
		EXAFS [b]	Crystal Str. [c]	EXAFS [d]	MD [e]
Fe–O_{oxo}	1	1.79(1)	1.748(3)	1.87(1)	1.92(4)
Fe–N	4	2.06(1)	2.08(1)	2.080(5)	2.11(3)
Fe–C_{α}	8	3.05(2)	3.08(2)	3.072(7)	3.12(4)
Fe–C_{β}	8	4.40(7)	4.30(2)	4.31(1)	4.37(5)
Fe–C_{meso}	4	3.4(1)	3.42(2)	3.388(8)	3.45(5)
Fe–Fe	1	3.5(9)	3.48	3.56(4)	3.83(7)
Fe–$O(1)_{wat}$ [f]	1	–	–	2.48(1)	3.31[h]
Fe–$O(2)_{wat}$ [g]	4	–	–	3.48(1)	3.25[h]

[a] Values are averages with standard deviations in parentheses for distances between Fe and all atoms of the type indicated. N is the number of Fe-X distances considered; [b] dried precipitate obtained from aqueous acetone solution; [c] μ-oxo ferrihaem dimethyl ester[72]; [d] frozen aqueous solution in 20% (v/v) aq. PEG400, 40% (v/v) DMSO, pH 10; [e] structure most closely resembling the average structure from 10 ns MD simulation in TIP4P-Ew water box, including five water molecules; [f] water molecule positioned on the unligated face of a given ferrihaem molecule; [g] water molecules surrounding the bridging oxide ligand; [h] initial water positions identified from the points of highest probability in the SDF.

Interestingly, the results of the EXAFS fit indicated that water molecules were found close to the sixth coordination site of iron centres of [μ-oxo ferrihaem]$^{4-}$. The fitted distance for these molecules fell approx. halfway between that expected for coordinated water (~2.1 Å) and for outer sphere solvation (~3.3 Å). This suggests that an equilibrium exists between five- and six-coordinate iron of [μ-oxo ferrihaem]$^{4-}$ where weak binding to the sixth coordination site occurs (see Fig. 5.15).

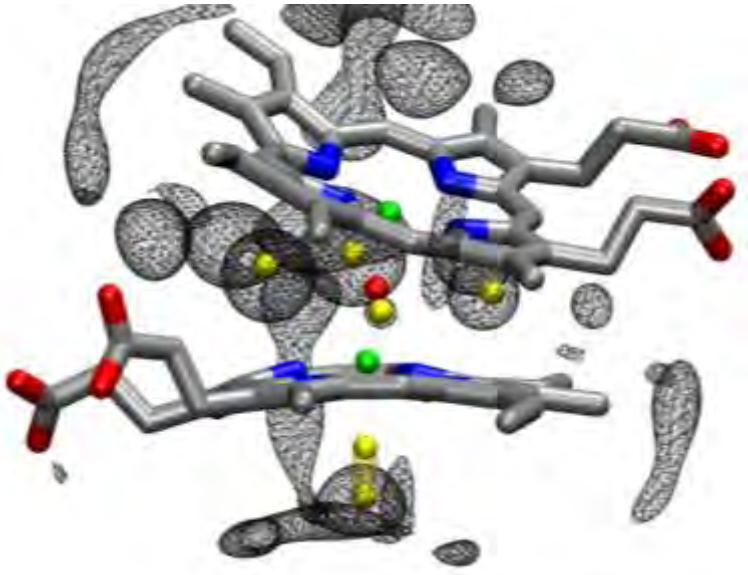


Figure 5.15. Aqueous solution structure of [μ-oxo ferrihaem]$^{4-}$ determined from MD simulation. Black wireframe represents SDF of water oxygen atoms at 50% greater density than bulk. Positions of water oxygen atoms calculated from fit to the EXAFS spectrum are shown in yellow and fall within SDF isosurfaces. The water oxygen atom on the unligated face was fitted at 2.48 Å from Fe and likely corresponds to an equilibrium between outer sphere solvation at 3.3 Å and coordinated water at 2.1 Å. This is indicated by two positions on the lower face. For clarity, solvation of the upper face is not shown.

Given that the SDFs of monomeric [HO-ferrihaem]$^{2-}$ and [H_2O-ferrihaem]^{0} showed similar strong solvation near the sixth coordinate site as seen in [μ-oxo ferrihaem]$^{4-}$, these former species may also exist in equilibrium between five- and six-coordinate states. This equilibrium can explain the experimental observations by Asher *et al.*[76] regarding the aggregation state of μ-oxo ferrihaem. In alkaline 40% aq. DMSO, the μ-oxo dimer does not form aggregates, presumably owing to the presence of the sixth ligand which prevents π-stacking interactions with other μ-oxo dimers. On the other hand, substantial aggregation was observed in aqueous solutions when employing high salt concentrations. These authors provided no specific explanation of why this occurred, however, this can now be rationalised in light of the EXAFS findings. Salts act as dehydrating agents as a result of strong water solvation and presumably remove the sixth axial ligand. This consequently facilitates π-stacking of the resulting unligated faces. The positions of water molecules surrounding the μ-oxo bridge produced fitted distances somewhat longer than the highest probability identified by SDF (3.48 vs. 3.25 Å). However, the fitted values were still well within the probability isosurface of the SDF (Fig. 5.15). The slightly longer fitted distance most likely occurred because this value represents the average of many structures in solution which have solvent molecules in several difference positions falling within the SDF isosurfaces.

5.5.2 Translational diffusion.

Translational diffusion coefficients, *D*, of all ferrihaem species were calculated from MD simulation using the MSD method described in Section 5.2.2 and are listed in Table 5.7. Only direct comparison with previously reported experimental values for the μ-oxo dimer and π-π dimeric [(HO-ferrihaem)$_2$]$^{2-}$ was possible.[75] The computed diffusion coefficient of the three monomeric species were compared to the experimental value measured for the bis-cyano-ferrihaem monomer, (CN)$_2$-ferrihaem. All calculated diffusion coefficients were overestimated as compared to their experimental counterparts, however, the trend across the species was reproduced. This is made clearer by normalising the average calculated diffusion coefficient of monomeric species to its experimental counterpart, where more reasonable agreement with experiment was observed. These relative values may thus be useful in aiding the identification of aggregation of ferrihaem species in aqueous solution when combined with experimental values.

Table 5.7. Experimental and computed diffusion coefficients (*D*) of selected ferrihaem species.

Ferrihaem Species	*D* ($\times 10^{-10}$ m^2 s^{-1})		
	Exp.	Calc.	Norm. Calc. [d]
Monomer	2.2(2) [a, b]	4.1(2) [c]	2.2
[HO-ferrihaem]$^{2-}$	-	3.9(3)	-
[H$_2$O-ferrihaem]$^{1-}$	-	4.1(2)	-
[H$_2$O-ferrihaem]^{0}	-	4.3(2)	-
[(HO- ferrihaem)$_2$]$^{4-}$	1.4(1) [b]	3.1(1)	1.7(1)
[(H$_2$O-ferrihaem)(HO-ferrihaem)]$^{3-}$	-	3.5(2)	-
[(H$_2$O-ferrihaem)$_2$]$^{2-}$	-	3.7(2)	-
[(H$_2$O-ferrihaem)$_2$]^{0}	-	3.5(2)	-
[μ-Oxo ferrihaem]$^{4-}$	1.6(1) [b]	3.0(2)	1.6(2)

[a] (CN)$_2$-ferrihaem; [b] from de Villiers *et al.*[75]; [c] average value computed for three monomeric species below; [d] normalized to the experimental monomer diffusion coefficient of (CN)$_2$-ferrihaem.

5.6 Conclusions

The focus of this study was to probe solution structures and solvation of ferrihaem species in aqueous medium using MD simulations supplemented by EXAFS data. To perform simulations representative of experimental conditions and to accurately describe free five-coordinate ferrihaem species, development of a new force field was crucial. Optimised atomic charges were able to generally reproduce QM electronic data well. More specifically, MM ESPs were in good agreement with QM counterparts, even after the starting structure was minimised (see Appendix B Fig. B1). MM WIEs obtained using both QM and MM-minimised geometries were comparable to QM values with MUE of 1.1 kcal mol^{-1} and below (see Appendix B Table B1 for MM-minimised values). Furthermore, MM structures were able to reproduce QM geometries with average MUEs in bond lengths of 0.03 Å, bond angles of 2° and iron out-of-plane deviations of 0.1 Å.

MD simulations strongly supported the feasibility of the formation of π-π dimers in aqueous solution in both the anionic and neutral state. These species were found to be essentially co-planar. Solvation of π-π dimers and monomers predominated around the axial ligand. Marked interaction of the unligated face with water molecules was also observed in the case of the latter species indicating this surface is not entirely hydrophobic and may in fact be weakly six-coordinate. Combining MD simulations with EXAFS data provided a useful method to elucidate the solution structure of [μ-oxo ferrihaem]$^{4-}$. Somewhat surprisingly, although not unreasonable, solvation of this species was found to dominate around the axial μ-oxo bridge. There was also evidence of weak coordination of water molecules to the sixth coordination site on the unligated face.

The findings from this work suggest the ferrihaem model established for MD simulation is robust and could be used to probe interactions with other molecules such as antimalarial drugs or in the presence of other substances, such as lipids or detergents.

CHAPTER 6

MOLECULAR DYNAMICS SIMULATION OF THE CQ-FERRIHAEM COMPLEX

6.1 Introduction

While a number of studies have employed computational techniques in an attempt to understand the interaction between ferrihaem and CQ or its analogues, the majority have made use of DFT calculations.[151, 175, 184, 186, 187] By contrast, relatively few investigations have been conducted using molecular mechanics or MD simulations. Unlike QM calculations which mainly involve determining properties for a single static structure, MD simulations are well suited to investigate the solution-state interaction of CQ and ferrihaem. This is because the time-dependent nature of the technique allows for equilibration of the system and provides information regarding the dynamic structure and solvation of the complex. The findings from these calculations are thus more representative of experimental conditions and can provide a clearer picture of the interaction of these compounds. Previous MD studies of antimalarial-ferrihaem complexes, specifically those relating to CQ and its analogues, have been reported by Marques *et al*.[165], Roepe and co-workers[172, 173, 347], Portela *et al*.[176] and Otelo *et al*.[348] These investigations, however, have been undertaken either using ferrihaem model compounds such as six-coordinate imidazole-iron(III)porphine,[165] or by simulating interactions with monomeric ferrihaem species such as HO-ferrihaem which are not immediately relevant to aqueous solution in which, as demonstrated in Chapter 3, CQ interacts with [μ-oxo ferrihaem]$^{4-}$.[348] On the other hand, those studies that used the μ-oxo dimer employed force fields that were not specifically developed for this ferrihaem species and did not report using implicit or explicit solvent in the calculation. Thus the accuracy of these calculations with regard to the solution-state interaction of CQ and ferrihaem is uncertain.

This chapter addresses the above-mentioned deficiencies of the previously reported studies by simulating the CQ-ferrihaem complex using the ferrihaem force field developed as described in Chapter 5. Since no force field exists for CQ, parameterisation of this molecule is also presented. MD simulation of CQ with ferrihaem in the μ-oxo dimeric form was conducted and the structural, conformational and solvation properties (in the form of SDFs) of the CQ-ferrihaem complex were determined. Computational data were directly compared with experiment. Specifically ^{1}H NMR T_1 relaxation measurements and EXAFS spectra were used to support the conclusions of the simulation. Finally, the aggregation state of the complex was determined using computed and experimental diffusion coefficients. The findings of this study have provided an experimentally supported model of the interaction of CQ and ferrihaem in aqueous solution.

6.2 Methods

6.2.1 Parameterisation procedure for CQ.

As suggested by the CHARMM parameterisation protocol,[246, 248] a fragment approach was used to create a model for CQ. This approach could be successfully applied to CQ but not ferrihaem species because significant charge delocalisation between propionate side chains and the porphyrin ring system was observed in the latter. The difference in charge density between the two separate fragments and the combined molecule thus could not be replicated by the fragment approach. In the case of CQ, however, such charge delocalisation between side chain and *N*-methyl-4-amino-7-chloroquinolinium moieties was not observed.

CQ was separated into two parts, one comprising the side chain and the other the *N*-methyl-4-amino-7-chloroquinolinium moiety. The latter was further fragmented into *N*-methyl-4-aminopyridinium and chlorobenzene. Chlorobenzene has been previously parameterised in the CHARMM General Force Field for drug-like molecules (CGenFF) and thus atomic charges and structure optimisation was not necessary for this compound.[248] The ParamChem program was used to assign the atom types, charges and structural parameters described by CGenFF for chlorobenzene.[349, 350] Similarly, the side chain required no parameterisation as the atom types, charges and structure parameters assigned by ParamChem were similar to molecules previously parameterised. This was confirmed by the low penalty scores assigned to parameters by ParamChem (1.2 and 1.3 for atomic charges and structural parameters respectively) which were well below the suggested re-parameterisation cutoff value of 10. By comparison, parameterisation was required for *N*-methyl-4-aminopyridinium and *N*-methyl-4-amino-7-chloroquinolinium molecules. Three new atom types (CG2DK2, CG2DK4 and NG2DK4) were required to reproduce QM-optimised geometries (MP2/6-31G*) and the crystal structure of CQ.[149] Lennard-Jones parameters were taken from the most closely related atom types CG2R62, CG2R64 and NG311 respectively. Atom typing of the two molecules is displayed in Fig. 6.1.

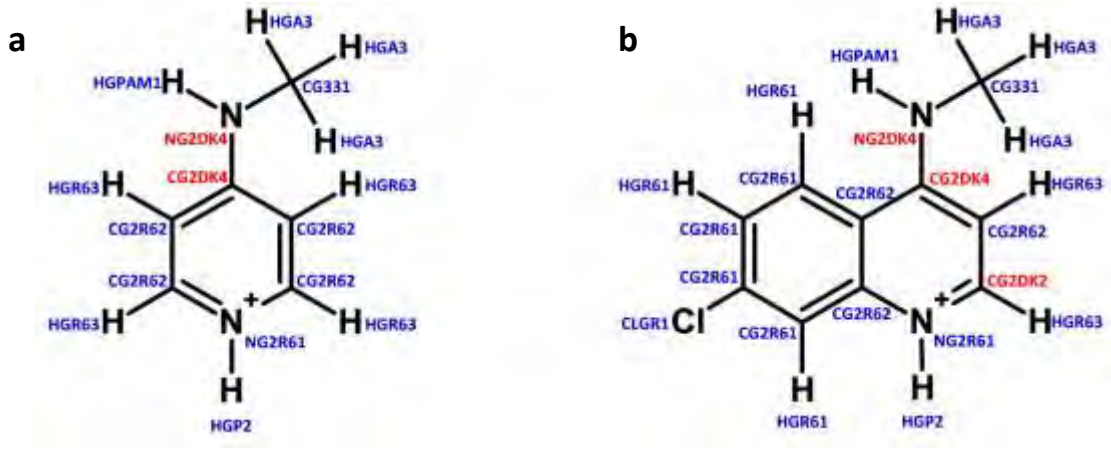


Figure 6.1. Atom typing for (**a**) *N*-methyl-4-aminopyridinium and (**b**) *N*-methyl-4-amino-7-chloroquinolinium. New atoms types and those assigned by ParamChem are shown in red and blue respectively.

Parameterisation of atomic charges was approached in the same manner as described for ferrihaem (see Section 5.2.1 for details), albeit with some minor changes. Firstly, because of the smaller size of these molecules, QM geometries used in determining WIEs were optimised using MP2/6-31G*. Starting structures for these QM geometry-optimisation calculations were made by modifying the crystal structure of CQ reported by Karle.[149] Secondly, WIEs could be calculated according to CHARMM parameterisation methodology using HF/6-31G* because these molecules were diamagnetic and thus did not suffer from spin-contamination as observed for ferrihaem species. Finally, initial atomic charges were obtained from ParamChem as opposed to Mulliken atomic charges as the former were already parameterised for similar structures and thus were likely to be closer to the optimum values. Optimisation of structural parameters was also conducted following that previously described for ferrihaem (see Section 5.2.1 for details). In the case of *N*-methyl-4-aminopyridinium, comparison of QM and MM vibrational frequencies was possible using the VIBRAN module in CHARMM which described vibrations using symmetry coordinates (see Appendix C Table C1). Symmetry coordinates were constructed using internal coordinates defined for *N*-methyl-4-aminopyridinium which were adapted from those reported for pyridine,[352] 4-ethylpyridine[353] and others[354, 355] (see Appendix C Fig. C1 for details). Bond lengths and angles that deviated from the QM values of *N*-methyl-4-aminopyridinium by more than 0.03 Å and 3° respectively were optimised to reproduce both QM target values and vibrations with the largest contribution from that bond or angle.

6.2.2 MD simulation and analysis.

Two models of the CQ-ferrihaem complex were constructed for use in MD simulation. The first involved CQ positioned above one of the unligated faces of [μ-oxo ferrihaem]$^{4-}$, while the second comprised CQ situated between the porphyrin rings with quinolinium nitrogen atom orientated towards the μ-oxo bridge. These conformations are referred to as the π-stacked and docked structures respectively. Initial geometries employed in both conformations were made using the structure of [μ-oxo ferrihaem]$^{4-}$ which most closely matched the average structure of this species over 10 ns of MD simulation in aqueous solution (see Chapter 5 for details). The MM-minimised structure of CQ was subsequently placed in the π-stacked or docked positions. Conformations were allowed to minimise in vacuum, after which they were solvated in a pre-equilibrated TIP4P-Ew water box, neutralised with two sodium ions placed in the box corners and further minimised using steepest decent (100 steps) and conjugate gradient (200 steps) methods. Equilibration and production MD simulations of the CQ-ferrihaem complex were performed using the same procedure as described in Section 5.2.2 and required the same amount of computing time.

SDFs were computed in the same manner as detailed in Section 5.2.2 using the structure of the complex most closely matching the average over the 10 ns simulation. Analysis of porphyrins was performed as described in Section 5.2.2. To determine geometric parameters for CQ relative to porphyrin rings, the least squares plane through the quinoline ring was determined (LSP_{CQ}) and the complex re-orientated such that the porphyrin $LSP_{A/B}$ of μ-oxo ferrihaem lay in the x,y-plane. Interplanar distances between CQ and porphyrin moieties were given by the z-component of LSP_{CQ}. Lateral shifts were calculated by measuring the distance between COM of the porphyrin $LSP_{A/B}$ and the COM of LSP_{CQ} projected onto $LSP_{A/B}$. Interplanar angles, α, were calculated using eq. 5.2.6 in Chapter 5 by substituting one of the porphyrin normal vectors and magnitudes with those for LSP_{CQ}. Interplanar torsion angles between CQ and porphyrin systems were determined from the dihedral defined using the meso carbon atom between propionate side chains, the COM of porphyrin $LSP_{A/B}$ and C4A and C8A atoms of CQ, $C_{meso,A/B}$-$COM_{A/B}$-C4A-C8A. Deviation of iron from the porphyrin plane was determined by measuring the distance of the iron atom from the COM of $LSP_{A/B}$. Diffusion coefficients, *D*, were determined using the MSD method described in Section 5.2.2. The COM of carbon and nitrogen atoms in the quinoline ring system of CQ was used to calculate *D*.

To calculate T_1 relaxation times of CQ protons in the CQ-ferrihaem complex, $T_1^{complex}$, the Solomon-Bloembergen relation shown in eq. 6.1 – 6.3 was used.[172, 356, 357] $T_1^{complex}$ values for each proton were computed for each of the 200 000 structures determined from MD simulation and then averaged.

$$T_1^{complex} = \frac{r^6}{\xi^2[7j(\omega_e) + 3j(\omega_N)]} \tag{6.1}$$

$$\xi^2 = \frac{2}{15}[S(S+1)]\left(\frac{\mu_0}{4\pi}\right)^2\left(\frac{g_e\gamma_e\gamma_N h}{2\pi}\right)^2 \tag{6.2}$$

$$[7j(\omega_e) + 3j(\omega_N)] = \left[\frac{7\tau_c}{1+{\omega_e}^2{\tau_c}^2} + \frac{3\tau_c}{1+{\omega_N}^2{\tau_c}^2}\right] \tag{6.3}$$

In equations 6.1 to 6.3, r is the Fe-H distance, S the electronic spin per iron of [μ-oxo ferrihaem]$^{4-}$, μ_0 the permeability of vacuum ($4\pi \times 10^{-7}$ J s^2 C^{-2} m^{-1}), g_e the electron g-value (2.00232), γ_e and γ_N the electron and nuclear magnetogyric ratios (1.7608×10^{11} and 2.6754×10^8 rad T^{-1} s^{-1} respectively), h the Planck constant (6.62608×10^{-34} J s), ω_e and ω_N the electron and nuclear Larmor frequency for a 300 MHz NMR spectrometer (2.5179×10^9 and 1.8852×10^9 s^{-1} respectively), and τ_c the electron correlation time. The *S(S+1)* term in eq. 6.2 was calculated using the expression in eq. 6.4 and the measured magnetic moment of 2.25 μ_B (see Chapter 3). The electron correlation time could be expressed according to eq. 6.5, where τ_s is the electron spin correlation time, τ_r the rotational relaxation time of the paramagnetic ion and τ_m the nuclear lifetime in the bound state. Since paramagnetic Fe^{3+} has relatively short τ_s (approx. 10^{-9} to 10^{-11} s) and the rotation of the CQ-ferrihaem complex is relatively low, τ_c is dominated by the τ_s term.[358, 359] The value of τ_c was systematically varied to give the best correlation between computed and experimental T_1 relaxation times. The optimal value used (8.5×10^{-11} s) fell within the expected range.

$$S(S+1) = \left(\frac{\mu_B}{2}\right)^2 = \left(\frac{2.25}{2}\right)^2 = 1.27 \tag{6.4}$$

$$\frac{1}{\tau_c} = \frac{1}{\tau_s} + \frac{1}{\tau_r} + \frac{1}{\tau_m} \tag{6.5}$$

6.2.3 Proton longitudinal relaxation time measurements (T_1).

Experimental measurement of T_1 relaxation times required recording a series of ^{1}H NMR spectra using an inversion-recovery sequence 180°, τ, 90°. This was performed on a Varian VXR-300 NMR spectrometer at 298 K. The inversion-recovery sequence involved applying an initial 180° pulse which inverted the magnetisation from the +z-direction to –z, causing nuclei to be in a higher energy state since they were aligned against the applied magnetic field. Nuclei were then allowed to relax for a specified length of time, τ, after which a 90° pulse was introduced. The subsequent return to equilibrium emits radiofrequency energy which was recorded. The recorded signal is known as the free induction decay (FID) and contains information regarding the intensity and frequency of relaxation for each proton in the system. Observed T_1 relaxation times, T_1^{obs}, were determined by recording spectra at varying τ values (see Table 6.1) and fitting the observed i^{th} peak intensity at a particular value of τ, I_i, to eq. 6.6, where I_0 is the intensity after the initial 180° pulse. Fitting of spectroscopic intensities was conducted using the T_1 analysis menu on the Varian NMR workstation. Spectra were recorded on eight aqueous CQ-ferrihaem solutions consisting of CQ (10 mM in D_2O) and linearly varying ferrihaem concentrations (0 to 1.6 mM). These solutions were prepared by diluting stock CQ (50 mM in D_2O) and haematin solutions (20 mM in 0.1 M NaOD) to the required concentrations and were adjusted to measured pD values in the range 7.3 – 7.5 by addition of 0.5 or 0.1 M NaOD.

$$I_i = I_0\left(1 - 2e^{-\tau_i / T_1{}^{obs}}\right) \qquad (6.6)$$

Table 6.1. Values of τ (in s) used to determine T_1^{obs} relaxation times of CQ-ferrihaem solutions.

[ferrihaem] / mM	Spectrum Number										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
0	0.0125	0.025	0.05	0.1	0.2	0.4	0.8	1.6	3.2	6.4	12.8
0.2	0.00625	0.125	0.025	0.05	0.1	0.2	0.4	0.8	1.6	3.2	–
0.4	0.00625	0.125	0.025	0.05	0.1	0.2	0.4	0.8	1.6	3.2	–
0.6	0.00625	0.125	0.025	0.05	0.1	0.2	0.4	0.8	1.6	3.2	–
0.8	0.00625	0.125	0.025	0.05	0.1	0.2	0.4	0.8	1.6	3.2	–
1.0	0.00625	0.125	0.025	0.05	0.1	0.2	0.4	0.8	1.6	3.2	–
1.2	0.0025	0.005	0.01	0.02	0.04	0.08	0.16	0.32	0.64	1.28	2.56
1.4	0.0025	0.005	0.01	0.02	0.04	0.08	0.16	0.32	0.64	1.28	2.56
1.6	0.0025	0.005	0.01	0.02	0.04	0.08	0.16	0.32	0.64	1.28	2.56

Assuming there is fast exchange between bound and free CQ,[172] observed T_1 relaxation times could be described by eq. 6.7, where χ_B and χ_f are the mole fraction of bound and free CQ respectively, and T_1^{free} is the relaxation time of free CQ. $T_1^{complex}$ could be determined by plotting T_1^{obs} against the total concentration of ferrihaem, $[M]_{tot}$, and fitting the resulting curve to eq. 6.9 (derived from substitution of eq. 6.8 into 6.7) using non-linear least-squares methods. This was readily computed since T_1^{free} was experimentally measured and χ_B could be calculated from eq. 6.8, where K_d is the measured dissociation constant ($^1/_{K'} = 3.16 \times 10^{-7}$, see Chapter 3 for K') and $[CQ]_{tot}$ the total CQ concentration (10 mM).

$$\frac{1}{T_1{}^{obs}} = \frac{\chi_B}{T_1{}^{complex}} + \frac{\chi_f}{T_1{}^{free}} = \left(\frac{1}{T_1{}^{complex}} - \frac{1}{T_1{}^{free}}\right)\chi_B + \frac{1}{T_1{}^{free}} \tag{6.7}$$

$$\chi_B = \frac{[M]_{tot}}{2(K_d + [CQ]_{tot})} \tag{6.8}$$

$$T_1{}^{obs} = \frac{2T_1{}^{free}(K_d + [CQ]_{tot})}{[M]_{tot}\left(\frac{T_1{}^{free}}{T_1{}^{complex}} - 1\right) + 2(K_d + [CQ]_{tot})} \tag{6.9}$$

6.2.4 EXAFS.

Collection of the EXAFS spectrum of CQ-ferrihaem and its processing was conducted as described in Section 5.2.3. The spectrum was recorded on a solution prepared by combining 0.10 mL HEPES buffer (0.2 M, pH 7.4), 0.25 mL haemin (20 mM in 0.1 M NaOH), 0.10 mL CQ (25 mM in water), 0.2 mL PEG400 and 0.35 mL water. The resulting solution was pH adjusted to 7.4 using 1 M perchloric acid. Fitting of the experimental spectrum was undertaken in a similar manner as detailed in Section 5.2.3 using solvated π-stacked and docked structures obtained from SDF calculations.

6.3 Parameterisation of CQ

To parameterise CQ, a fragment approach was adopted. This involved generating parameters for separate parts of the molecule and then combining the fragments once optimised. In the case of CQ, it was fragmented into the side chain and quinolinium moieties. The latter was further fragmented into *N*-methyl-4-aminopyridinium and chlorobenzene constituents (see Fig. 6.2). The quinolinium form of CQ was specifically chosen rather than the free base quinoline because, based on the measured pK_a of the quinoline nitrogen atom (7.94),[360] 73% of CQ would be diprotic under the experimental conditions used (aqueous, pH 7.5). While a previous study in aqueous methanol solution suggested CQ interacted with ferrihaem in the singly protonated form (i.e. quinoline nitrogen free base form),[163] a more recent investigation showed that a correlation between ferrihaem association constant in aqueous DMSO and DFT-calculated atomic charges for a series of 7-substituted 4-aminoquinolines was only obtained for molecules with protonated quinoline nitrogen atoms.[151]

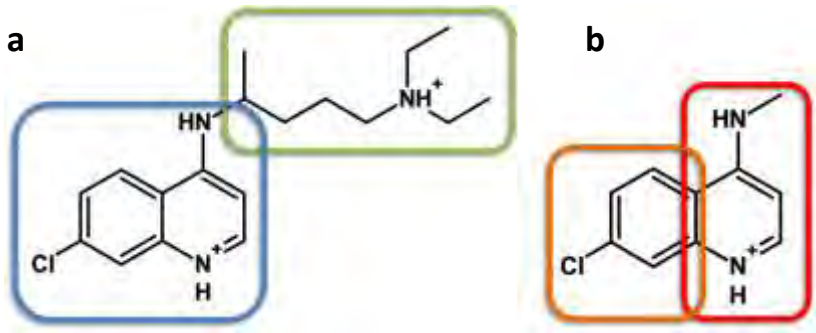


Figure 6.2. Fragments used to parameterise CQ. (**a**) CQ was built by combining separately parameterised side chain (green) and quinolinium moieties (blue). (**b**) The *N*-methyl-4-aminoquinolinium fragment was further separated into *N*-methyl-4-aminopyridinium (red) and chlorobenzene (orange) moieties.

Unlike ferrihaem, the CQ fragments are organic and thus could be described by parameters developed for drug-like molecules using the CHARMM General force field, CGenFF.[248] In fact, parameters already existed for chlorobenzene and parameters for the side-chain of CQ could be readily adopted from similar parameterised molecules. While parameters were available for pyridine, none existed for the pyridinium cation and thus to properly describe the *N*-methyl-4-aminopyridinium molecule refinement of parameters was required.

6.3.1 *N*-Methyl-4-aminopyridinium.

In a similar parameterisation procedure to that used for ferrihaem species, atomic charge parameterisation of *N*-methyl-4-aminopyridinium was conducted such that optimised MM charges reproduced both corresponding QM WIEs and ESPs. The QM structure of this molecule was optimised using MP2/6-31G* which was subsequently used to determine the required ESPs and WIEs with selected atoms (see Fig. 6.3). To ensure consistency with the previously parameterised ferrihaem species, the two sets of target data were calculated using OPBE/6-31G*. In the case of WIEs, these were similarly scaled by 1.12 which was the value determined to scale the interaction energy of a water dimer calculated using OPBE/6-31G* to that obtained using HF/6-31G*. Since this molecule did not suffer from the spin contamination problem experienced for ferrihaem species, interaction energies were also computed for comparison using HF/6-31G*. In general, reasonably good agreement between scaled OPBE/6-31G* interaction energies and those calculated using HF/6-31G* was observed with a MUE of 0.67 kcal mol^{-1} (see Table 6.2).

Initial MM atomic charges were obtained from the ParamChem interface (see Scheme 6.1 for initial charges) which automatically assigns atom types, atomic charges and structural parameters for molecules based on the CGenFF force field.[248, 349, 350] While, WIEs calculated using initial ParamChem atomic charges produced reasonable agreement with QM values (MUE of 1.3 kcal mol^{-1}), further refinement was possible. This involved systematic alteration of atomic charges and fitting to the QM ESP using the FitCharge module in CHARMM (see Sections 5.2.1 and 5.2.2 for more details regarding this process) which produced better agreement with a MUE of less than 1 kcal mol^{-1}. In accordance with the CHARMM parameterisation protocol, charges of the aliphatic hydrogen atoms were set to 0.09 e and excess charged summed into the adjacent carbon atom. WIEs calculated using optimised atomic charges are shown in Table 6.2. Good agreement between QM and MM ESPs were also obtained and are shown in Fig. 6.4. Final optimised charges and atom labelling for *N*-methyl-4-aminopyridinium are shown in Scheme 6.1.

Table 6.2. WIEs (in kcal mol^{-1}) calculated for *N*-methyl-4-aminopyridinium using QM and CHARMM.

	Interaction Type [a]	Water Interaction Energy			
		OPBE/6-31G* [b]	HF/6-31G*	CHARMM initial [e]	CHARMM opt. [f]
1	N1-H1 --- OH_2	-15.95	-15.11	-17.26	-16.89
2	C2-H2 --- OH_2	-9.29	-9.73	-10.66	-10.38
3	C3-H3 --- OH_2	-9.19	-8.85	-9.91	-9.71
4	N4-H4 ---OH_2	-13.62	-12.56	-10.52	-13.67
5	N4 --- H-OH	*DC* [c]	3.08	2.57	3.16
	MUE	*0.67* [d]		*1.34*	*0.90*

[a] See Fig. 6.3 for numbering; [b] QM interaction energy scaled by 1.12; [c] DC = did not converge; [d] MUE between scaled OPBE/6-31G* and HF/6-31G* interaction energies; [e] Interaction energies calculated using atomic charges suggested by ParamChem; [f] Interaction energies calculated using optimised atomic charges.

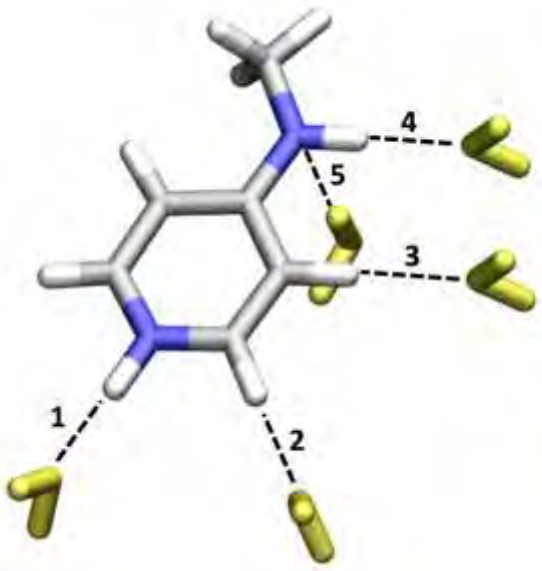


Figure 6.3. Water interactions with selected *N*-methyl-4-aminopyridinium atoms determined from DFT calculation (HF/6-31G*). Water molecules have been coloured yellow and interactions numbered for clarity.

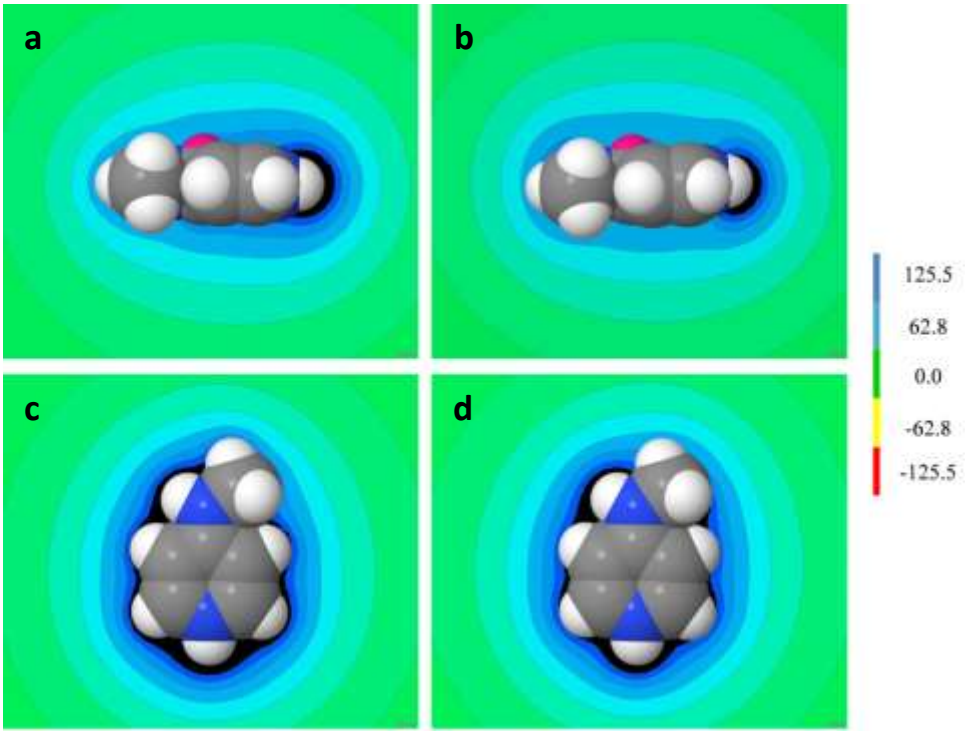
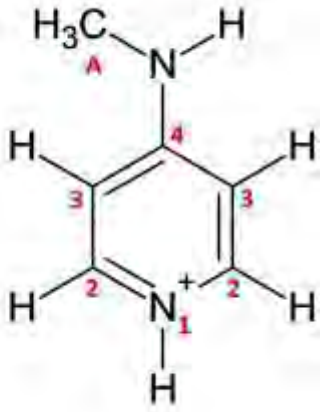


Figure 6.4. QM (left) and MM (right) ESPs of *N*-methyl-4-aminopyridinium from (**a, b**) side and (**c, d**) top views. The pink sphere in (a) and (b) is a dummy atom placed to orientate the ESP plane. ESPs were generated using the Jmol software package. Energies displayed are in kcal mol^{-1}.

Scheme 6.1:

Atom Label	MM Atomic Charge	
	ParamChem	Optimised
N1	-0.489	-0.520
C2	0.149	0.141
C3	-0.050	-0.075
C4	0.196	0.143
N4	-0.496	-0.391
CA	-0.198	-0.203
H1	0.451	0.457
H2	0.190	0.190
H3	0.159	0.142
H4	0.370	0.448
HA	0.090	0.090
Dipole: 16.6 (MP2/6-31G)*	*19.9*	*18.9*



While, the pyridinium moiety of the MM-minimised structure using ParamChem structural parameters was in good agreement with the corresponding QM counterpart, reproduction of the *N*-methyl-4-amino substituent was poor (see Table 6.3). In particular, the structure of this substituent in the MM geometry showed considerable deviation from planarity with an improper dihedral of about 34° out of plane. By contrast, the substituent in the QM geometry was planar. Consequently, deviation in H4-N4-C4, H4-N4-CA and CA-N4-C4 angles were also observed between QM and MM structures. Furthermore, the MM structure also contained a markedly longer C4-N4 bond length (1.410 vs. 1.338 Å).

To address the deviations in MM geometry, two new atom types were created for C4 and N4 atoms, CG2DK4 and NG2DK4 respectively, and planarity of the substituent enforced with the addition of an improper dihedral, N4*-CA-C4-H4. Parameters describing the C4-N4 bond and H4-N4-C4, H4-N4-CA and CA-N4-C4 angles were optimised to reproduce QM values as well as QM-calculated vibrational frequencies (see Section 6.2.1 for details). Selected bond lengths, angles and dihedrals from the structure minimised using optimised parameters are listed in Table 6.3 and a selection of calculated QM and MM vibrational frequencies are shown in Table 6.4. Reasonable agreement was observed for CA-N4-H4 bending and rocking motions as well as for C4-N4 wagging. While poorer agreement was found for the C4-N4 stretching motion, all bond lengths and angles calculated using the optimised parameters had deviations less than the values suggested by the CHARMM parameterisation protocol (0.03 Å and 3° respectively). Finally, addition of the improper dihedral necessitated refinement of parameters describing the H4-N4-C4-C3 and CA-N4-C4-C3 torsion angles. This was conducted by performing a QM dihedral scan around the C4-N4 bond using MP2/6-31G* and then altering the MM dihedral force constant, K_ϕ, of the two torsion angles to reproduce the computed QM energies in CHARMM. Final optimised parameters produced MM dihedral scan energies in good agreement with QM values (see Fig. 6.5). A selection of initial and optimised parameters describing bonds, angles, dihedrals, improper dihedrals and non-bonded terms is shown in Table 6.5.

Table 6.3. Selected bond lengths, angles and dihedrals for the *N*-methyl-4-aminopyridinium structure optimised using QM and CHARMM.

	MP2/6-31G*	PC Parameters [a]	Opt. Parameters [b]
N1-C2	1.358	1.375	1.379
C2-C3	1.372	1.371	1.370
C3-C4	1.423	1.372	1.430
C4-N4	1.338	1.410	1.349
N4-CA	1.460	1.480	1.481
N4-H4	1.015	1.024	1.025
N1-H1	1.017	1.017	1.018
N1-C2-C3	120.6	121.4	121.4
C2-C3-C4	119.9	119.7	119.3
C3-C4-N4	121.4	116.9	121.3
C3-C4-C3	117.3	119.8	118.6
C2-N1-C2	121.7	118.0	120.0
H1-N1-C2	119.2	121.0	120.1
H4-N4-C4	117.6	116.3	116.0
CA-N4-C4	124.8	115.7	125.9
H4-N4-CA	117.6	105.9	118.1
C3-C4-N4-CA	180.0	-0.8	-179.6
C3-C4-N4-H4	180.0	-56.7	-178.1
C8-C4-H4-*N4	0.0	-34.3	-1.3

[a] Calculated using parameters suggested by ParamChem; [b] Calculated using optimised parameters.

Table 6.4. Selected vibrational frequencies (in cm^{-1}) calculated for *N*-methyl-4-aminopyridinium using MP2/6-31G* and CHARMM with optimised parameters. [a, b]

Vibrational label	Vibrational mode	MP2/6-31G*	CHARMM
V_3	C4-N4 stretch	1344 (18%)	1221 (17%)
V_{11}	CA-N4-H4 bend	1278 (20%)	1331 (19%)
V_{22}	CA-N4-H4 rock	221 (33%)	266 (36%)
V_{23}	C4-N4 wag	221 (48%)	266 (34%)

[a] see Appendix C Fig. C1 and Table C1 for description of vibrational labels and modes; [b] only frequencies with highest vibrational mode contribution (given in parenthesis) are listed.

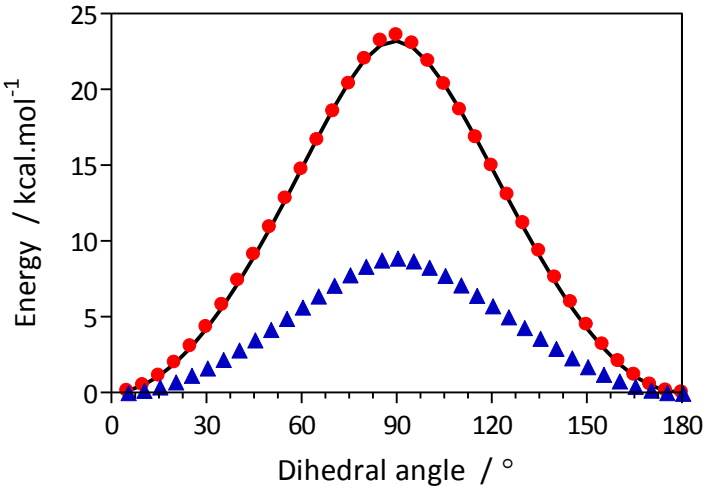


Figure 6.5. Dihedral scan around the C4-N4 bond in *N*-methyl-4-aminopyridinium calculated using MP2/6-31G* (solid line), CHARMM with ParamChem parameters (blue triangles) and CHARMM with optimised parameters (red circles).

Table 6.5. Selected initial and optimised parameters for *N*-methyl-4-aminopyridinium.

	Initial Guess			**Optimised**		
Bonds:	K_b (kcal mol^{-1} Å^{-2})		b_0 (Å)	K_b (kcal mol^{-1} Å^{-2})		b_0 (Å)
C4-N4	360.00		1.37 [a]	360.00		1.34
Angles:	K_θ (kcal mol^{-1} rad^{-2})		b_θ (°)	K_θ (kcal mol^{-1} rad^{-2})		b_θ (°)
C3-C4-N4	81.00		118.40 [a]	100.00		118.40
C4-N4-CA	43.00		106.00 [b]	40.00		118.00
CA-N4-H4	45.00		104.00 [b]	25.00		118.00
Dihedrals:	K_Φ (kcal mol^{-1})	*n*	δ (°)	K_Φ (kcal mol^{-1})	*n*	δ (°)
C3-C4-N4-H4	2.00	2	180.00 [a]	2.00	2	180.00
C3-C4-N4-CA	2.37	2	180.00 [b]	4.40	2	180.00
Improper:	K_ψ (kcal mol^{-1} rad^{-2})		ψ_0 (°)	K_ψ (kcal mol^{-1} rad^{-2})		ψ_0 (°)
N4*-H4-C4-CA	-2.50		0.00 [a]	-2.50		0.00
Non-bonded:	ε^{min} (kcal mol^{-1})		R_{min} (Å)	ε^{min} (kcal mol^{-1})		R_{min} (Å)
C4	–0.040		2.10 [a]	–0.040		2.10
N4	–0.045		2.00 [b]	–0.045		2.00

[a] Nearest analogues parameters in CHARMM; [b] assigned by ParamChem.

6.3.2 *N*-Methyl-4-amino-7-chloroquinolinium.

N-methyl-4-amino-7-chloroquinolinium was built in CHARMM by combining structural parameters for chlorobenzene obtained from ParamChem and the optimised parameters determined for *N*-methyl-4-aminopyridinium. Initial atomic charges used for *N*-methyl-4-amino-7-chloroquinolinium were taken as those suggested by ParamChem and were optimised to reproduce WIEs calculated using HF/6-31G* (see Fig. 6.6) as well as QM ESPs using the FitCharge module in CHARMM. WIEs and ESPs calculated using optimised atomic charges are shown in Table 6.6 and Fig. 6.7 respectively. QM counterparts have been included for comparison.

Figure 6.6. Water interactions with selected *N*-methyl-4-amino-7-chloroquinolinium atoms determined from DFT calculation (HF/6-31G*). Water molecules have been coloured yellow and interactions numbered for clarity.

Table 6.6. WIEs (in kcal mol^{-1}) calculated for *N*-methyl-4-amino-7-chloroquinolinium using QM and CHARMM.

	Interaction Type [a]	Water Interaction Energy			
		OPBE/6-31G* [b]	HF/6-31G*	CHARMM initial [d]	CHARMM opt. [e]
1	C3-H3 --- OH_2	-7.50	-7.55	-7.97	-7.53
2	C2-H2 --- OH_2	-8.81	-9.45	-10.80	-9.03
3	N1-H1 --- OH_2	-15.29	-14.69	-18.00	-15.09
4	C8-H8 ---OH_2	-8.34	-8.07	-7.86	-7.91
5	Cl --- H-OH	0.89	0.90	-0.15	0.10
6	C6-H6 --- OH_2	-6.15	-6.09	-5.98	-6.51
7	C5-H5 --- OH_2	-10.92	-10.78	-8.57	-11.05
8	N4-H4 --- OH_2	-10.32	-10.24	-7.30	-9.61
	MUE	*0.23* [c]		*1.45*	*0.39*

[a] See Fig. 6.5 for numbering; [b] QM interaction energy scaled by 1.12; [c] MUE between scaled OPBE/6-31G* and HF/6-31G* interaction energies; [d] Interaction energies calculated using atomic charges suggested by ParamChem; [e] Interaction energies calculated using optimised atomic charges.

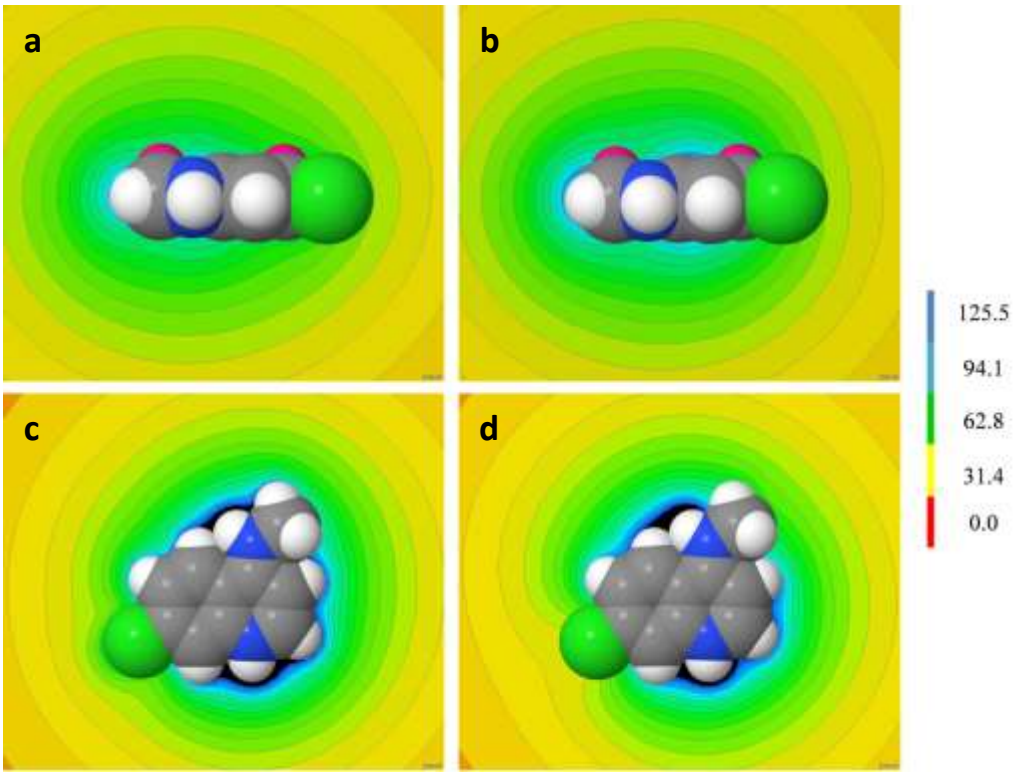


Figure 6.7. QM (left) and MM (right) ESPs of *N*-methyl-4-amino-7-chloroquinolinium from (**a, b**) side and (**c, d**) top views. The pink sphere in (a) and (b) are dummy atoms placed to orientate the ESP plane. ESPs were generated using the Jmol software package. Energies displayed are in kcal mol^{-1}.

As observed for *N*-methyl-4aminopyridinium, WIEs calculated using OPBE/6-31G* and scaled by 1.12 were in good agreement with those computed using HF/6-31G* where the MUE between the two was 0.23 kcal mol^{-1}. These findings reinforce the validity of using scaled OPBE/6-31G* WIEs for ferrihaem species to parameterise atomic charges. Optimised atomic charges of *N*-methyl-4-amino-7-chloroquinolinium reproduced QM WIEs well, where a MUE of 0.4 kcal mol^{-1} was obtained and no deviations greater than 0.6 kcal mol^{-1} were observed. Reasonable agreement between ESPs was produced using optimised charges, although the region surrounding the chlorine substituent was not as good. The atomic charge on this atom could not be further altered without significantly affecting WIEs of the adjacent hydrogen atoms. Furthermore, because the atomic charges on the chlorobenzene molecule had been previously parameterised, it was deemed prudent not to substantially alter the atomic charge of this substituent far from this initial value. Optimised atomic charges for *N*-methyl-4-amino-7-chloroquinolinium are given in Scheme 6.2.

Scheme 6.2:

Atom Label	Atomic Charge	
	ParamChem	Optimised
N1	-0.519	-0.556
C2	0.149	0.130
C3	-0.062	-0.081
C4	0.172	0.144
C4A	0.064	0.051
C5	-0.062	-0.070
C6	-0.114	-0.121
C7	0.073	0.077
C8	0.067	0.068
C8A	0.272	0.244
N4	-0.491	-0.457
Cl	-0.144	-0.134
CA	-0.198	-0.214
H1	0.451	0.406
H2	0.190	0.164
H3	0.159	0.164
H4	0.390	0.480
H5	0.115	0.153
H6	0.165	0.175
H8	0.073	0.107
HA	0.090	0.090
Dipole: 30.8 (MP2/6-31G)*	*31.1*	*29.5*

Describing *N*-methyl-4-amino-7-chloroquinolinium using the optimised structural parameters for *N*-methyl-4-aminopyridinium in conjunction with those suggested by ParamChem for chlorobenzene, produced a geometry that had good structural agreement with the corresponding QM geometry (see Table 6.7 for bond length comparisons). MUEs for bond lengths and angles were 0.02 Å and 1° respectively with no deviation exceeding 0.03 Å or 4°. However, comparing the values with those obtained for the corresponding moiety in the crystal structure of CQ indicated that the N1-C2 bond was too long by about 0.04 Å. To reproduce the shorter bond, a new atom type was created for the C2 atom (CG2DK2) and given structural parameters used in describing the free base form of quinoline. Using this modified parameter set produced better agreement between both experimental and QM values with MUEs obtained for bond lengths and angles of 0.01 Å and 1° respectively, and all deviations being less than 0.03 Å and 3°.

Table 6.7. Experimental and computed bond lengths of *N*-methyl-4-amino-7-chloroquinolinium.

Bonds	Exp. [a]	MP2/6-31G*	Initial Parameters [b]	Modified Parameters [c]
N1-C2	1.324	1.344	1.359	1.326
C2-C3	1.361	1.376	1.353	1.367
C3-C4	1.404	1.413	1.424	1.430
C4-C4a	1.446	1.446	1.434	1.455
C4a-C5	1.413	1.415	1.408	1.418
C5-C6	1.370	1.379	1.400	1.398
C6-C7	1.402	1.411	1.396	1.392
C7-C8	1.371	1.381	1.396	1.392
C7-Cl	1.724	1.720	1.738	1.737
C8-C8a	1.395	1.406	1.405	1.406
C8a-C4a	1.417	1.417	1.388	1.403
C8A-N1	1.371	1.378	1.363	1.378
C4-N4	1.335	1.336	1.352	1.358
N4-CA	1.482	1.459	1.479	1.481
N4-H4	-	1.007	1.021	1.024
N1-H1	-	1.013	1.014	1.015
Dipole:		*30.8*	*26.6*	*26.5*

[a] From the crystal structure of CQ; [b] computed using optimised parameters for *N*-methyl-4-aminopyridinium and parameters for chlorobenzene from ParamChem; [c] computed using modified structural parameters involving the C2 atom.

6.3.3 Chloroquine.

The MM model of CQ was constructed using the atomic charges and structural parameters optimised for *N*-methyl-4-amino-7-chloroquinolinium in conjunction with those suggested by ParamChem for the tertiary amine side chain. To ensure that rotation about the N4-CA bond joining the side chain to quinolinium moiety was correctly described, a QM dihedral scan around the corresponding bond in the model compound *N*-isopropyl-4-aminopyridinium was performed using MP2/6-31G*. This molecule was built in CHARMM using optimised parameters for *N*-methyl-4-aminopyridinium and those suggested by ParamChem. Only the MM dihedral force constant describing the C3-C4-N4-CA torsion angle required alteration to produce MM dihedral energies that matched corresponding QM values (see Fig. 6.8). These parameters were then used for the corresponding dihedral in CQ.

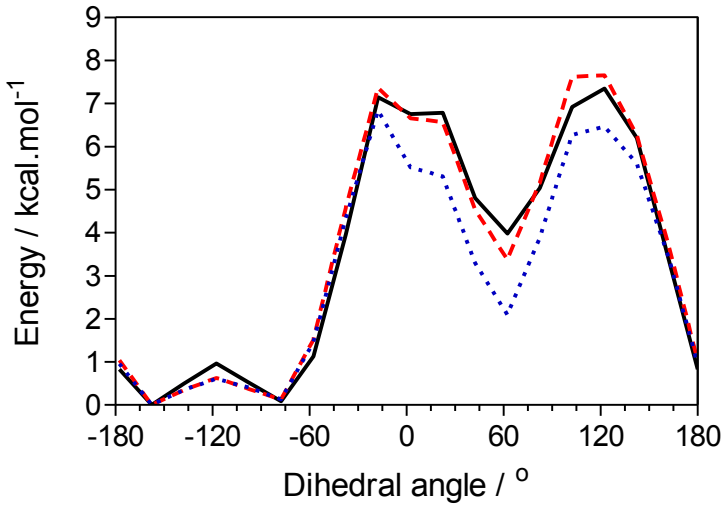


Figure 6.8. Dihedral scan around the N4-CA bond in *N*-isopropyl-4-aminopyridinium calculated using MP2/6-31G* (solid line), CHARMM with parameters suggested by ParamChem (blue dotted line) and CHARMM with optimised parameters (red dashed line). Optimised K_{φ} value for the C3-C4-N4-CA torsion was 3.57 kcal mol^{-1}.

6.4 Structure and Solution Behaviour of the CQ-Ferrihaem Complex

6.4.1 Determining the position of CQ in the CQ-ferrihaem complex.

The experimental findings reported in Chapter 3 indicated that CQ induces the μ-oxo dimeric form of ferrihaem in aqueous solution. Consequently, MD simulations were conducted using the parameterised CQ model and that developed for [μ-oxo ferrihaem]$^{4-}$ (see Chapter 5). A number of studies have suggested that CQ interacts with the unligated face of the μ-oxo dimer through π-stacking interactions[172, 173, 176, 347] and have subsequently performed MD simulations of the complex with CQ in this orientation (see Fig. 6.9a for representative structure). However, an additional conformation was considered possible in which CQ was positioned between the porphyrin rings (see Fig. 6.9b for representative structure). Consequently, two separate simulations of the CQ-ferrihaem complex in the π-stacked and docked conformations were performed.

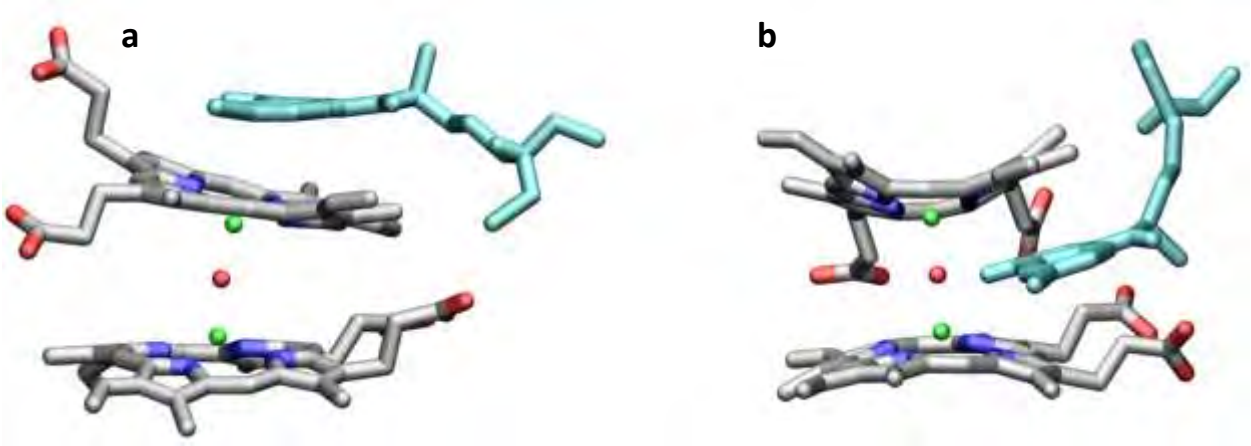


Figure 6.9. Conformations of CQ in the CQ-ferrihaem complex. (**a**) CQ π-stacking with the unligated face of [μ-oxo ferrihaem]$^{4-}$. (**b**) CQ docked between the two porphyrin rings of [μ-oxo ferrihaem]$^{4-}$. For clarity, only N-H hydrogen atoms are shown and CQ is coloured in cyan.

To experimentally distinguish between these structures, proton longitudinal (T_1) relaxation measurements were performed. The presence of the paramagnetic iron centre affects proton relaxation times by increasing the effective field at the nucleus. As a result, NMR peaks tend to shift to lower field strengths and become broader in the presence of paramagnetic ions (see Fig. 6.10a and b).[361-363]

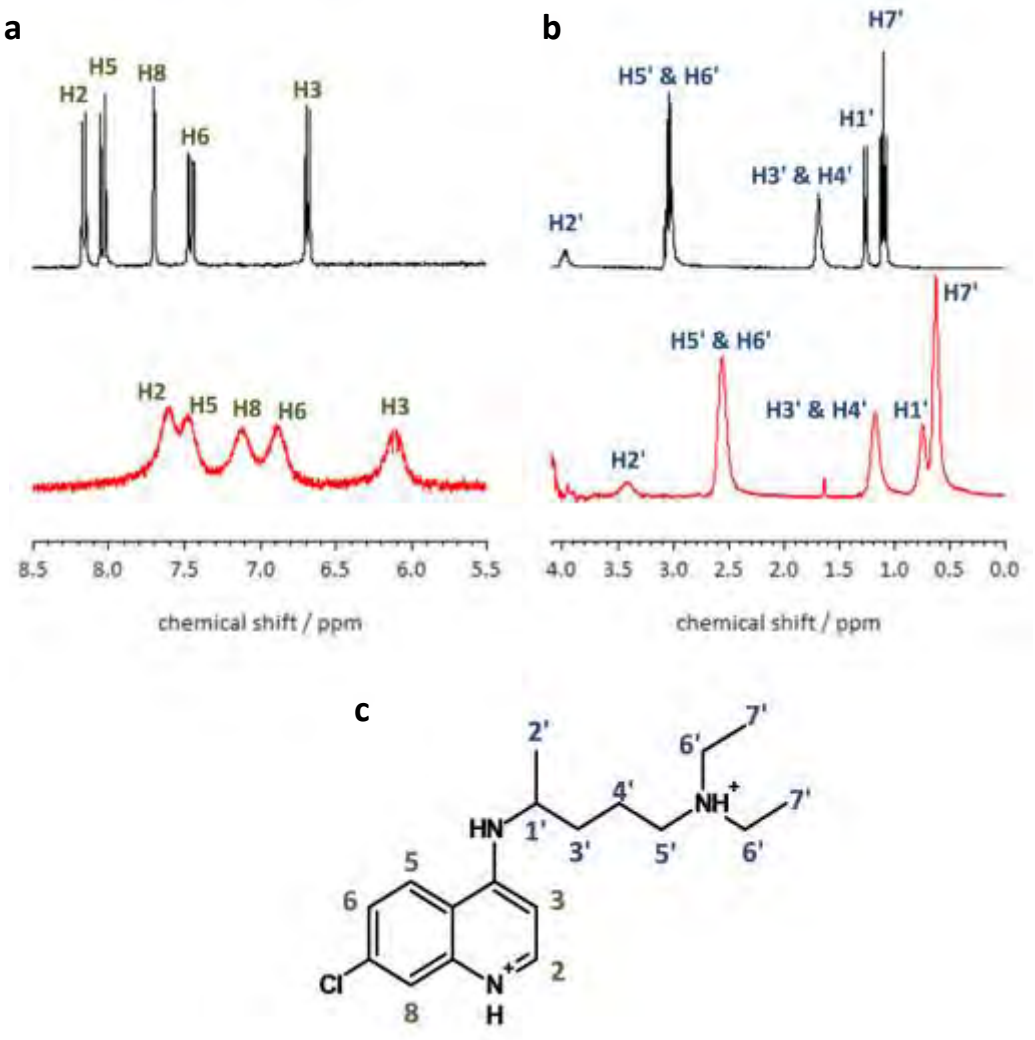


Figure 6.10. (**a**) Aromatic and (**b**) aliphatic regions of the ^{1}H NMR spectra of CQ (black) and CQ in the presence of 1.4 mM paramagnetic ferrihaem (red) showing peak shifts and broadening that occurs in the latter. (**c**) Numbering scheme of CQ.

T_1 relaxation times have been shown to have a distance-dependence through the Solomon-Bloembergen relation (see Section 6.2.2). Consequently, protons which are positioned closer to the iron centre in the complex will have shorter T_1 relaxation times than those further away and thus this relationship can be used to probe proton distances from the iron centre. Measurement of T_1 relaxation times required recording a series of spectra using an inversion-recovery sequence and fitting observed intensities to eq. 6.6 (see Section 6.2.3). A representative example of the inversion-recovery spectra and corresponding fit to the data to determine the observed T_1 relaxation time, T_1^{obs}, is shown in Fig. 6.11a and b respectively.

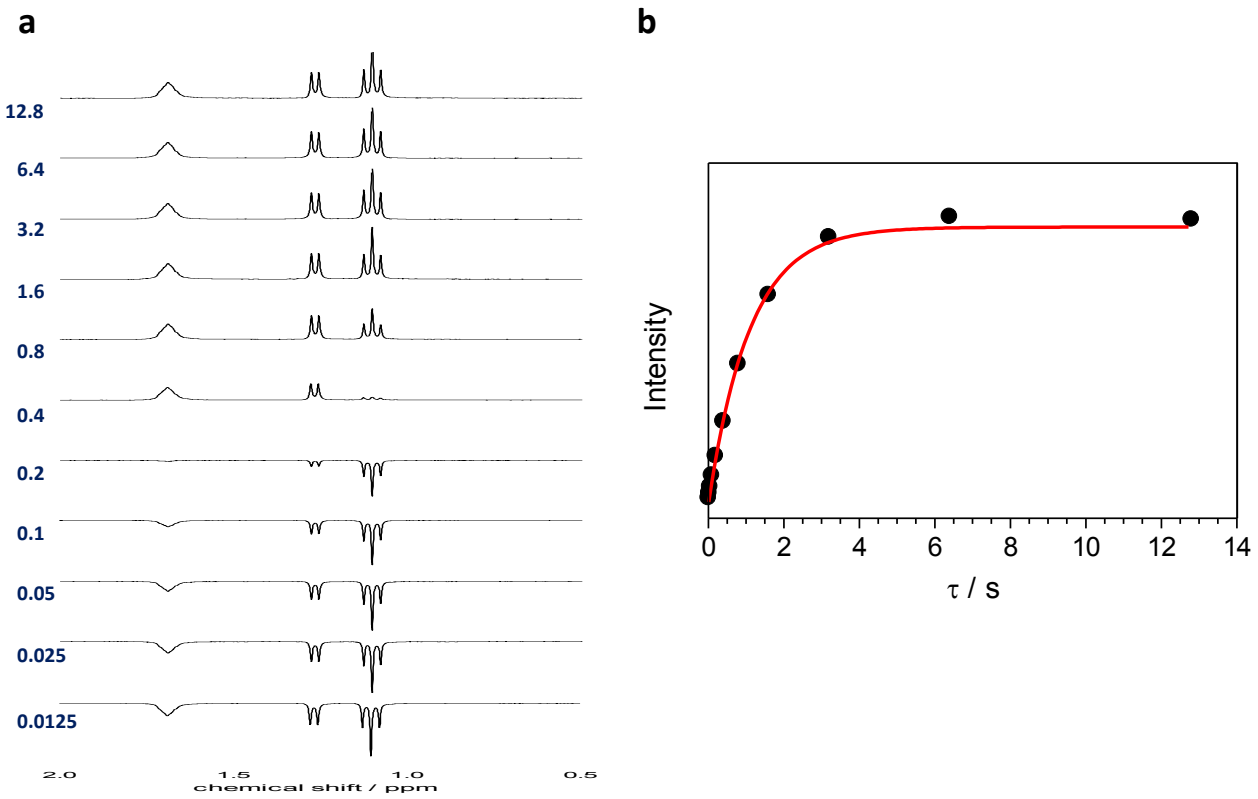


Figure 6.11. (**a**) Representative inversion-recovery ^{1}H NMR experiment for CQ over the range 0.5 to 2 ppm. Values of τ (in s) used are shown on the left in blue. (**b**) Plot of intensity vs. τ for the peak corresponding to H7' at 1.1 ppm used to determine T_1^{obs} from eq. 6.6 (see Section 6.2.3).

In a previous study, such investigations on the CQ-ferrihaem complex in aqueous solution were conducted by Leed *et al*.[172] where T_1 relaxation times were used to determine putative average proton distances from the iron centre which were subsequently used as constraints in MD simulations. From these simulations, structures of the complex in solution were proposed, although they consisted of conformations only in which CQ was π-stacked with [μ-oxo ferrihaem]$^{4-}$. However, several important criticisms of the reported study can be made. Firstly, the authors used a spin-state of S = ½ for the iron centre of [μ-oxo ferrihaem]$^{4-}$ which was incorrect. As described in Section 2.5, the antiferromagnetic coupling between the two centres produces a magnetic moment similar but not identical to that expected for a S = ½ spin-state. Ideally, the spin-state calculated from the measured magnetic moment should have been used. This was correctly addressed in the present study using eq. 6.4 in Section 6.2.2. Secondly, the authors concluded that the structures produced were representative of those likely in the solution state. Indeed, while the distance constraints imposed were obtained from solution data and thus did infer some solution-state information, the reported MD simulations did not mention the use of implicit or explicit solvent molecules. By not incorporating solvent in the simulation important solution-phase hydrogen bonding and electrostatic dampening effects could not be properly described and hence unrealistic structures may still have been produced. To avoid this, MD simulations performed in this work were undertaken using explicit solvent water molecules. Lastly, the authors made a statistical error by incorrectly equating the distance calculated from the average T_1 relaxation time to the average Fe-H distance, *r*. As detailed in Section 1.5, the error arises from the fact that $\langle r \rangle \neq \sqrt[6]{\langle r^6 \rangle}$.[177] This error was circumvented in this study by rather calculating T_1^{obs} using *r* values from MD simulation as opposed to Leed *et al*.[172] who used experimental T_1^{obs} to determine *r*. This approach had the added advantage that MD simulations were performed independently of T_1 NMR experiments and thus the latter could be used to test the feasibility of the computed complex conformations.

T_1 relaxation times for each proton of CQ in the complex, $T_1^{complex}$, were determined by fitting eq. 6.9 (see Section 6.2.3 for details) to the non-linear plot between the T_1^{obs} and concentration of ferrihaem, $[M]_{tot}$. This approach differed to that used by Leed and co-workers where $T_1^{complex}$ was determined using the linear relationship between $1/T_1^{obs}$ and mole fraction of bound CQ, χ_B. For the sake of comparison, the same linear relationship was constructed from the data obtained in this study and is displayed alongside that reported by Leed *et al*. in Fig. 6.12a and b.

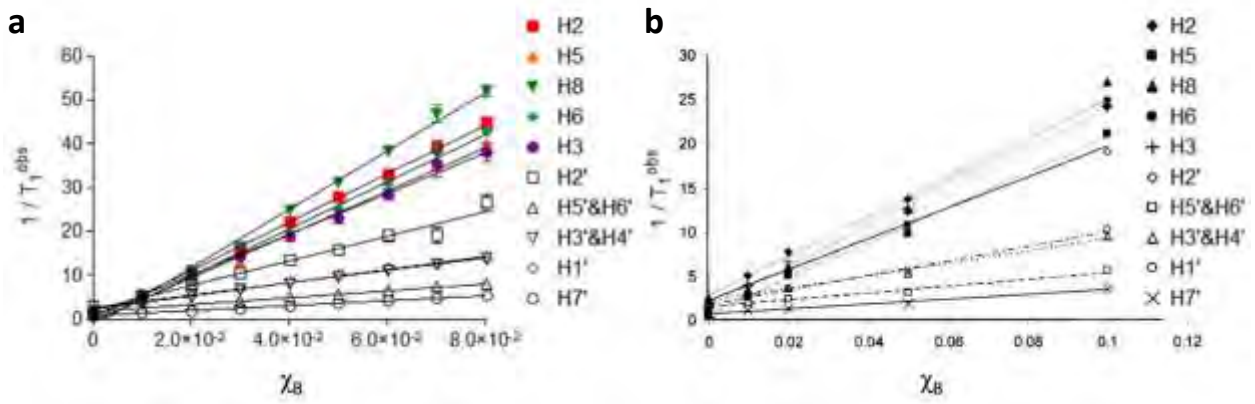


Figure 6.12. Plots of $1/T_1^{obs}$ against mole fraction of bound CQ, χ_B, (**a**) from this study and (**b**) that reported by Leed *et al*.[172] See Fig. 6.10 for labelling of CQ protons. In (**a**), χ_B was calculated from eq. 6.8. Figure 6.8b has been adapted with permission from Leed *et al*.[172] © 2002, American Chemical Society.

$T_1^{complex}$ values determined in the current study are listed in Table 6.8 and corresponding $T_1^{complex}$ values calculated from Fig. 6.12b have been included for comparison. A significant linear correlation was obtained between the two sets of data (r^2 = 0.996 and P-stat < 0.0001), however, a slope of 1.93 ± 0.04 was observed which indicated the values determined in this study were approximately double those obtained by Leed and co-workers. It was suspected that this discrepancy likely arose through incorrect determination of $[CQ]_{bound}$ by the latter authors. In the reported study, it was mentioned that CQ binds to [μ-oxo ferrihaem]$^{4-}$ in a 1:1 stoichiometry which is correct, however, they may have instead calculated $[CQ]_{bound}$ according to a 1:1 CQ:ferrihaem ratio in error. Unfortunately, details regarding concentrations were not reported and so this could not be confirmed. Nevertheless, the same trend was obtained and thus results remain qualitatively comparable.

Table 6.8. Experimental $T_1^{complex}$ (× 10^{-3} s) values for the CQ-ferrihaem complex in aqueous solution determined in this study and reported by Leed *et al*.[172]

Proton label [a]	Experiment	
	This work	Leed *et al*.
H2	2.05	4.55
H3	2.36	4.55
H5	2.51	4.99
H6	2.40	4.29
H8	1.92	3.81
H1'	6.40	11.16
H2'	3.24	6.01
H3' & H4'	7.44	13.01
H5' & H6'	12.92	25.93
H7'	19.33	37.15

[a] See Fig. 6.10c for labels.

Calculated $T_1^{complex}$ values for each proton of CQ in the π-stacked and docked conformation of the CQ-ferrihaem complex were determined using the Solomon-Bloembergen equation for every structure over the course of the MD simulation (200 000) and averaged. Comparison of average computed $T_1^{complex}$ values determined from both MD simulations with experimental values produced significant linear correlations (see Fig. 6.13a and c). The linear relationship observed for the π-stacked conformation produced a similar correlation to that obtained from the data of Leed and co-workers (see Fig. 6.13e), although exhibited a slightly lower significance. Conversely, a strong correlation was observed for the docked structure which retained significance even when the $T_1^{complex}$ values of only the aromatic and H2' protons were considered (see Fig. 6.9b). By contrast, using the same protons produced no significant correlation for the π-stacked conformations simulated in this study or that reported by Leed *et al*.[172] (see Fig. 6.9d and f). Thus the π-stacked structures did not reproduce the trend whereas the docked conformation did.

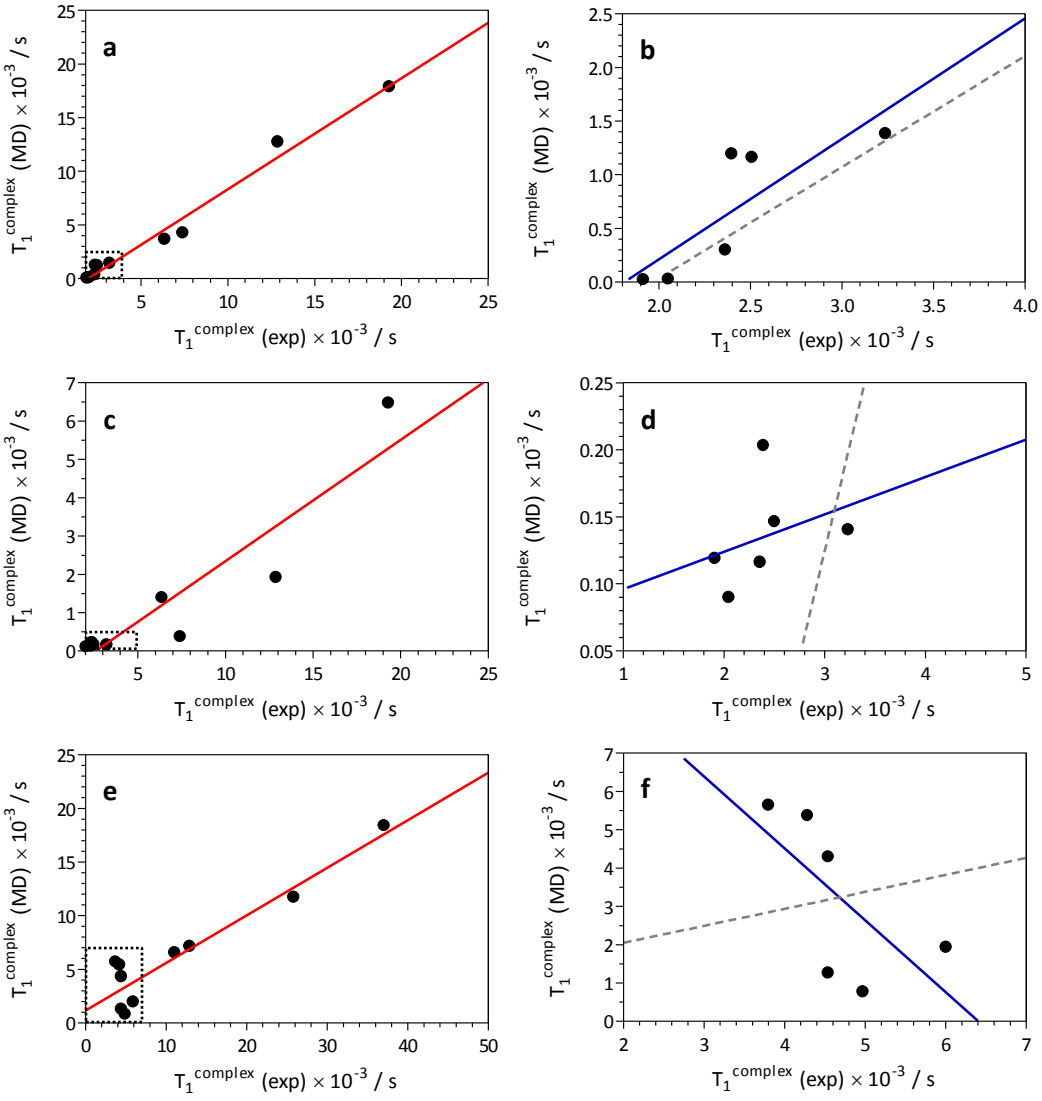


Figure 6.13. Left. Linear correlations (red lines) between computed MD and experimental $T_1^{complex}$ relaxation times of complexed CQ. **Right.** Expansions of dotted area on the left. Blue lines are the linear correlations obtained using aromatic protons and H2' only. Grey dashed lines are the linear correlations for all protons determined on the left and are included for comparison. (**a**) $T_1^{complex}(MD)$ values computed using CQ in the docked conformation. Regression described by $T_1^{complex}(MD)$ = 1.04(5) $T_1^{complex}(exp.)$ – 2.0(4) × 10^{-3}; r^2 = 0.98, P-stat < 0.0001. (**b**) Regression (blue line) described by $T_1^{complex}(MD)$ = 1.1(4) $T_1^{complex}(exp.)$ – 2(1) × 10^{-3}; r^2 = 0.68, P-stat = 0.044. (**c**) $T_1^{complex}(MD)$ values computed using CQ in the π-stacked conformation. Regression described by $T_1^{complex}(MD)$ = 0.32(4) $T_1^{complex}(exp.)$ – 8.3(7) × 10^{-3}; r^2 = 0.86, P-stat = 0.0001. (**d**) Regression (blue line) not significant, P-stat = 0.52. (**e**) Correlation obtained using $T_1^{complex}$ data reported by Leed *et al.*[172] Regression described by $T_1^{complex}(MD)$ = 0.44(6) $T_1^{complex}(exp.)$ – 1.2(9) × 10^{-3}; r^2 = 0.88, P-stat < 0.0001. (**f**) Correlation (blue line) not significant, P-stat = 0.16.

In a further attempt to experimentally distinguish between the docked and π-stacked conformation of the CQ-ferrihaem complex, EXAFS spectroscopy was employed. Given the good agreement obtained when using solvation information determined from SDF calculations for [μ-oxo ferrihaem]$^{4-}$, the same approach was applied to the CQ-ferrihaem complex which first required determining the SDFs from MD simulation. SDFs calculated for the docked and π-stacked conformations of the CQ-ferrihaem complex in aqueous solution are shown in Fig. 6.14. The SDF of CQ has been also included for comparison.

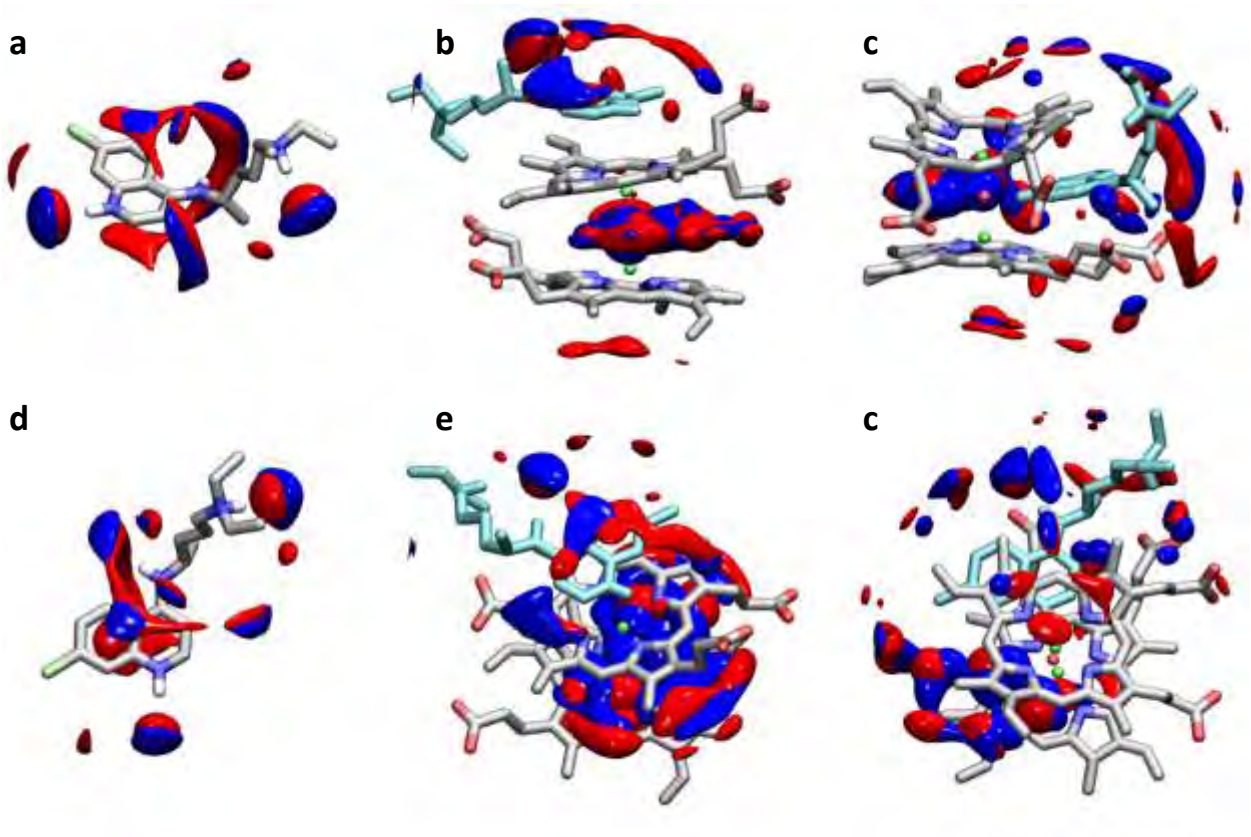


Figure 6.14. SDFs of CQ (**a, d**) and the CQ-ferrihaem complex in π-stacked (**b, e**) and (**c, f**) docked conformations at 50% greater density than bulk. Red and blue isosurfaces indicate solvation by oxygen and hydrogen atoms of water respectively. Porphyrin and selected hydrogen atoms of CQ have been removed for clarity and CQ in the CQ-ferrihaem complex shown in cyan to aid visualisation.

As seen in the SDF of [μ-oxo ferrihaem]$^{4-}$ (see Section 5.4.3), solvation of the μ-oxo bridge is prominent in both conformations of the CQ-ferrihaem complex. In the case of the π-stacked orientation, the bridging ligand was solvated by three waters and showed a marked second solvation shell around these molecules (Fig. 6.14b and e). By contrast, the μ-oxo bridge in the docked conformation was only solvated by two water molecules and had a slightly less prominent second solvation shell (Fig. 6.14c and f). These SDFs indicated that CQ in the docked position replaced one of the solvating water molecules. Furthermore, the hydrogen bonding interaction between the quinolinium nitrogen and μ-oxo ligand in the complex replaced a water molecule solvating CQ in the same position in the absence ferrihaem and there was reduced solvation of the aromatic ring of CQ in the complex (Fig. 6.14a and d). The release of the two waters in the docked conformation, reduction in second solvation shell around the μ-oxo bridge and decreased solvation of the CQ aromatic ring suggests that a favourable increase in entropy would likely occur upon formation of this structure. Indeed, this is consistent with experimental UV-visible measurements of CQ and ferrihaem in 40% aq. DMSO solution reported by Egan *et al*. where a positive entropy component of 21.5 cal mol^{-1} K^{-1} was obtained.[124]

To further investigate the conformation of the CQ-ferrihaem complex in aqueous solution, the solvated structures obtained from SDFs for the π-stacked and docked conformations were then used in conjunction with EXAFS spectroscopy. The spectrum of the CQ-ferrihaem complex in frozen aqueous 20% (v/v) PEG400 solution (pH 7.5) was recorded and, with the exception of the region around 3 Å, was remarkably similar to that obtained for [μ-oxo ferrihaem]$^{4-}$ in aq. DMSO (see Fig. 6.15a). In a similar procedure as described for [μ-oxo ferrihaem]$^{4-}$, the EXAFS spectrum of the CQ-ferrihaem complex was fitted employing the solvated structure obtained from SDF calculations. Using the π-stacked conformation in which N1, C2 and C8A CQ atoms as well as three water molecules surrounding the oxide ligand were included, produced a poor fit to the experimental spectrum (Fig. 6.15b). On the other hand, remarkably good agreement was obtained using the docked structure which included two water molecules solvating the axial μ-oxo bridge, N1 C2 and C8A atoms of CQ as well as one water molecule solvating the unligated face (see Fig. 6.15c). Excluding the solvent and CQ atoms from this model considerably worsened the fit (see Fig. 6.15d).

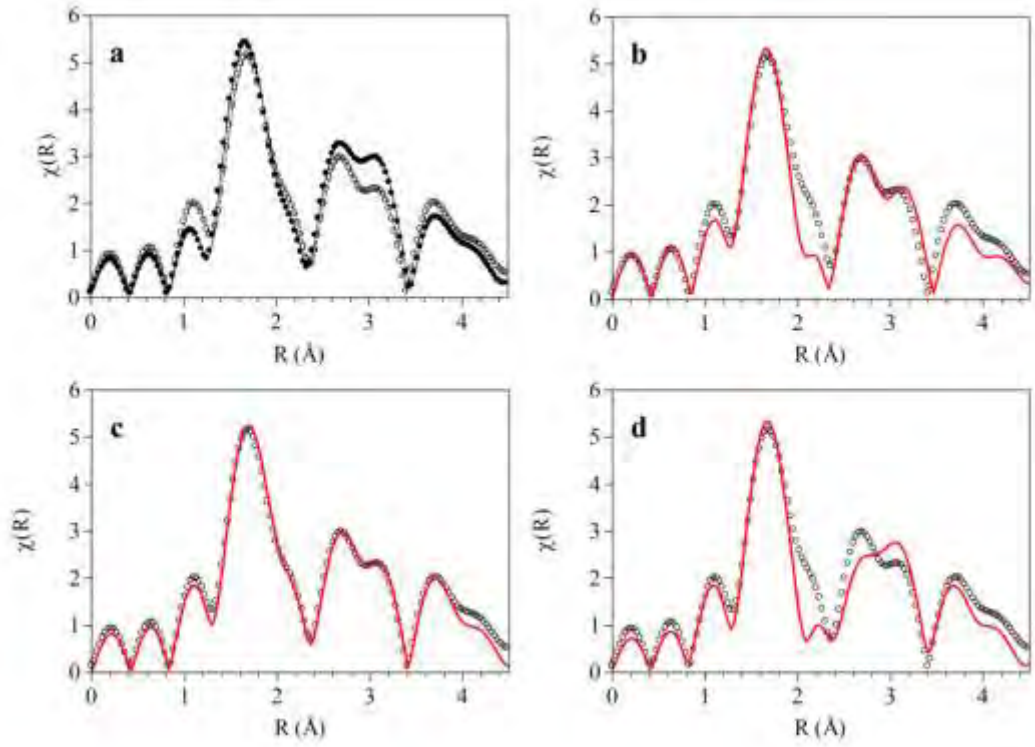


Figure 6.15. Experimental and fitted EXAFS spectra. (**a**) Experimental spectra of [μ-oxo ferrihaem]$^{4-}$ in frozen aqueous 20% (v/v) PEG400, 40% DMSO solution at pH 10 (filled circles) and the CQ-ferrihaem complex in frozen 20% aq. PEG400 solution at pH 7.5 (open circles). (**b**) Experimental spectrum of CQ-ferrihaem complex (open circles) and the fit obtained using the π-stacked structure including CQ atoms and water identified from SDF calculation (solid red line). (**c**) Experimental spectrum of CQ-ferrihaem complex (open circles) and the fit obtained using the docked structure including CQ atoms and water identified from SDF calculation (solid red line). (**d**) Experimental spectrum of CQ-ferrihaem complex (open circles) and the fit obtained using the docked structure excluding water and CQ atoms (solid red line).

EXAFS spectra were fitted in the same manner as described for [μ-oxo ferrihaem]$^{4-}$ where adjustable scaling factors were used (see Section 5.4.3). In the case of the π-stacked structure, five α scaling factors were employed which included two describing the axial oxide ligand and iron centre of the adjacent porphyrin, α_{oxo} and α_{Fe}; one describing CQ atoms N1, C2 and C8A, α_{CQ}; one describing the three water molecules surrounding the μ-oxo bridge, $\alpha(2)_{wat}$; and one to represent all the rest. Four β parameters were required to describe the oxide ligand, iron centre of adjacent porphyrin, CQ atoms and all other atoms (β_{oxo}, β_{Fe}, β_{CQ} and β respectively). In the case of the docked structure, the same fitting

parameters were used with the exception that $\alpha(2)_{wat}$ only described two water molecules and an additional water atom on the unligated face was also included and represented by $\alpha(1)_{wat}$. The latter was not included in the fit for the π-stacked species as one ligated face interacts with CQ and the other was likely π-stacked with another CQ-ferrihaem complex to give a 2:4 aggregate (see Section 3.3.3) and thus a water molecule in this position would not be expected for this conformation. Table 6.9 lists the fitting parameters for the π-stacked and docked conformations including solvation and CQ atoms as well as the docked structure excluding CQ and water atoms. Additional statistics generated from EXAFS fits are collected in Appendix C Table C2.

Table 6.9. *R*-factor and scaling parameters for EXAFS spectra of the CQ-ferrihaem complex in π-stacked and docked conformations. [a]

Parameter	Fitted value		
	π-Stacked [d]	**Docked** [e]	**Docked** [f]
R-factor	0.089	0.009	0.100
α	–0.02(2)	–0.004(4)	–0.013(7)
α_{oxo}	0.04(6)	–0.03(1)	–0.04(2)
α_{Fe}	–0.01(9)	–0.05(3)	–0.06(3)
α_{CQ}	0.0(2)	0.08(4)	– [g]
$\alpha(1)_{wat}$ [b]	– [g]	–0.279(9)	– [g]
$\alpha(2)_{wat}$ [c]	–0.04(4)	–0.23(1)	– [g]
β	1(3)	1(1)	1(1)
β_{oxo}	2(8)	1.5(8)	1.70(1)
β_{Fe}	1(12)	1.5(8)	1.70(1)
β_{CQ}	0.005(3) [h]	0.3(2)	– [g]

[a] Errors reported in parenthesis; [b] Water molecule situated on unligated face; [c] Water molecules situated around μ-oxo bridge; [d] Fitted using the MD data including three water molecules at positions identified from the SDF and N1, C2 and C8A atoms of CQ; [e] Fitted using the MD data including solvent water molecules and CQ N1, C2 and C8A atoms at positions identified from the SDF; [f] Fitted using the MD data excluding solvent water molecules and CQ N1, C2 and C8A atoms; [g] Not applicable; [h] value corresponds to σ^2_{CQ}.

In general fitted EXAFS distances agree with those determined from MD simulation of the docked conformation, with the majority falling within one standard deviation (see Table 6.10). As observed for the μ-oxo dimer in Section 5.5.1, a single water molecule is positioned around 2.4 Å from the iron centre on the unligated face and thus suggests an equilibrium between five- and six-coordinate species. The second hydration site around the μ-oxo ligand is slightly closer to the iron than that

predicted from SDF calculations but is well within the SDF isosurface determined for the complex (see Fig. 6.16a). This is in contrast to that seen for the μ-oxo species in the absence of CQ where water molecules were positioned somewhat further away (3.48 Å). Fitted distances from the iron centre to both C2 and C8A atoms of CQ also agree well with MD values. Fitted Fe-N1 distances are slightly longer than that predicted from MD simulation, however, are reasonably well sampled throughout the simulation (see Fig. 1.16b).

Table 6.10. Comparison of fitted EXAFS interatomic distances and those obtained from the MD simulation of CQ-ferrihaem in aqueous solution. [a]

Distance	EXAFS [b]	MD [c]
Fe–O_{oxo}	1.88(2)	1.93(4)
Fe–N	2.099(9)	2.13(3)
Fe–C_{α}	3.13(1)	3.14(4)
Fe–C_{β}	4.39(2)	4.40(5)
Fe–C_{meso}	3.44(2)	3.48(5)
Fe–Fe	3.7(1)	3.86(5)
Fe–$O(1)_{wat}$ [d]	2.44(3)	3.67 [f]
Fe–$O(2)_{wat}$ [e]	2.97(4)	3.39 [f]
Fe-N_{CQ}	3.6(1)	3.37(9)
Fe-$C2_{CQ}$	4.1(1)	3.9(1)
Fe-$C8A_{CQ}$	4.2(2)	4.2(1)

[a] Values are averages with standard deviations in parentheses for distances between Fe and all atoms of the type indicated; [b] frozen aqueous solution in 20% (v/v) aq. PEG400, pH 7.5; [c] structure most closely resembling the average structure from 10 ns MD simulation in TIP4P-Ew water box, including three water molecules and three CQ atoms; [d] one water molecule positioned on the unligated face of a given ferrihaem molecule; [e] two water molecules surrounding the bridging oxide ligand; [f] initial water positions identified from the points of highest probability in the SDF.

Independent experimental measurements in the form of EXAFS spectroscopy and ^{1}H NMR T_1 relaxation times both strongly supported the docked model of the CQ-ferrihaem complex in aqueous solution. Importantly, the methods used to probe the complex were not only independent of each other but were also both independent of the MD simulations used to model the docked and π-stacked conformations. This is in contrast to studies such as that reported by Leed *et al*.[172] where MD simulations were constrained to NMR-determined distances. Thus the formation of the docked CQ-ferrihaem complex in aqueous solution rather than the π-stacked species seems highly likely.

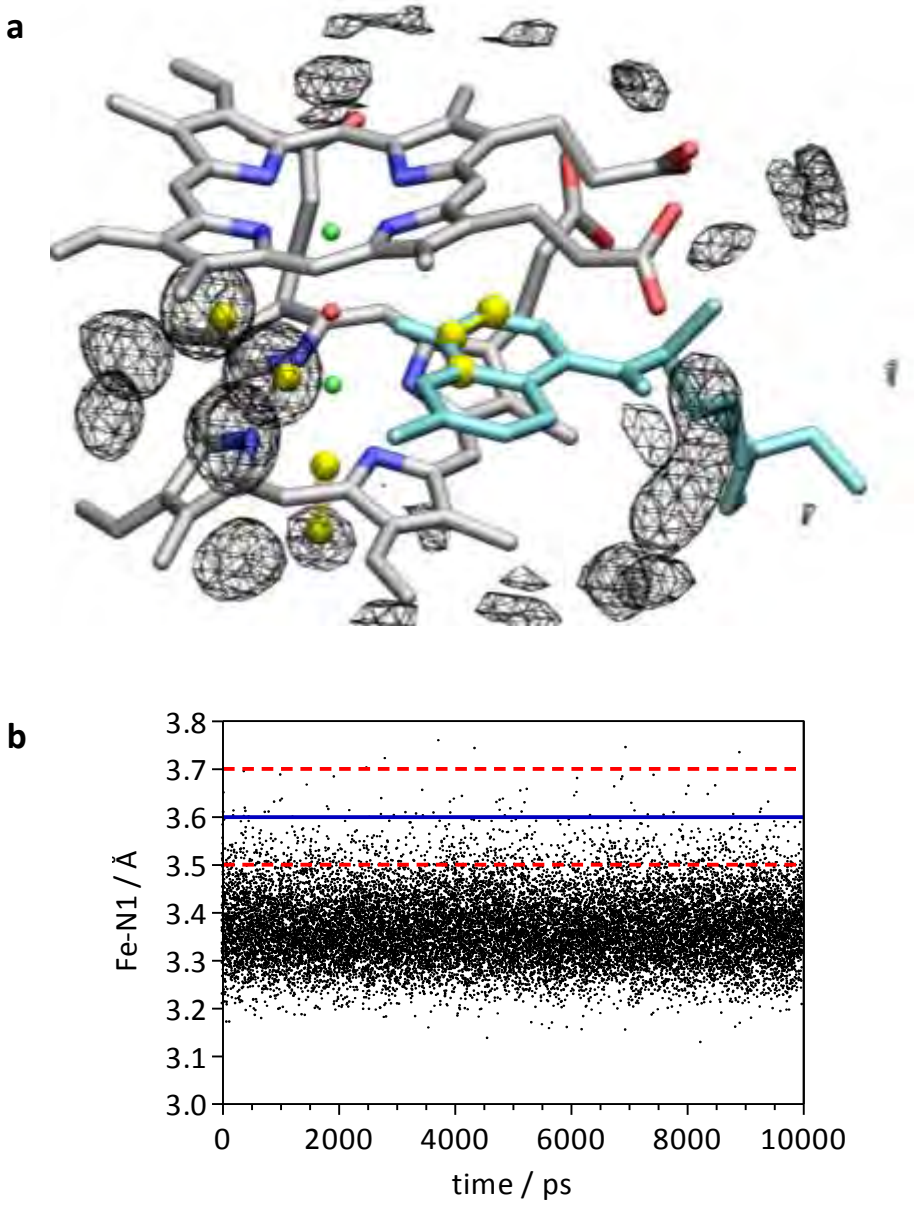


Figure 6.16. (**a**) CQ-ferrihaem complex used to fit the EXAFS spectrum. Yellow spheres around the μ-oxo bridge represent water and CQ atoms at distances determined from fitting. The water molecule fitted at 2.4 Å likely corresponds to an equilibrium between coordinated water at 2.1 Å and outer sphere solvation at 3.6 Å which has been represented by the two yellow spheres on the lower face. The SDF isosurface for oxygen atoms of water at 50% greater density than bulk is represented by the black mesh. CQ is coloured cyan and solvation of the upper face omitted for clarity. (**b**) Average Fe-N1 distance over 10 ns of MD simulation (black circles). Fitted Fe-N1 distance (blue line) and one standard deviation (red dashed lines) are overlaid.

6.4.2 Dynamic structure of the docked CQ-ferrihaem complex.

The structure of the docked CQ-ferrihaem complex over the course of the simulation was analysed in a similar manner to that used to describe the dynamic structure of π-π dimeric ferrihaem and [μ-oxo ferrihaem]$^{4-}$ species (see Sections 5.4.2 and 5.4.3). This included determining the average interplanar distance between CQ and ferrihaem porphyrins; the lateral shift of CQ relative to the centre of mass of the porphyrin rings (Fig. 6.16a); the hydrogen bond angle between N1-H1 of CQ and oxide bridge of [μ-oxo ferrihaem]$^{4-}$ (Fig. 6.16b); the interplanar angle between porphyrins (Fig. 6.16c) as well as between CQ and porphyrin rings (Fig. 6.16d); and the planar torsion angle between porphyrin rings (Fig. 6.16e), and CQ and the porphyrin rings (Fig. 6.16f).

The docked conformation of the CQ-ferrihaem complex consisted of CQ positioned between the porphyrin rings of [μ-oxo ferrihaem]$^{4-}$. The average interplanar distance of CQ from each porphyrin ring was 3.2 ± 0.2 Å which corresponded closely to that observed between π-stacking porphyrin rings of π-π dimeric ferrihaem species (3.4 ± 0.1 Å). This suggested that CQ was able to π-stack with both porphyrin rings when docked between the two. The lateral shift of CQ was very similar with respect to both porphyrin rings (Fig. 6.16a), where an average value of approx. 4.2 ± 0.2 Å was observed. The average lateral shift between porphyrin rings of [μ-oxo ferrihaem]$^{4-}$ in the complex was slightly smaller than the corresponding value observed for the species in the absence of CQ (0.8 ± 0.2 Å vs. 1.1 ± 0.2 Å) but still covered a narrow range of between 0.2 and 1.5 Å. The hydrogen bond angle between CQ and the μ-oxo bridge covered a wide range between 140 and 180°. This indicated the hydrogen bonding interaction was not locked in one specific conformation, although from Fig. 6.16b and Fig. 6.17a, angles closer to linearity were seemingly preferred. The average angle calculated over the course of the simulation was 169 ± 6°, while the average H1-O_{oxo} distance was 1.72 ± 0.09 Å. According to Jeffrey,[364] these values fall within those expected for a moderately strong hydrogen bond that is mostly electrostatic in nature (1.5 – 2.2 Å, > 130°). This distance compares well with the average reported between an oxygen atom of water and pyridinium hydrogen atom in the solid state (1.78 ± 0.01 Å).[365] The average interplanar angle between porphyrin rings of [μ-oxo ferrihaem]$^{4-}$ in the presence of CQ was markedly smaller than the corresponding angle in the absence of CQ (approx. 15° vs. 23°). Interestingly, the interplanar angle between CQ and porphyrin rings of [μ-oxo ferrihaem]$^{4-}$ showed slight differences. Different angles with each porphyrin were observed which had average values of 9 ± 6 and 13 ± 6°.

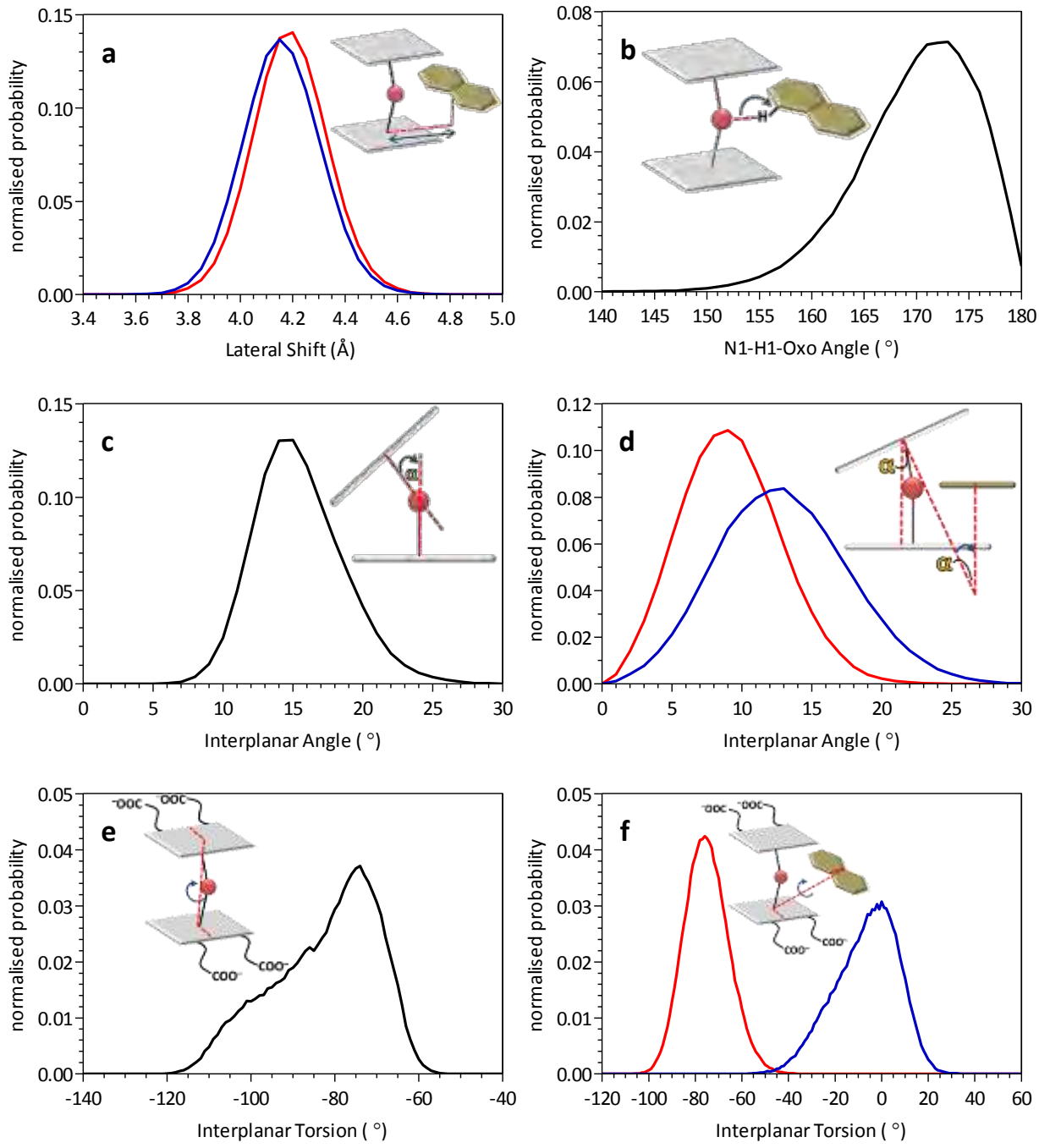


Figure 6.16. (**a**) Lateral shift histogram of CQ relative to the centre of mass of porphyrin rings. (**b**) Histogram of hydrogen bond angle between the oxide bridge of [μ-oxo ferrihaem]$^{4-}$ and N1-H1 of CQ. (**c**) Histogram of interplanar angle between porphyrin rings of [μ-oxo ferrihaem]$^{4-}$ in the CQ-ferrihaem complex. (**d**) Histogram of interplanar angle between CQ and the two porphyrin rings of [μ-oxo ferrihaem]$^{4-}$. (**e**) Histogram of interplanar torsion between porphyrin rings of [μ-oxo ferrihaem]$^{4-}$ in the CQ-ferrihaem complex. (**f**) Histogram of interplanar torsion between CQ and porphyrin ring of [μ-oxo ferrihaem]$^{4-}$ in the CQ-ferrihaem complex. All interactions between CQ and a particular porphyrin ring are represented by the same colour (either blue or red).

A notable difference in structure between [μ-oxo ferrihaem]$^{4-}$ in the presence and absence of CQ was the interplanar torsion angle. In the case of the of the latter, the torsion angle was approx. 180° and indicated that propionate side chains on one porphyrin were orientated anti to those on the other. By contrast, propionate side chains of [μ-oxo ferrihaem]$^{4-}$ in the presence of CQ had torsion angles of between −120 and −60° and averaged −83 ± 12° (see Fig. 6.16e). From the time series in Fig. 6.17b, it was observed that orientations in fact rapidly alternated between those having a torsion angle of approx. −75° and those with torsion angles of approx. −110°. The time series of torsion angles of CQ relative to porphyrin rings indicated that one ferrihaem moiety had less variation in torsion angle (Fig. 6.17c) relative to the other (Fig. 6.17d). Indeed the latter porphyrin system also showed a wider range of torsion angle (Fig. 6.16d).

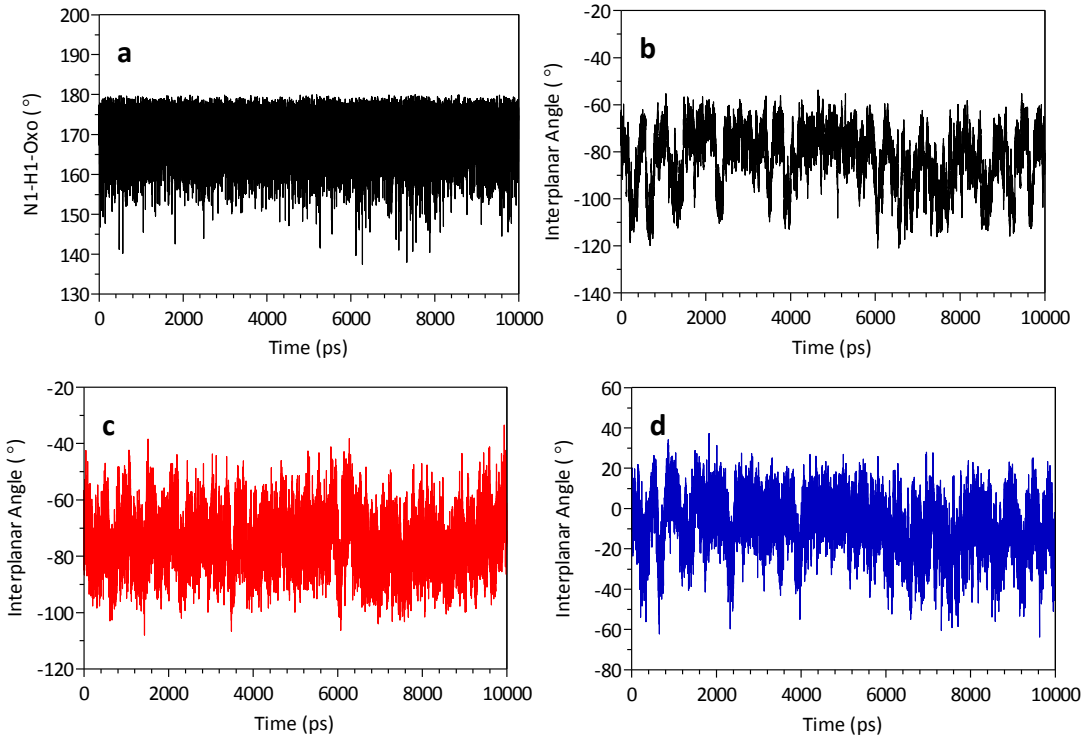


Figure 6.17. (**a**) Time series of N1-H1-O_{oxo} angle. (**b**) Time series of interplanar torsion angle between porphyrin rings of [μ-oxo ferrihaem]$^{4-}$ in the CQ-ferrihaem complex. (**c** and **d**) Interplanar torsion angle between CQ and porphyrin rings of [μ-oxo ferrihaem]$^{4-}$ in the CQ-ferrihaem complex.

The difference in torsion angle between [μ-oxo ferrihaem]$^{4-}$ in the presence and absence of CQ was not a result of a marked change in initial starting geometry of the CQ-ferrihaem complex, which had an interplanar torsion angle of approx. –150° after minimisation in TIP4P-Ew water. Thus the more gauche orientations of the propionate side chains were formed during the equilibration stage of the MD simulation. These conformations were likely favoured because interactions between propionate groups and CQ side chain and/or H4 atom were possible. Furthermore, as described for similar orientations of the π-π dimeric species, these conformations facilitate solvation shell sharing between propionate groups and CQ which minimise the entropy loss associated with solvent ordering around these groups.

6.4.3 Aggregation state of the docked CQ-ferrihaem complex.

To investigate the aggregation state of the CQ-ferrihaem complex, a plot was created of the logarithm of experimental diffusion coefficient, l*ogD*$_{exp}$, against the logarithm of diffusion coefficients calculated from MD simulation, *logD*$_{calc}$ (see Fig. 6.18). These values included those reported in Section 5.4.5 for monomeric ferrihaem species, [(HO-ferrihaem)$_2$]$^{4-}$ and [μ-oxo ferrihaem]$^{4-}$ in the absence of CQ. To increase the number of number of data points, diffusion coefficients were calculated from MD simulation for CQ and benzene and related to their corresponding experimental values. The experimental value of benzene used in the plot was taken from that reported by Niesner and Heintz,[335] while that of CQ was measured using the diffusion cell method reported in Section 3.2.8. A linear correlation between *logD*$_{exp}$ and *logD*$_{calc}$ was obtained. This supported the experimental finding that the CQ-ferrihaem complex did not exist as large aggregates in aqueous solution. Incorporating the calculated diffusion coefficient of a 1:2 CQ-ferrihaem complex into the plot showed the value was outside the 95% confidence interval of the linear regression (dashed lines in Fig. 6.18). This indicated that it was unlikely that the CQ-ferrihaem complex exists in the 1:2 form. By contrast, when the tetrameric complex (2:4), [CQ-(μ-oxo ferrihaem)]$_2$, was modelled using MD simulation and the diffusion coefficient for this species calculated, it fell within the 95% confidence interval. The structure of [CQ-(μ-oxo ferrihaem)]$_2$ is shown in Fig. 6.19 and diffusion coefficients used to construct the plot in Fig. 6.18 are reported in Table 6.11.

Table 6.11. Experimental and calculated diffusion coefficients, D_{exp} and D_{calc}, ($\times 10^{-10}$ m^2 s^{-1}) used to construct the linear relationship in Fig. 6.18. [a]

Compound	D_{exp}	D_{calc}	$logD_{exp}$	$logD_{calc}$
Benzene	10.3(2) [b]	13.2(9)	-8.99(1)	-8.88(3)
CQ	3.3(3) [c]	6.6(3)	-9.48(4)	-9.18(2)
Ferrihaem monomer	2.2(2) [d, e]	4.1(2) [f]	-9.66(4)	-9.39(2)
[HO-/HO- ferrihaem] $^{4-}$	1.4(1) [e]	3.1(1)	-9.85(3)	-9.51(1)
[μ-Oxo ferrihaem] $^{4-}$	1.6(1) [e]	3.0(2)	-9.80(3)	-9.52(3)
CQ-ferrihaem (1:2)	0.6(2) [c]	2.6(1)	-10.2(1)	-9.59(2)
CQ-ferrihaem (2:4)		2.0(1)		-9.70(2)

[a] Values are averages with standard deviations in parentheses; [b] From Niesner and Heintz[335]; [c] determined in this study using the diffusion cell method; [d] $(CN)_2$-ferrihaem; [e] from de Villiers *et al.*[75]; [f] average value computed for three monomeric ferrihaem species (see Section 5.4.5).

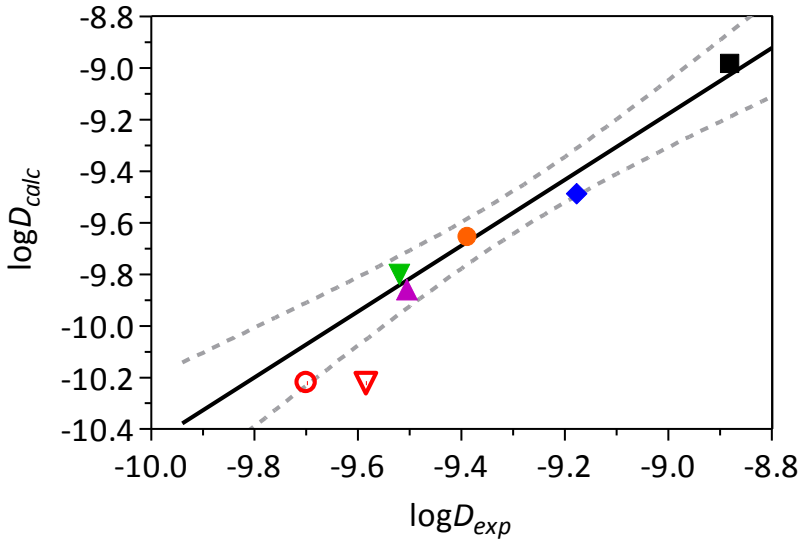


Figure 6.18. Plot of the logarithm of calculated diffusion coefficient, $logD_{calc}$, against the logarithm of experimental diffusion coefficient, $logD_{exp}$. Linear regression (solid line) was calculated using filled data points and is described by $logD_{calc}$ = 1.3(1) $logD_{exp}$ − 2(1); r^2 = 0.98, P-stat = 0.001. The 95% confidence interval is represented by grey dashed lines. Benzene, CQ, monomeric ferrihaem, $[(HO\text{-ferrihaem})_2]^{4-}$ and [μ-oxo ferrihaem]$^{4-}$ are represented by the filled black square, blue diamond, orange circle, purple triangle and green triangle respectively. Unfilled triangle and circle represent the 1:2 and 2:4 CQ-ferrihaem complexes respectively.

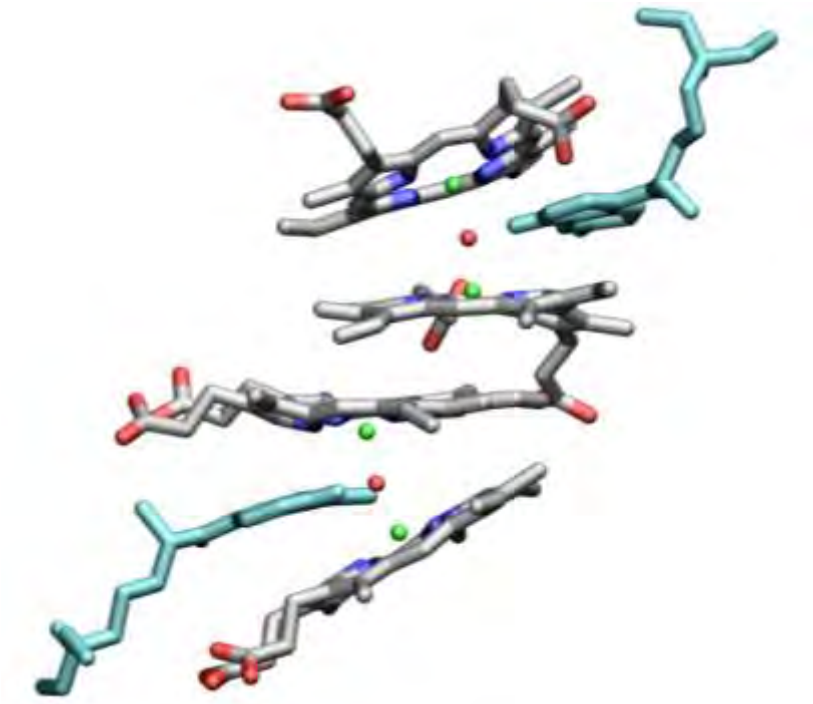


Figure 6.19. Structure of the tetrameric 2:4 CQ-ferrihaem complex, [CQ-(μ-oxo ferrihaem)]$_2$, that likely exists in aqueous solution. CQ and porphyrin hydrogen atoms have been removed for clarity and CQ coloured in cyan to aid visualisation.

While the finding that the complex exists in a tetrameric [CQ-(μ-oxo ferrihaem)]$_2$ form in aqueous solution is somewhat uncertain given the errors in the experimental and computed diffusion coefficients, empirical evidence from Section 3.3.3 suggested the same aggregation state as being the most likely to exist in solution. Furthermore, mass spectrometry findings reported by Schwedhelm *et al*.[170] also implicated a stoichiometry of 2:4 CQ:ferrihaem in solution. Given these three sets of findings, it seems likely that the [CQ-(μ-oxo ferrihaem)]$_2$ species is the dominant form of the CQ-ferrihaem complex in aqueous solution.

Structural analysis of the 2:4 CQ:ferrihaem adduct showed that out of plane (OOP) distortion (doming) of the iron centre of the two outer porphyrin rings was similar to that observed for μ-oxo ferrihaem (see Fig. 6.20a). In the latter molecule, this displacement corresponded to relatively minor doming of the porphyrin system and presumably favoured the coordination of water to produce an equilibrium between five- and six-coordinate states. Given the similarities in iron OOP deviations, an analogous equilibrium likely exists for the 2:4 CQ:ferrihaem adduct. Conversely, significantly larger iron OOP deviations and consequently greater distortion of the porphyrin rings were observed for the 1:2 CQ:ferrihaem species. Interestingly, while the angle between CQ and porphyrin planes in the 1:2 CQ:ferrihaem species was similar for both porphyrins, in the 2:4 CQ:ferrihaem adduct there was a marked increase in interplanar angle between CQ and the porphyrins partaking in π-stacking interactions (Fig. 6.20b). On the other hand, there was little difference observed in interplanar angle for the outer porphyrin rings in direct contact with solvent (Fig. 6.20c). This suggests that there is a preference for CQ to interact with the outer porphyrin system (presumably through π-stacking interactions) over the porphyrin which π-stacks with a neighbouring 1:2 CQ:ferrihaem complex.

The observations from MD simulation relating to iron OOP deviation may explain why the 2:4 CQ:ferrihaem species is seemingly preferred in aqueous solution. The weak doming of the porphyrin ring in μ-oxo ferrihaem presumably shifts the equilibrium between five- and six-coordinate species towards the latter and thus, as detailed in Section 5.5.1, prevents aggregation of μ-oxo dimers in aqueous solution. By extension the 2:4 CQ:ferrihaem adduct would similarly not aggregate. In the case of the 1:2 CQ:ferrihaem complex, distortion of the porphyrin ring probably disfavours ligation of water and thus a shift in the equilibrium towards five-coordinate species occurs. In this form, π-stacking interactions are favoured since the unligated face is no longer strongly hydrated and thus will likely π-stack with another five-coordinate 1:2 CQ:ferrihaem species to form a 2:4 CQ:ferrihaem adduct. Upon formation, the 2:4 CQ:ferrihaem complex the outer porphyrin rings would become less domed, shift the equilibrium in favour of six-coordinate species once again and thus prevent further aggregation. This process has been summarised in Fig. 6.21.

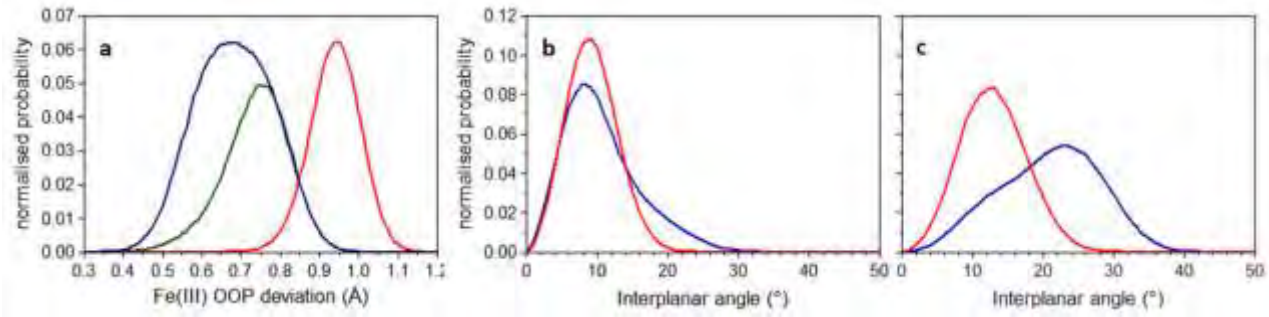


Figure 6.20. Structural histograms calculated for μ-oxo ferrihaem (green) and the 1:2 and 2:4 CQ:ferrihaem complexes (red and blue respectively). (**a**) Average deviation of the iron centre from the porphyrin plane; (**b**) Average interplanar angle between CQ and the porphyrin ring corresponding to the outer face in the 2:4 complex; (**c**) Average interplanar angle between CQ and the porphyrin ring which π-stacks with a neighbouring complex in the 2:4 adduct.

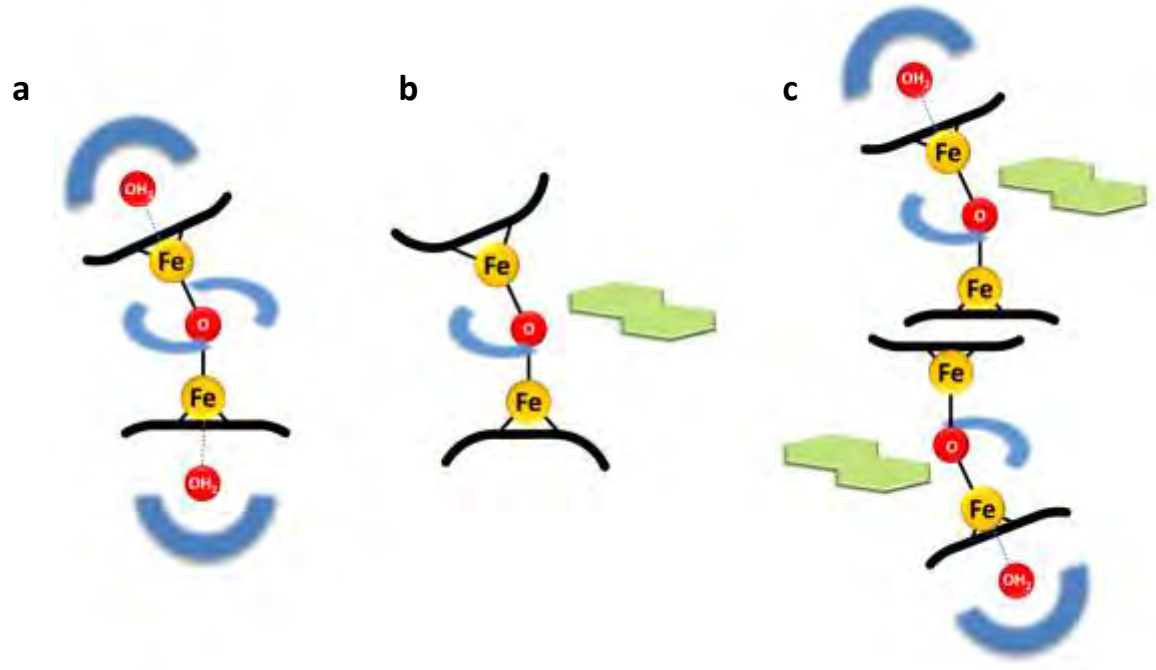


Figure 6.21. Schematic depicting the proposed preference for 2:4 CQ:ferrihaem in aqueous solution. (**a**) Six-coordinate μ-oxo dimer is favoured by relatively little doming of porphyrin rings and disfavours aggregation. (**b**) Docking of CQ between porphyrin rings of μ-oxo dimer enhances doming of 1:2 CQ:ferrihaem complex and favours the five-coordinate species. (**c**) Five-coordinate species π-stack forming a tetrameric 2:4 CQ:ferrihaem adduct. Porphyrin rings become less domed allowing coordination of water and formation of six-coordinate species which disfavour further aggregation. Blurred semi-circles represent water solvation.

6.5 Comparison of the Docked CQ-Ferrihaem Model to Literature

The docked structure of the CQ-ferrihaem complex is the most reliable model proposed to date. This is because it was identified using a multi-technique approach in conjunction with computational simulation. Furthermore, it was verified by two independent experimental techniques. The structure itself is markedly different from those previously proposed for the CQ-ferrihaem complex (see Fig. 1.22 for examples). Complexes proposed by Moreau *et al*.[60, 167], Webster *et al*.[171] and de Dios *et al*.[174] can be likely discounted based on their inconsistency with experimentally determined binding stoichiometries of 1:2 CQ:ferrihaem. In the case of the former (Fig. 1.22a), a 1:4 CQ:ferrihaem complex was proposed where one CQ is sandwiched between two μ-oxo dimers. The latter two authors proposed CQ-ferrihaem complexes having a 1:1 binding ratio with Webster *et al*. suggesting two CQ molecules capping a μ-oxo dimer through π-stacking interactions (Fig. 1.22c), while de Dios *et al*. proposed CQ bound to monomeric ferrihaem through its quinoline nitrogen atom (Fig. 1.22f). The structure suggested by de Dios *et al*. can be further discounted on the basis of magnetic susceptibility measurements and binding stoichiometries which suggest ferrihaem is in the μ-oxo dimeric form. Experimental magnetic susceptibility findings also discount the structure proposed by Constantinidis and Satterlee[164] who suggested CQ interacts with π-π dimeric urohaemin. While the model proposed by Schwedhelm *et al*.[170] displayed the correct binding ratio as well as aggregation state, it was inconsistent with IR evidence. This structure involved two CQ molecules capping two μ-oxo dimers which were reciprocally bound by propionate side chains (Fig. 1.22b), however, the characteristic coordinated propionate IR stretching peaks similar to that observed for haemozoin[126] have never been observed in the IR spectrum of the complex. Furthermore, the explanation of the low magnetic moment proposed by these authors is invalid.

The remaining models proposed by Leed *et al*.[172], Natarajan *et al*.[173] and Portela *et al*.[175, 176] all involved CQ π-stacking with the unligated face of μ-oxo ferrihaem (Fig. 1.22d, e, g and h). This structure was not consistent with ^{1}H NMR T_1 relaxation times or the experimental EXAFS spectrum of the CQ-ferrihaem complex. Moreover, to explain the 1:2 binding stoichiometry, these authors suggested that two CQ molecules cap the unligated faces of two π-stacking μ-oxo dimers. However, in the presence of excess CQ, one would expect the μ-oxo dimers in this structure to disaggregate and CQ to cap both unligated faces to give a 1:1 binding stoichiometry. This is not observed experimentally and thus provides further evidence against these models.

In contrast to the previously proposed models, the docked CQ-(μ-oxo ferrihaem) complex was consistent with all experimental data. The model supports the 1:2 CQ:ferrihaem binding stoichiometry even in an excess of CQ since the structure can support only one CQ molecule per μ-oxo dimer. The low magnetic susceptibility was consistent with the μ-oxo dimeric form of ferrihaem and the experimental EXAFS spectrum gave an excellent fit using the solvated structure. In addition, experimental T_1 relaxation times were well correlated with the MD model using the Solomon-Bloembergen relationship. Importantly, this was the only model to maintain correlation using solely the aromatic protons of the quinoline ring.

Closer analysis of the docked complex revealed three interactions between CQ and μ-oxo ferrihaem which were likely crucial for complex formation. The first was the formation of an intramolecular hydrogen bond between the protonated quinoline nitrogen atom and μ-oxo bridge; the second was π-stacking between the quinoline ring of CQ and porphyrin system of ferrihaem; and the third was an interaction between the 7-chloro substituent and hydrogen atoms of a methyl group on the porphyrin ring system which formed close contacts within 3 Å (sum of van der Waals radii of Cl and H, 1.8 Å + 1.2 Å, respectively) in approx. 30% of the structures over the course of the MD simulation. These interactions are highlighted in Fig. 6.22. The former two interactions (i) and (ii) are well known in the literature and while the latter type (iii) is less common, much work has been devoted to investigating halogen acceptors in hydrogen bonding.[366-368] An example of this class of interaction has been reported by Gibb and co-workers for the encapsulation of halogenated adamantine molecules in resorcinarene "molecular baskets".[369, 370] Interestingly, the same halogen-hydrogen bond was also implicated in the haemozoin-inhibition model proposed by Buller *et al*.[148] Unlike that reported by Natarajan *et al*.[173], no substantial interaction was observed between propionate side chains of ferrihaem and the tertiary nitrogen group of the CQ side chain. Instead, SDF calculations generated in this work indicated this moiety as well as the 4-amino group is strongly hydrated and thus it is likely that interactions may rather occur through solvation shell sharing between CQ and ferrihaem groups.

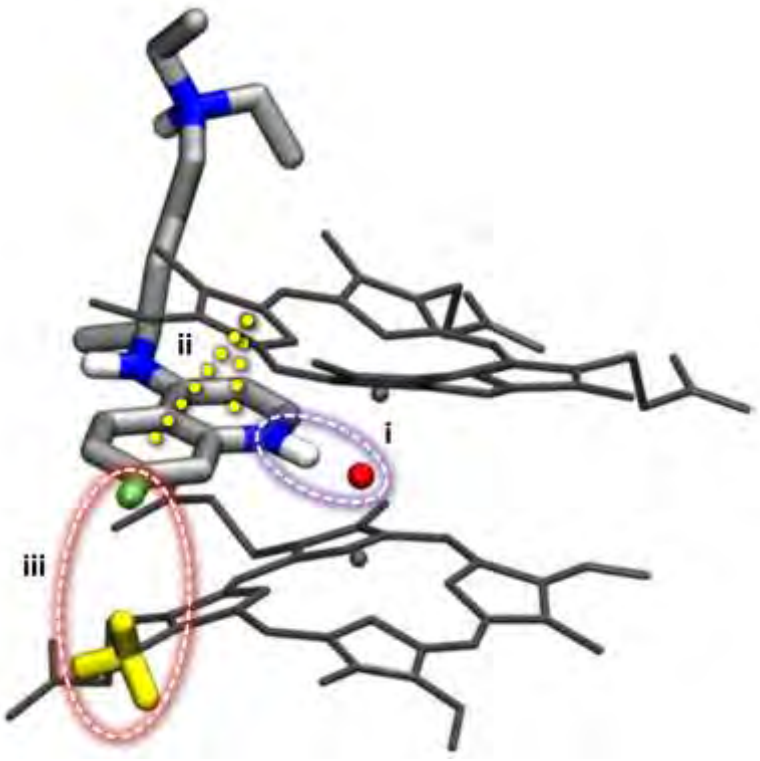


Figure 6.22. Interactions identified in the docked CQ-(μ-oxo ferrihaem) complex. Only a selection of hydrogen atoms are displayed for clarity and the μ-oxo dimer is coloured in grey to aid visualisation. (**i**) Hydrogen bonding between the μ-oxo bridge and protonated quinoline nitrogen (purple dashes). Average hydrogen bond distance (H···O) is 1.72 ± 0.09 Å. (**ii**) π-Stacking interactions between aromatic quinoline ring and CQ and porphyrin system of ferrihaem (yellow dots). Average interplanar distance between CQ and ferrihaem porphyrins is 3.4 ± 0.1 Å. (**iii**) Close contacts (< 3.0 Å) made between 7-chloro substituent of CQ and methyl protons (coloured yellow for clarity) of the porphyrin ring (red dashes). Average distance for structures displaying close contacts is 2.8 ± 0.1 Å.

The three interactions identified could be adapted into a SAR model which could be used to establish criteria for molecules binding to ferrihaem in the same manner as CQ. The SAR model is summarised in Fig. 6.23 and features discussed in Sections 6.5.1 and 6.5.2.

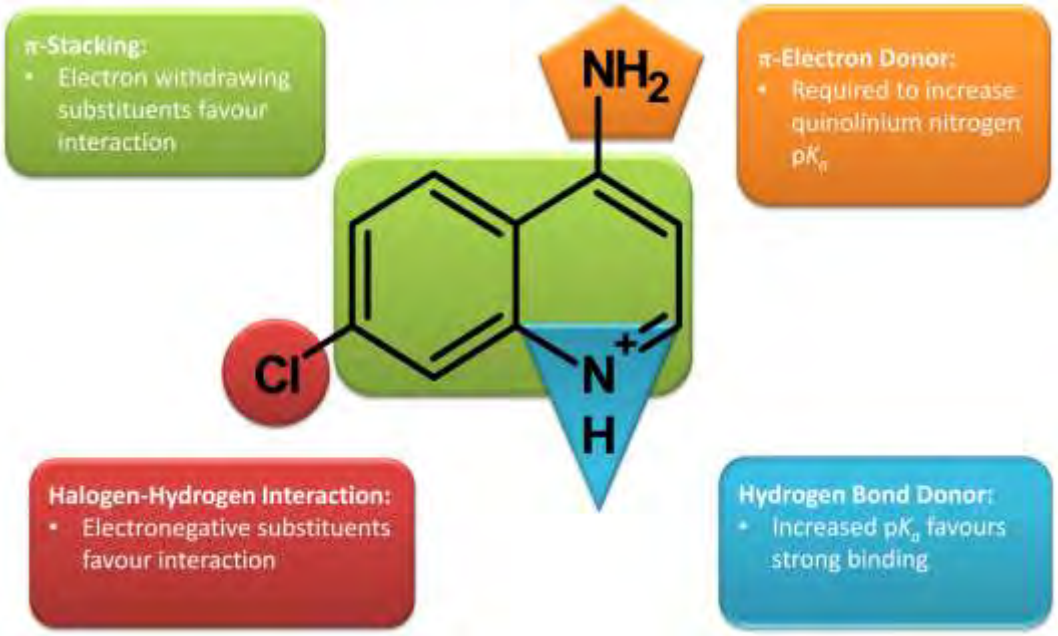


Figure 6.23. Proposed SAR model for CQ binding to μ-oxo ferrihaem in a docked conformation. The 4-aminoquinolinium moiety of CQ has been used as it is the smallest fragment displaying the three important interactions identified in the CQ-(μ-oxo ferrihaem) complex (see Fig. 6.22).

6.5.1 Hydrogen bonding and pK_a of quinolinium nitrogen.

The hydrogen bond between quinolinium and μ-oxo bridge is a crucial interaction required to form the docked complex. A number of features of the proposed model can affect the formation and strength of this hydrogen bond. The first is the presence of the quinoline nitrogen atom itself (see blue triangle in Fig. 6.23). Without this hydrogen bond donor, no interaction can occur and thus removal of this atom would likely have a detrimental effect on ferrihaem association. Experimental precedence for this proposed hypothesis has been reported by Cheruku *et al.*[159] where a CQ isostere synthesised with a CH group replacing the quinoline nitrogen failed to produce any binding affinity with ferrihaem (Fig. 6.24). Further evidence to support this interaction was reported by Nsumiwa *et al.*[151] where SARs involving atomic charges computed using DFT and ferrihaem association constants were only observed for quinolinium species and not the quinoline free base.

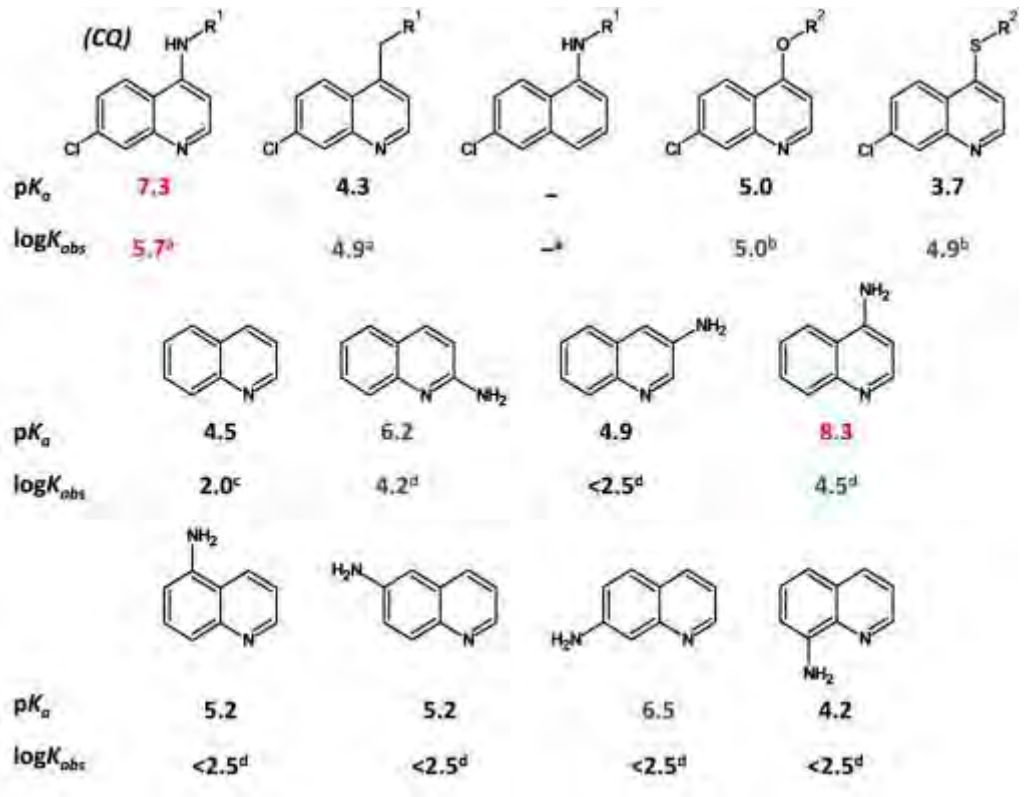


Figure 6.24. CQ analogues with corresponding quinoline nitrogen pK_a values (predicted using MarvinSketch[371]) and association constants (K_{obs}) with ferrihaem as reported by (a) Cheruku *et al*.[159]; (b) Natarajan *et al*.[173]; (c) Kuter *et al*.[183] and (d) Egan *et al*.[127] Red, blue and black values indicate strong, moderate and weak association constants and high, moderate and low pK_a values respectively. R^1 = $CH(CH_3)CH_2CH_2CH_2N(CH_2CH_3)_2$ and R^2 = $CH_2CH_2CH_2CH_2N(CH_2CH_3)_2$.

The second feature of the SAR model relating to hydrogen bonding strength is the amino group at position 4 in the quinoline ring (see orange hexagon in Fig. 6.23). This group drastically affects the pK_a of the quinoline nitrogen atom through resonance π-electron donation. The pK_a (predicted using MarvinSketch[371]) is increased from 4.5 in quinoline to 8.3 in 4-aminoquinoline and there is a corresponding marked increase in the logarithm of the ferrihaem association constant (logK_{obs}) from 2.0 to 4.4 (Fig. 6.24).[127, 129] This increased association strength presumably occurs as a result of the stronger hydrogen bonding character of the quinoline nitrogen but also because there is a greater concentration of protonated quinolinium species which are able to bind to ferrihaem in the latter at neutral and low pH. Indeed, the effect of pK_a on ferrihaem association was also observed in the work reported by Natarajan *et al*.[173] and Cheruku *et al*.[159] in which the 4-amino substituent of CQ or CQ-analogues was replaced by carbon, oxygen or sulfur groups. These molecules all had lower predicted pK_a values than the corresponding amino counterpart and produced weaker ferrihaem association strengths (Fig. 6.24).

Positioning of the amino group in aminoquinoline compounds also has a profound effect on ferrihaem association, where the amino substituent in any position other than at carbon-2 or -4 produced low logK_{obs} values.[127] This likely occurs because these molecules have low predicted quinoline nitrogen pK_a values (< 5.2). An exception is 7-aminoquinoline which has a moderate value (6.5) but violates a feature of the proposed SAR model (see Section 6.5.2) which presumably accounts for the weaker than expected interaction with ferrihaem. In the case of 8-aminoquinoline, unfavourable steric clashes are also likely to occur between the substituent and μ-oxo bridge, thus producing a low binding affinity for ferrihaem. Indeed, it is well known that the 8-aminoquinoline antimalarial drug, primaquine, does not bind to ferrihaem,[124] presumably for these two reasons stated above. Interestingly, despite the marked difference in predicted pK_a values between 2- and 4-aminoquinoline (6.2 vs. 8.3, respectively), both produce association constants of similar strength (logK_{obs} = 4.2 vs. 4.4). Based on the discussion above, the former molecule would be expected to give a somewhat weaker association constant than the latter. However, given the proximity of the amino group in the 2-position, an additional hydrogen bond between amino and μ-oxo moieties is feasible. Thus the formation of a bidentate hydrogen-bonded structure may be the cause of the unexpectedly stronger ferrihaem association constant of 2-aminoquinoline.

6.5.2 Quinoline 7-position substitution and π-stacking interactions.

The final two features of the SAR model are somewhat related. They involve π-stacking interactions between quinoline ring and porphyrin system of ferrihaem (see green rectangle in Fig. 6.23) and close contacts between electronegative substituents in the 7-position and methyl protons of the porphyrin ring (see red oval in Fig. 6.23). This is because π-stacking interactions are known to be favoured by the inclusion of electron withdrawing substituents such as halogen, nitro or cyano groups. Those with negative potentials would thus also favour the close contact interaction. Conversely, π-stacking interactions are destabilised by electron donating substituents such an amino group owing to increased electron repulsion between π-systems.[372, 373] The close contact feature would also be less favourable in the case of groups having greater positive potential.

There is much experimental support for the proposed π-stacking and close contact features. Egan and co-workers[120, 127, 151] as well as Vippagunta *et al*.[374] have noted electron withdrawing substituents in the 7-position generally produce stronger association constants with ferrihaem. For example, 4-amino-7-chloroquinoline (Fig. 6.25) has a ferrihaem association constant roughly 1.6 log units larger than that of 4,7-diaminoquinoline (Fig. 6.24). Support for the close contact feature is also seen in the findings of the above-mentioned authors. In the case of Vippagunta *et al*.[374], strong binding was observed for CQ which has a chlorine atom in the 7-position, however, the analogue in which the chlorine atom was synthesised in the 6-position produced no detectable ferrihaem association. In the case of Nsumiwa *et al*.[151], strong correlations were produced between 7-substituted CQ-analogues and the DFT-calculated atomic charge of the substituent in the 7-position, where groups with greater negative potential favoured strong ferrihaem binding.

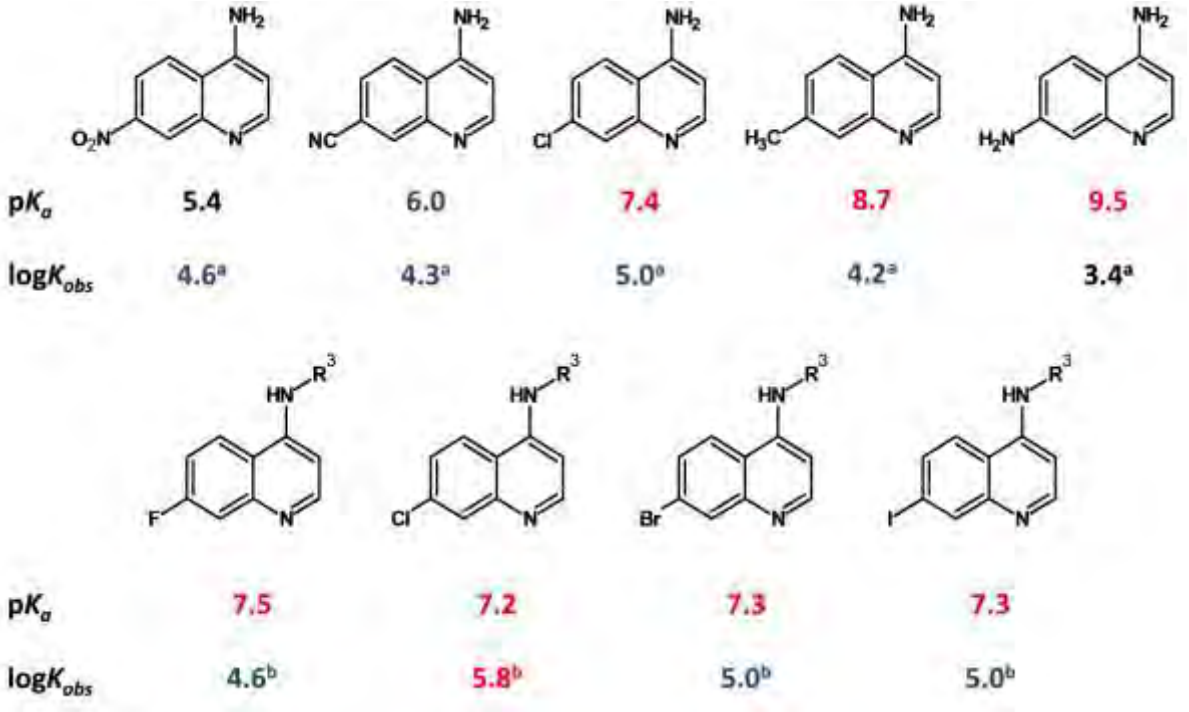


Figure 6.25. CQ analogues with corresponding quinoline nitrogen pK_a values (predicted using MarvinSketch[371]) and association constants (K_{obs}) with ferrihaem as reported by (a) Nsumiwa *et al*.[151] and (b) Kaschula *et al*.[120] Red, blue and black values indicate strong, moderate and weak association constants and high, moderate and low pK_a values respectively. $R^3 = CH_2CH_2N(CH_2CH_3)_2$.

While the features of the SAR model may at first glance seem straight-forward, there is in fact a complex interplay between them. This is because the substituent at the 7-position not only affects the close contact and π-stacking interactions but also influences the pK_a of the quinolinium nitrogen. For example, 4-amino-7-chloroquine (Fig. 6.25) has a predicted pK_a of 7.4 while that of 4-aminoquinoline is 8.4 (Fig. 6.24). Thus, the presence of this 7-chloro substituent weakens the hydrogen bonding interaction described in Section 6.5.1 but strengthens the close contact and π-stacking interactions which results in an overall increased association strength. Another example is 4-amino-7-nitroquinoline (Fig. 6.25). While the electron withdrawing character and electronegative potential of this molecule favours π-stacking and close contact interactions, respectively, it has a predicted pK_a of 5.4 which causes substantial weakening of the hydrogen bonding feature. Consequently, its association with ferrihaem is overall weaker than that of the chloro-analogue. The same effect is seen for 4-amino-7-cyanoquinoline which has a predicted pK_a of 6.0 and a weaker association constant than would be expected based purely on π-stacking and close contact interactions (Fig. 6.25). Conversely, weakening of the π-stacking and close contact interactions weaken overall association strength even although the substituent can strengthen the hydrogen bonding interactions through increased quinolinium pK_a values. This is probably the case for 4-amino-7-methylquinoline which has a weakly inductively-donating methyl group and high predicted quinolinium nitrogen pK_a value of 8.7 (Fig. 6.25). Consequently, halogens are likely the best substituents to have in the 7-position for ferrihaem binding as they produce the strongest π-stacking and close contact interactions with least adverse effect on quinoline pK_a. Chlorine is particularly favourable since steric factors are also likely to play a role in the docked structure. Indeed, reported association constants of bromo- and iodo-substituted CQ analogues have a lower association constant than CQ presumably for this reason (Fig. 6.25).[120] By contrast, the fluoro-substituted CQ analogue produces weaker association constant (Fig. 6.25). [120] This is not unexpected, however, given the fact that it is a poor hydrogen bond acceptor arising from its weak polarisibility and suppressed lone-pair donating character.[375] On the other hand, the larger, less electronegative halogens are more polarisable and better able to donate lone pairs.

It is likely that this complex interplay between the features of the proposed SAR model is the reason why there has never been a structural model of the CQ-ferrihaem complex proposed based solely on the information obtained from SARs. Identifying the structure of the CQ-ferrihaem complex has thus allowed a means to rationalise the previously reported SAR findings.

6.6 Conclusions

MD simulations of the CQ-ferrihaem complex in aqueous solution were successfully conducted using the optimised parameters of CQ developed in this study. By combining independent experimental techniques in the form of EXAFS spectroscopy and ^{1}H NMR T_1 measurements with data obtained from MD simulation, strong support was obtained for the conformation in which CQ was situated between porphyrin rings in μ-oxo ferrihaem forming a docked complex. While the docked conformation of the CQ-(μ-oxo ferrihaem) complex displayed similarities to [μ-oxo ferrihaem]$^{4-}$, there were notable differences in both structure and solvation. Solvation surrounding the μ-oxo bridge was somewhat disrupted by the presence of CQ, leading to reduced solvent hydration in this area. The interplanar torsion angle between propionate side chains of each ferrihaem group adopted a more gauche orientation in contrast to the anti conformation observed for [μ-oxo ferrihaem]$^{4-}$. Other differences in structure included a smaller lateral shift between ferrihaem moieties and smaller interplanar angle between porphyrin rings. On the other hand, the same weak ligation by water to the sixth coordination site found for [μ-oxo ferrihaem] persisted in the CQ-(μ-oxo ferrihaem) complex. Support for the solvated structure determined from MD simulation was obtained by an excellent fit to the experimental EXAFS spectrum. Comparison of experimental and computed diffusion coefficients indicated that this conformation likely exists in a 2:4 aggregation state and supported the similar finding obtained when using the empirical relationship between diffusion coefficient and molar volume in Section 3.3.3.

Three key interactions were identified in the docked structure, namely a hydrogen bond between protonated quinoline nitrogen and μ-oxo bridge; π-stacking interactions between CQ and porphyrin ring systems; and close contacts between the chlorine atom of CQ and hydrogen atoms of a methyl substituent of ferrihaem. These interactions were developed into a SAR model to describe ferrihaem binding by CQ analogues and, importantly, could provide a comprehensive qualitative rationalisation of previously reported SAR data.

The results presented in this chapter illustrate the potential of using complimentary experimental techniques in conjunction with computational simulation to better understand the interaction of ferrihaem with compounds such as antimalarial drugs.

CHAPTER 7

OVERALL CONCLUSIONS AND FUTURE WORK

7.1 Overall Conclusions

CQ is regarded as the quintessential antimalarial drug and has been the subject of many investigations. Its mechanism of action is widely considered to involve inhibition of haemozoin formation and indeed, this has recently been shown *in vivo* where a decrease in haemozoin and corresponding increase in ferrihaem was observed.[128] While early studies indicated a strong association existed between CQ and ferrihaem,[99] there has subsequently been much controversy in the literature regarding the interaction between these two molecules. This has primarily arisen from a lack of understanding with regard to ferrihaem speciation, the frequent reliance on single experimental technique to investigate interactions and use of varying experimental conditions between studies including solvent, pH, salt concentration, etc. In the absence of a crystal structure, the exact structure of the CQ-ferrihaem complex remains inherently uncertain. However, even if a crystal structure was obtained, this structure would be highly dependent on the environment from which it was crystallised and, since it exists in the solid state, may not necessarily be representative of that found in solution. With this in mind, the only way to investigate the structure of the CQ-ferrihaem complex in solution is through the combined use of experimental and computational methods. To date, this type of approach has not been systematically applied to determine the structure of the complex. Instead, many models have been proposed often based on qualitative interpretation or intuition. A few studies have provided structures using potentially more robust methods, but these have been based largely on NMR evidence only and models have not been tested against experimental data.

The aim of the present study was to comprehensively investigate the interaction of CQ and ferrihaem in solution employing a variety of complimentary experimental techniques and computational methods. A careful and methodical approach was applied in this regard which first involved a detailed experimental investigation of the CQ-ferrihaem complex in aqueous solution to determine the stoichiometry of interaction and the species of ferrihaem. This was achieved using, amongst others, UV-visible and MCD spectroscopy supplemented by magnetic susceptibility measurements. Employing more than one technique to determine, for example, the species of ferrihaem proved crucial for a reliable interpretation of the data. This was particularly evident for the solid precipitated from aqueous solution at pH 5 where IR spectra lacked discernable peaks associated with the μ-oxo dimer, whereas magnetic susceptibility measurements convincingly showed a low magnetic moment which could only be

interpreted as evidence of the antiferromagnetically coupled μ-oxo species. Simply relying on one technique, such as IR spectroscopy in this case, would have (and has previously) led to incorrect conclusions regarding the complex. These investigations clearly showed the complex in aqueous solution occurred in a 1:2 CQ:ferrihaem stoichiometry with ferrihaem existing in the μ-oxo dimeric form. Contrary to recent reports, the same complex was observed in the solid state. Importantly, while the μ-oxo dimeric form of ferrihaem has previously been implicated in the CQ-ferrihaem complex in solution, CQ has been described as shifting the monomer-(μ-oxo dimer) equilibrium in favour of the latter. The UV-visible, MCD and magnetic susceptibility evidence unequivocally demonstrated that is not the case, but rather that CQ induces μ-oxo dimer formation in conditions where it is not otherwise present (aqueous solution).

To accurately describe the CQ-ferrihaem complex using MM and MD simulations, force fields specific for five-coordinate ferrihaem species were established using QM-derived data. *Ab initio* calculations of ferrihaem species produced structures in good agreement with related iron(III)porphyrin crystal structures and facilitated a fundamental understanding of the origins of UV-visible spectroscopic features. The newly developed ferrihaem force fields used in MD simulations of monomeric, π-π dimeric and μ-oxo dimeric ferrihaem species performed well, particularly in the case of the latter species where excellent agreement with experimental solution-phase EXAFS spectra was obtained. Subsequent MD simulation of the CQ-ferrihaem complex led to the proposal of a docked CQ-ferrihaem complex in which CQ was positioned between the porphyrin rings of μ-oxo ferrihaem. This was strongly supported by two independent experimental data sets (EXAFS and ^{1}H NMR T_1 relaxation times). Furthermore, the aggregation state of the complex was shown to likely exist in a 2:4 CQ:ferrihaem form based on an empirical relationship between diffusion coefficient and molecular size reported by Gustafson and Dickhut,[268] as well as from linear correlation of experimental and computed diffusion coefficients. A summary of the CQ-ferrihaem interaction is depicted schematically in Fig. 7.1.

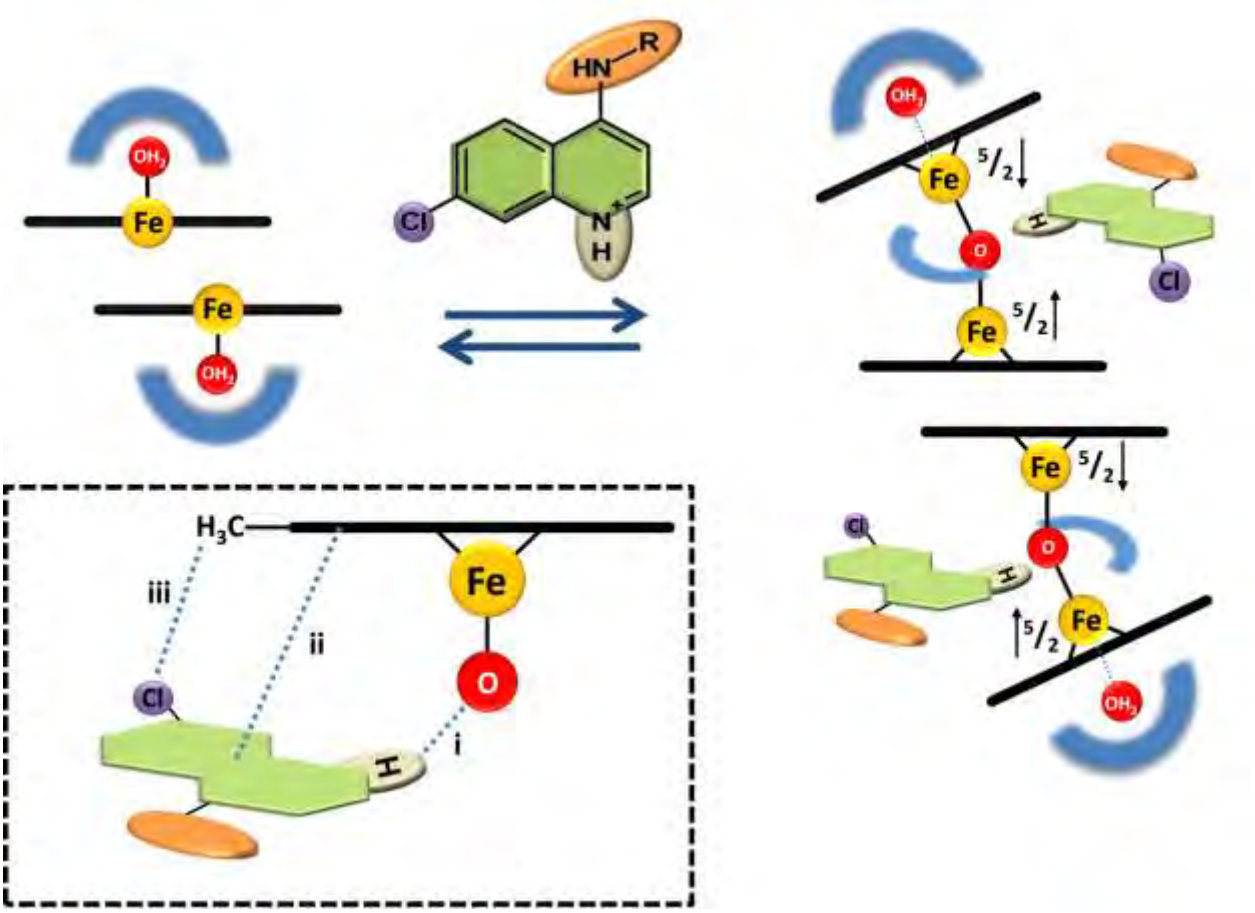


Figure 7.1. Schematic summary of the CQ-ferrihaem interaction in aqueous solution. Ferrihaem in aqueous solution exists as a π-π dimer (upper left) with axial water ligands that are highly solvated (represented by blue smearing). Addition of CQ induces formation of the antiferromagnetically coupled μ-oxo ferrihaem species (indicated by $^{5}/_{2}$ black arrows) and docks between the porphyrin planes of the μ-oxo dimer to form the CQ-(μ-oxo ferrihaem) complex. Strong hydration of the oxide bridging ligand still persists. The CQ-(μ-oxo ferrihaem) complex exists as a tetramer adduct, $[CQ\text{-}(\mu\text{-oxo ferrihaem})]_2$, in which the outer unligated faces exist in equilibrium between five- and six-coordinate states. **Inset**. An expansion of the CQ-(μ-oxo ferrihaem) complex showing three key interactions essential for ferrihaem binding in the docked conformation, including (i) an intermolecular hydrogen bond between protonated quinoline nitrogen of CQ (grey oval) and the oxide bridge of μ-oxo ferrihaem; (ii) π-stacking interactions between the aromatic quinoline ring of CQ and the porphyrin ring system of ferrihaem; and (iii) close contacts (≤ 3 Å) between the chlorine substituent of CQ and methyl hydrogen atoms on the porphyrin ring. Lower ferrihaem moiety of the μ-oxo dimer omitted for clarity.

The summary in Fig. 7.1 includes the structure of the CQ-(μ-oxo ferrihaem) complex proposed in the current study. To date, this is the most reliable model because it is able to account for experimental findings observed not only in this study, but also those reported by a number of other authors. Furthermore, this model is the first that is able to provide a comprehensive qualitative account of reported SARs of CQ derivatives based on three key structural features. These include the hydrogen bond between protonated quinoline nitrogen; π-stacking interactions between the aromatic moiety of CQ and the porphyrin ring system of ferrihaem; and close contact between the 7-chloro substituent and methyl hydrogen atoms on the porphyrin (see Fig 7.1). Strong ferrihaem association is favoured by high pK_a values of the quinolinium nitrogen so as to produce a large concentration of protonated species and stronger hydrogen bonding; electron-withdrawing groups at the 7-position to reduce electron density, thus increasing π-stacking interactions; non-bulky groups at the 8-position to enable docking of the quinoline ring; electronegative and hydrophobic substituents at the 7-position to provide favourable interaction with methyl hydrogen atoms.

Additionally, the docked model can explain an anomaly not touched upon in previously suggested π-stacked models relating to the binding stoichiometry. Experimentally, a 1:2 CQ:ferrihaem binding ratio is observed even when high concentrations of CQ are present. In the case of the π-stacked models proposed by others in which CQ caps the ends of two associated μ-oxo dimers, one would expect an excess of CQ to disaggregate the complex and form a disaggregated μ-oxo species with CQ π-stacking on both unligated faces. This structure, however, would produce a 1:1 binding stoichiometry which is not consisted with experiment. In the case of the docked model proposed in this work, only one CQ molecule can be accommodated between the porphyrin rings. Thus even in an excess of CQ, further binding of CQ to the complex is not possible. Furthermore, the presence of ferrihaem in the form of the μ-oxo dimer is consistent with previously reported Mössbauer spectroscopic measurements of the CQ-ferrihaem complex precipitated at pH 5.[180] Induction of the μ-oxo dimer by CQ also supports the observed entropy increase upon complex formation reported by Egan *et al.*[124] Lastly, the 2:4 CQ-ferrihaem complex is supportive of the reported increased solubility of the complex in aqueous solution[155] since the highly aggregated states previously suggested would tend to cause precipitation at higher concentrations. It should be noted, however, that while the interactions identified in the docked CQ-(μ-oxo ferrihaem) complex provide some understanding of previously reported experimental

observations and SARs, they are qualitative in nature and thus further investigation in this regard is strongly warranted (see Section 7.2 for suggested future studies).

The work presented in this study highlights the advantages of employing multiple experimental techniques to investigate and address the interaction of antimalarial drugs with ferrihaem. It also emphasises the utility and advantage in combining these methods with computational simulation. It is important to note, however, that while the findings produced as a result of this combined experimental and computational approach have provided greater insight regarding the CQ-ferrihaem complex in aqueous solution, it is not proposed that complex formation between CQ and ferrihaem is necessarily the basis of haemozoin inhibition by this antimalarial. In fact, recent evidence tends to support the hypothesis that CQ exerts its antimalarial activity by docking into the fastest growing crystal face of haemozoin.[154] Nevertheless, understanding the properties of the CQ-ferrihaem complex in solution is still of outstanding interest. This is because the complex is likely present in the parasite owing to the putative high concentrations of CQ in the DV and the increased ferrihaem concentrations that occur as a result of haemozoin inhibition. Thus the presence of the CQ-ferrihaem complex in the DV probably plays a key role in determining the toxic effects of haem which lead to the death of the parasite. Indeed, recent *in vivo* evidence has shown that the concentrations of free ferrihaem required to halt the growth of malaria parasites differ between haemozoin-inhibiting antimalarial drugs (Egan and co-workers, unpublished data), an observation which may very well relate to properties of the complex.

7.2 Future Work

A number of findings from this study warrant further investigation. These have been divided into two separate categories relating to experimental and computational investigations.

7.2.1 Experimental investigations.

As mentioned in Section 7.1, the formation of the μ-oxo dimer may well play an important role with regard to toxicity effects causing death of the malaria parasite. This could be investigated by determining the ability of various antimalarial drugs to induce μ-oxo dimer formation using magnetic susceptibility measurements and relating them to the free ferrihaem concentrations required to kill parasites employing the method described by Combrinck *et al*.[128] The study could be further expanded to identify whether different ferrihaem species produce more reactive oxygen species or partition to a greater extent into lipid membranes. These investigations would be of considerable importance for the design of new potent antimalarial drugs.

Three further investigations could be conducted to confirm some of the observations proposed in this study. The first would be to identify the Fe-O-Fe stretching frequency that was obscured in the IR spectrum of the solid precipitated at low pH. Conducting nuclear resonance vibrational spectroscopy on the precipitated complex would be ideal for this purpose as only ^{57}Fe-X vibrations would be observed. RR spectroscopy could also be applied, however, it should be noted that Dziedzic-Kocurek *et al*. have been unable to detect this peak even in a well-characterised μ-oxo ferrihaem sample.[376] The second study of interest relates to the entropy associated with π-π dimer formation in aqueous solution. This could be investigated using ITC or in a manner analogous to that reported by Egan *et al*.[124] using UV-visible spectroscopy. Finally, the dimerisation of 1:2 CQ-ferrihaem complexes to give 2:4 complexes could be probed in a similar manner to that used to determine dimerisation constants for π-π dimeric ferrihaem by de Villiers *et al*.[75] This could further confirm or refute the existence of the 2:4 form in aqueous solution.

7.2.2 Computational investigations.

Given the newly-developed force field parameters for free five-coordinate ferrihaem species, a range of computational investigations are possible. These include simulating the solution-phase interactions of ferrihaem and a variety of antimalarial drugs such as quinine, quinidine, amodiaquine as well as their analogues. In addition, probing the solvation of ferrihaem species in aqueous mixtures of water-miscible solvents such as in 40% (v/v) aq. DMSO may provide further information on the mechanism of μ-oxo dimer formation. A more immediately relevant study, however, would be to probe the solvation and structure of ferrihaem species at a lipid-water interface in order to uncover more information about the early stages of haemozoin nucleation. This is particularly important since little is known regarding this initial process at the atomic level. MD simulations of ferrihaem species and the CQ-ferrihaem complex at a lipid-water interface could also be conducted in a similar approach to that used in the present study to better understand partitioning, orientation and interactions between ferrihaem and neutral lipids, such as monopalmitolglycerol, implicated in haemozoin formation.

To provide further support for the docked CQ-ferrihaem complex and the qualitative rationalisation of SARs, free energy perturbation calculations could be performed on a range of 4-aminoquinoline molecules with varying 7-position substituents. The differences in association free energy between different compounds could be computed and related to previously-determined experimental association constants. Furthermore, the interaction between μ-oxo ferrihaem and 4-aminoquinoline analogues with bulky substituents in the 8-position could be simulated in order to see if any disfavour the docked conformation. Synthesis of these compounds and subsequent association constant measurement could then be conducted to ascertain whether binding has substantially decreased.

The proposed experimental and computational investigations would provide valuable further insight into the interaction between ferrihaem and CQ as well as their relation to haemozoin formation and inhibition. This knowledge might assist in understanding the role of ferrihaem under biological conditions and aid the design of sorely-needed new antimalarial drugs.

REFERENCES

1. J.-J. Chen, R. N. V. S. Suragani, H. Ding, B. D. Sumer, J. Gao, F. Gao, L. Zhang, Y. Li, R. Sessoms, D. P. Huynh, A. Arvey, C. Leslie, T. Chernova and A. G. Smith, in *Heme Biology. The secret life of heme in regulating diverse biological processes*, World Scientific, Singapore, 2011.
2. F. C. Bernstein, T. F. Koetzle, G. J. Williams, E. E. Meyer Jr, M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi and M. Tasumi, *J. Mol. Biol.*, 1977, **112**, 535-542.
3. *http://www.rcsb.org/pdb/*, Accessed 6 January 2014.
4. I. G. Denisov, T. M. Makris, S. G. Sligar and I. Schlichting, *Chem. Rev.*, 2005, **105**, 2253-2277.
5. C. Garrido, L. Galluzzi, M. Brunet, P. E. Puig, C. Didelot and G. Kroemer, *Cell Death Differ.*, 2006, **13**, 1423-1433.
6. M. D. Maines, *Annu. Rev. Pharmacol. Toxicol.*, 1997, **37**, 517-554.
7. S. M. Mense and L. Zhang, *Cell Res.*, 2006, **16**, 681-692.
8. P. R. Ortiz de Mondellano, Hemes in Biology, in *Wilely Encyclopedia of Chemical Biology*, T. Begley, John Wiley & Sons Ltd, West Sussex, 2009, vol 2, pp. 240-249.
9. C. Savino, A. E. Miele, F. Draghi, K. A. Johnson, G. Sciara, M. Brunori and B. Vallone, *Biopolymers*, 2009, **91**, 1097-1107.
10. N. Mast, M. A. White, I. Bjorkhem, E. F. Johnson, C. D. Stout and I. A. Pikuleva, *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 9546-9551.
11. A. Messerschmidt, R. Huber, T. Poulos and K. Wieghardt, in *Handbook of metalloproteins*, John Wiley & Sons Ltd, West Sussex, 2001.
12. C. J. Reedy and B. R. Gibney, *Chem. Rev.*, 2004, **104**, 617-649.
13. A. S. Tsiftsoglou, A. I. Tsamadou and L. C. Papadopoulou, *Pharmacol. Therapeut.*, 2006, **111**, 327-345.
14. J.-H. Fuhrhop, *Langmuir*, 2013, **30**, 1-12.
15. N. D. Ishii and G. M. Maniatis, *Nature*, 1978, **274**, 372-374.
16. J.-J. Chen and I. M. London, *Cell*, 1981, **26**, 117-122.
17. A. S. Tsiftsoglou and S. H. Robinson, *Int. J. Cell Cloning*, 1985, **3**, 349-366.
18. F. C. Monette, S. A. Holden, M. J. Sheehy and E. A. Matzinger, *Exp. Hematol.*, 1984, **12**, 782-787.
19. A. S. Tsiftsoglou, I. S. Pappas and I. S. Vizirianakis, *Pharmacol. Therapeut.*, 2003, **100**, 257-290.
20. V. Jeney, J. Balla, A. Yachie, Z. Varga, G. M. Vercellotti, J. W. Eaton and G. Balla, *Blood*, 2002, **100**, 879-887.
21. S. C. Larsson, H.-O. Adami, E. Giovannucci and A. Wolk, *J. Natl. Cancer Inst.*, 2005, **97**, 232-233.
22. F. Pierre, S. Tache, C. R. Petit, R. van der Meer and D. E. Corpet, *Carcinogenesis*, 2003, **24**, 1683-1690.
23. H. Atamna and W. H. Frey II, *Proc. Natl. Acad. Sci. USA*, 2004, **30**, 11153-11158.
24. D. Pramanik and S. G. Dey, *J. Am. Chem. Soc.*, 2011, **133**, 81-87.
25. R. L. Aft and G. C. Mueller, *J. Biol. Chem.*, 1983, **258**, 12069-12072.
26. H. B. Suliman, M. S. Carraway, L. W. Velsor, B. J. Day, A. J. Ghio and C. A. Piantadosi, *Free Radic. Biol. Med.*, 2002, **32**, 246-256.
27. D. J. Schaer, P. W. Buehler, A. I. Alayash, J. D. Belcher and G. M. Vercellotti, *Blood*, 2013, **121**, 1276-1284.
28. M. Beppu, M. Nagoya and K. Kikugawa, *Chem. Pharm. Bull.*, 1986, **34**, 5063-5070.
29. S. Kumar and U. Bandyopadhyay, *Toxicol. Lett.*, 2005, **157**, 175-188.
30. T. H. Schmitt, W. A. Frezzatti and S. Schreier, *Arch. Biochem. Biophys.*, 1993, **307**, 96-103.
31. P. R. Montellano, *Curr. Opin. Chem. Biol.*, 2000, **4**, 221-227.
32. S. R. Robinson, T. N. Dang, R. Dringen and G. M. Bishop, *Redox Rep.*, 2009, **14**, 228-235.
33. A. A. Kassim and M. R. DeBaun, *Annu. Rev. Med.*, 2013, **64**, 456-466.
34. M. Krugliak, J. Zhang and H. Ginsburg, *Mol. Biochem. Parasitol.*, 2002, **119**, 249-256.

35. *World Malaria Report 2012*, World Health Organisation, Geneva, http://www.who.int/malaria/publications/world_malaria_report_2012/en/, 6 January 2014.
36. N. A. Campbell and J. B. Reece, in *Biology*, 6th edn., Benjamin Cummings, San Francisco, 2002, pp 557-558.
37. L. M. Prescott, J. P. Harley and D. A. Klein, in *Microbiology*, 6th edn., McGraw-Hill, New York, 2005, pp 929-931.
38. G. Jiang, M. Shi, S. Conteh, N. Richie, G. Banania, H. Geneshan, A. Valencia, P. Singh, J. Aguiar, K. Limbach, K. I. Kamrud, J. Rayner, J. Smith, J. T. Bruder, C. R. King, T. Tsuboi, S. Takeo, Y. Endo, D. L. Doolan, T. L. Richie and W. R. Weiss, *PLOS one*, 2009, **4**, e6559.
39. I. Petersen, R. Eastman and M. Lanzer, *FEBS Lett.*, 2011, **585**, 1551-1562.
40. J. Cox-Singh, *Curr. Opin. Infect. Dis.*, 2012, **25**, 530-536.
41. J. N. Hays, in *Epidemics and pandemics: their impacts on human history*, ABC-CLIO, California, 2005.
42. *National Institute of Allergy and Infectious Diseases*, http://www.niaid.nih.gov/topics/malaria/pages/lifecycle.aspx, Accessed 27 January, 2014.
43. P. Loria, S. Miller, M. Foley and L. Tilley, *Biochem. J.*, 1999, **339**, 363-370.
44. T. J. Egan, J. M. Combrinck, J. Egan, G. R. Hearn, H. M. Marques, S. Ntenteni, B. T. Sewell, P. J. Smith, D. Taylor, D. A. van Schalkwyk and J. C. Walden, *Biochem. J.*, 2002, **365**, 343-347.
45. D. J. Krogstad, P. H. Schlesinger and I. Y. Gluzman, *J. Cell. Biol.*, 1985, **101**, 2302-2309.
46. R. Hayward, K. J. Saliba and K. Kirk, *J. Cell. Sci.*, 2006, **119**, 1016-1025.
47. R. Banerjee, J. Liu, W. Beatty, L. Pelosof, M. Klemba and D. E. Goldberg, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 990-995.
48. P. J. Rosenthal, P. S. Sijwali, A. Singh and B. R. Shenai, *Curr. Pharm. Des.*, 2002, **8**, 1659-1672.
49. K. K. Eggleson, K. L. Duffin and D. E. Goldberg, *J. Biol. Chem.*, 1999, **274**, 32411-32417.
50. T. J. Egan, *Targets*, 2003, **2**, 115-124.
51. T. J. Egan, *Mol. Biochem. Parasitol.*, 2008, **157**, 127-136.
52. T. J. Egan, *J. Inorg. Biochem.*, 2008, **102**, 1288-1299.
53. S. Pagola, P. W. Stephens, D. S. Bohle, A. D. Kosar and S. K. Madsen, *Nature*, 2000, **404**, 307-310.
54. M. F. Oliveira, J. R. Silva, M. Dansa-Petretski, W. de Souza, U. Lins, C. M. S. Braga, H. Masuda and P. L. Oliveira, *Nature*, 1999, **400**, 517-518.
55. M. F. Oliveira, J. C. P. d'Avila, C. R. Torres, P. L. Oliveira, A. J. Tempone, F. D. Rumjanek, C. M. S. Braga, J. R. Silva, M. Dansa-Petretski, M. A. Oliveira, W. de Souza and S. T. Ferreira, *Mol. Biochem. Parasitol.*, 2000, **111**, 217-221.
56. K. Ziele and F. Reuter, *Z. Physiol. Chem.*, 1933, **221**, 101-116.
57. F. Haurowitz, *Z. Physiol. Chem.*, 1938, **254**, 266-274.
58. J. Shack and W. M. Clark, *J. Biol. Chem.*, 1947, **171**, 143-187.
59. G. Blauer and B. Zvilichovsky, *Arch. Biochem. Biophys.*, 1968, **127**, 749-755.
60. S. Moreau, B. Perly and J. Biguet, *Biochimie*, 1982, **64**, 1015-1025.
61. V. Srinivas and C. M. Rao, *Biochem. Int.*, 1990, **21**, 849-855.
62. K. Kuzelová, M. Mrhalová and Z. Hrkal, *Biochim. Biophys. Acta*, 1997, **1336**, 497-501.
63. S. B. Brown, T. C. Dean and P. Jones, *Biochem. J.*, 1970, **117**, 741-744.
64. S. B. Brown, P. Jones and I. R. Lantzke, *Nature*, 1969, **223**, 960-961.
65. L. Kiong-Lam, J. W. Buchler, J. E. Kenny and W. R. Scheidt, *Inorg. Chim. Acta*, 1986, **123**, 91-97.
66. S. H. Strauss, M. J. Pawlik, J. Skowyra, J. R. Kennedy, O. P. Anderson, K. Spartalian and J. L. Dye, *Inorg. Chem.*, 1987, **26**, 724-730.
67. M. Li, M. Shang, H. F. Duval and W. R. Scheidt, *Acta Crystallogr. Sect. C-Cryst. Struct. Commun.*, 2000, **56**, 1206-1207.
68. A.-R. Li, H.-H. Wei and L.-L. Gang, *Inorg. Chim. Acta*, 1999, **290**, 51-56.

69. A. Gold, K. Jayaraj, P. Doppelt, J. Fischer and R. Weiss, *Inorg. Chim. Acta*, 1988, **150**, 177-181.
70. K. D. Karlin, A. Nanthakumar, S. Fox, N. N. Murthy, N. Ravi, B. H. Huynh, R. D. Orosz and E. P. Day, *J. Am. Chem. Soc.*, 1994, **116**, 4753-4763.
71. H. M. Lee, M. M. Olmstead, G. G. Gross and A. L. Balch, *Cryst. Growth Des.*, 2003, **3**, 691-697.
72. L. Cheng, J. Lee, D. R. Powell and G. B. Richter-Addo, *Acta Crystallogr.*, 2004, **E60**, m1340.
73. *Cambridge Structural Database and Cambridge Structrural Database System V 1.13; Cambridge Crystallographic Data Centre, University Chemical Laboratory: Cambridge*, 2011.
74. D. H. O'Keeffe, C. H. Barlow, G. A. Smythe, W. H. Fuchsman, T. H. Moss, H. R. Lilienthal and W. S. Caughey, *Bioinorg. Chem.*, 1975, **5**, 125-147.
75. K. A. de Villiers, C. H. Kaschula, T. J. Egan and H. M. Marques, *J. Biol. Inorg. Chem.*, 2007, **12**, 101-117.
76. C. Asher, K. A. de Villiers and T. J. Egan, *Inorg. Chem.*, 2009, **48**, 7994-8003.
77. A. P. Gorka, A. C. de Dios and P. D. Roepe, *J. Med. Chem.*, 2013, **56**, 5231-5246.
78. K. A. de Villiers, H. M. Marques and T. J. Egan, *J. Inorg. Biochem.*, 2008, **102**, 1660-1667.
79. O. Q. Munro, J. C. Bradley, R. D. Hancock, H. M. Marques, F. Marsicano and P. W. Wade, *J. Am. Chem. Soc.*, 1992, **114**, 7218-7230.
80. H. M. Marques, O. Q. Munro, N. E. Grimmer, D. C. Levendis, F. Marsicano, G. Pattrick and T. Markoulides, *J. Chem. Soc. Faraday Trans.*, 1995, **91**, 1741-1749.
81. A. F. G. Slater and A. Cerami, *Nature*, 1992, **355**, 167-169.
82. M. Chugh, V. Sundararaman, S. Kumar, V. S. Reddy, W. A. Siddiqui, K. D. Stuart and P. Malhotra, *Proc. Natl. Acad. Sci. USA*, 2013, **110**, 5392-5397.
83. D. J. Sullivan, I. Y. Gluzman and D. E. Goldberg, *Science*, 1996, **271**, 219-222.
84. K. Bendrat, B. J. Berger and A. Cerami, *Nature*, 1995, **378**, 138-139.
85. N. T. Huy, Y. Shima, A. Maeda, T. T. Men, K. Hirayama, A. Hirase, A. Miyazawa and K. Kamei, *PLOS one*, 2013, **8**, e70025.
86. J. M. Pisciotta, I. Coppens, A. K. Tripathi, P. F. Scholl, J. Schuman, S. Bajad, V. Shulaev and D. J. Sullivan, *Biochem. J.*, 2007, **402**, 197-204.
87. M. A. Ambele and T. J. Egan, *Malaria J.*, 2012, **11**, 337.
88. A. N. Hoang, K. K. Ncokazi, K. A. de Villiers, D. W. Wright and T. J. Egan, *Dalton Trans.*, 2010, **39**, 1235-1344.
89. T. J. Egan, J. Y.-J. Chen, K. A. de Villiers, T. E. Mabotha, K. J. Naidoo, K. K. Ncokazi, S. J. Langford, D. McNaughton, S. Pandiancherri and B. R. Wood, *FEBS Lett.*, 2006, **580**, 5105-5110.
90. M. A. Ketchum, K. N. Olafson, E. V. Petrova, J. D. Rimer and P. G. Vekilov, *J. Chem. Phys.*, 2013, **139**, 121911.
91. K. A. de Villiers, M. Osipova, T. E. Mabotha, I. Solomonov, Y. Feldman, K. Kjaer, I. Weissbuch, T. J. Egan and L. Leiserowtiz, *Cryst. Growth Des.*, 2009, **9**, 626-632.
92. M. A. Ambele, B. T. Sewell, F. R. Cummings, P. J. Smith and T. J. Egan, *Cryst. Growth Des.*, 2013, **13**, 4442-4452.
93. S. Kapishnikov, A. Weiner, E. Shimoni, P. Guttman, G. Schneider, N. Dahan-Pasternak, R. Dzikowski, L. Leiserowtiz and M. Elbaum, *Proc. Natl. Acad. Sci. USA*, 2012, **109**, 11188-11193.
94. S. Kapishnikov, A. Weiner, E. Shimoni, G. Schneider, M. Elbaum and L. Leiserowtiz, *Langmuir*, 2013, **29**, 14595-14602.
95. P. Olliaro, *Pharmacol. Therapeut.*, 2001, **89**, 207-219.
96. G. Padmanaban, V. A. Nagaraj and P. N. Rangarajan, *Curr. Sci.*, 2007, **92**, 1545-1555.
97. H. J. Painter, J. M. Morrisey, M. W. Mather and A. B. Vaidya, *Nature*, 2007, **446**, 88-91.
98. U. Eckstein-Ludwig, R. J. Webb, I. D. A. van Goethem, J. M. East, A. G. Lee, M. Kimura, P. M. O'Neill, P. G. Bray, S. A. Ward and S. Krishna, *Nature*, 2003, **424**, 957-961.
99. A. C. Chou, R. Chevli and C. Fitch, *Biochemistry*, 1980, **19**, 1543-1549.

100. Y. Sugioka, M. Suzuki, K. Sugioka and M. Nakano, *FEBS Lett.*, 1987, **223**, 251-254.
101. J. Ziegler, R. Linck and D. W. Wright, *Curr. Med. Chem.*, 2001, **8**, 171-189.
102. J. Achan, A. O. Talisuna, A. Erhart, A. Yeka, J. K. Tibenderana, F. N. Baliraine, P. J. Rosenthal and U. D'Alessandro, *Malaria J.*, 2011, **10**, 144.
103. S. R. Meshnick and M. J. Dobson, The history of antimalarial drugs, in *Antimalarial Chemotherapy: Mechanisms of Action, Resistance, and New Directions in Drug Discovery*, ed. P. J. Rosenthal, Humana Press Inc, Totowa, 2001.
104. K. Krafts, E. Hempelmann and A. Skórska-Stania, *Parasitol. Res.*, 2012, **111**, 1-6.
105. J. E. Hyde, *FEBS J.*, 2007, **274**, 4688-4698.
106. F. Kwakye-Berko and S. R. Meshnick, *Mol. Biochem. Parasitol.*, 1989, **35**, 51-55.
107. F. Kwakye-Berko and S. R. Meshnick, *Mol. Biochem. Parasitol.*, 1990, **39**, 275-278.
108. S. R. Meshnick, *Ann. Trop. Med. Parasitol.*, 1996, **90**, 367-372.
109. E. Konigk and B. Putfarken, *Trop. Med. Parasitol.*, 1983, **34**, 1-3.
110. E. Konigk and B. Putfarken, *Trop. Med. Parasitol.*, 1985, **36**, 81-84.
111. N. Surolia and G. Padmanaban, *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 4786-4790.
112. R. Zidovetzki, I. W. Sherman and L. O'Brien, *J. Parasitol.*, 1993, **79**, 565-570.
113. R. Zidovetzki, I. W. Sherman, J. Prudhomme and J. Crawford, *Parasitology*, 1994, **108**, 249-255.
114. I. Y. Gluzman, S. E. Francis, A. Oksman, C. E. Smith, K. L. Duffin and D. E. Goldberg, *J. Clin. Invest.*, 1994, **93**, 1602-1608.
115. H. Ginsburg, O. Famin, J. Zhang and M. Krugliak, *Biochem. Pharmacol.*, 1998, **56**, 1305-1313.
116. D. F. Platel, F. Mangou and J. Tribouley-Duret, *Mol. Biochem. Parasitol.*, 1999, **98**, 215-223.
117. J. Zhang, M. Krugliak and H. Ginsburg, *Mol. Biochem. Parasitol.*, 1999, **99**, 129-141.
118. D. J. Sullivan, *Int. J. Parasitol.*, 2002, **32**, 1645-1653.
119. C. A. Homewood, D. C. Warhurst, W. Peters and V. C. Baggaley, *Nature*, 1972, **235**, 50-52.
120. C. H. Kaschula, T. J. Egan, R. Hunter, N. Basilico, S. Parapini, D. Taramelli, E. Pasini and D. Monti, *J. Med. Chem.*, 2002, **45**, 3531-3539.
121. D. J. Sullivan, I. Y. Gluzman, D. G. Russell and D. E. Goldberg, *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 11865-11870.
122. T. J. Egan, D. C. Ross and P. A. Adams, *FEBS Lett.*, 1994, **352**, 54-57.
123. J. M. Karle, I. L. Karle, L. Gerena and W. K. Milhous, *Antimicrob. Agents Chemother.*, 1992, **36**, 1538-1544.
124. T. J. Egan, W. W. Mavuso, D. C. Ross and H. M. Marques, *J. Inorg. Biochem.*, 1997, **68**, 137-145.
125. A. Dorn, S. R. Vippagunta, H. Matile, C. Jaquet and J. L. Vennerstrom, *Biochem. Pharmacol.*, 1998, **55**, 727-736.
126. T. J. Egan, E. Hempelmann and W. W. Mavuso, *J. Inorg. Biochem.*, 1999, **73**, 101-107.
127. T. J. Egan, R. Hunter, C. H. Kaschula, H. M. Marques, A. Misplon and J. C. Walden, *J. Med. Chem.*, 2000, **43**, 283-291.
128. J. M. Combrinck, T. E. Mabotha, K. K. Ncokazi, M. A. Ambele, D. Taylor, P. J. Smith, H. C. Hoppe and T. J. Egan, *ACS Chem. Biol.*, 2012, **8**, 133-137.
129. T. J. Egan and D. Kuter, *Future Microbiol.*, 2013, **8**, 475-489.
130. H. Noedl, Y. Se, K. Schaecher, B. L. Smith, D. Socheat and M. M. Fukuda, *N. Engl. J. Med.*, 2008, **359**, 2619-2620.
131. A. M. Dondorp, F. Nosten, P. Yi, D. Das, A. Phyo, J. Tarning, K. M. Lwin, F. Ariey, W. Hanpithakpong, S. J. Lee, P. Ringwald, K. Silamut, M. Imwong, K. Chotivanich, P. Lim, T. Herdman, S. S. An, S. Yeung, P. Singhasivanon, N. P. J. Day, N. Lindegardh, D. Socheat and N. J. White, *N. Engl. J. Med.*, 2009, **361**, 455-467.

132. A. P. Phyo, S. Nkhoma, K. Stepniewska, E. A. Ashley, S. Nair, R. McGready, C. ler Moo, S. Al-Saai, A. M. Dondorp, K. M. Lwin, P. Singhasivanon, N. P. J. Day, N. J. White, T. J. C. Anderson and F. Nosten, *Lancet*, 2012, **379**, 1960-1966.
133. C. Sisowath, J. Strömberg, A. Mårtensson, M. Msellem, C. Obondo, A. Björkman and J. P. Gil, *J. Infect. Dis.*, 2005, **191**, 1014-1017.
134. T. N. C. Wells, New medicines to combat malaria: an overview of the global pipeline of therapeutics, in *Milestones in Drug Therapy. Treatment and Prevention of Malaria*, eds. H. M. Staines and S. Krishna, Spinger Basel AG, Switzerland, 2012, pp. 227-247.
135. C. Biot and D. Dive, *Top. Organomet. Chem.*, 2010, **2010**, 155-193.
136. M. Foley and L. Tilley, *Pharmacol. Therapeut.*, 1998, **79**, 55-87.
137. K. J. Saliba, P. I. Folb and P. J. Smith, *Biochem. Pharmacol.*, 1998, **56**, 313-320.
138. A. Yayon, Z. I. Cabantchik and H. Ginsburg, *Proc. Natl. Acad. Sci. USA*, 1985, **82**, 2784-2788.
139. R. L. Summers, M. N. Nash and R. E. Martin, *Cell Mol. Life Sci.*, 2012, **69**, 1967-1995.
140. R. E. Martin and K. Kirk, *Mol. Biol. Evol.*, 2004, **21**, 1938-1949.
141. M. B. Reed, K. J. Saliba, S. R. Caruana, K. Kirk and A. F. Cowman, *Nature*, 2000, **403**, 906-909.
142. J. Mu, M. T. Ferdig, X. Feng, D. A. Joy, J. Duan, T. Furuya, G. Subramanian, L. Aravind, R. A. Cooper, J. C. Wootton, M. Xiong and X. Z. Su, *Mol. Microbiol.*, 2003, **49**, 977-989.
143. J. K. Lekostaj, J. K. Natarajan, M. F. Paguio, C. Wolf and P. D. Roepe, *Biochemistry*, 2008, **47**, 10394-10406.
144. R. E. Martin, R. V. Marchetti, A. I. Cowan, S. M. Howitt, S. Bröer and K. Kirk, *Science*, 2009, **325**, 1680-1682.
145. S. K. Martin, A. M. J. Oduola and W. K. Milhous, *Science*, 1987, **235**, 899-901.
146. A. Ecker, A. M. Lehane, J. Clain and D. A. Fidock, *Trends Parasitol.*, 2012, **28**, 504-514.
147. D. C. Warhurst, J. C. Craig and I. S. Adagu, *Lancet*, 2002, **360**, 1527-1529.
148. R. Buller, M. L. Petersen, Ö. Almarsson and L. Leiserowtiz, *Cryst. Growth Des.*, 2002, **2**, 553-562.
149. J. M. Karle, *Acta Cryst. C*, 1988, **44**, 1605-1608.
150. R. M. Martin, in *Electron structure: basic theory and practical methods*, Cambridge University Press, Cambridge, 2008, pp 152-171.
151. S. Nsumiwa, D. Kuter, S. Wittlin, K. Chibale and T. J. Egan, *Bioorg. Med. Chem.*, 2013, **21**, 3738-3748.
152. M. S. Walczak, K. Lawniczak-Jablonska, A. Wolska, M. Sikora, A. Sienkiewicz, L. Suárez, A. J. Kosar, M. J. Bellemare and D. S. Bohle, *J. Phys. Chem. B*, 2011, **115**, 4419-4426.
153. D. S. Bohle, E. L. Dodd, A. J. Kosar, L. Sharma, P. J. Stephens, L. Suárez and D. Tazoo, *Angew. Chem., Int. Ed.*, 2011, **50**, 6151-6154.
154. J. Gildenhuys, T. le Roux, T. J. Egan and K. A. de Villiers, *J. Am. Chem. Soc.*, 2013, **135**, 1037-1047.
155. L. M. B. Ursos, K. DuBay and P. D. Roepe, *Mol. Biochem. Parasitol.*, 2001, **112**, 11-17.
156. K. A. de Villiers, J. Gildenhuys and T. le Roux, *ACS Chem. Biol.*, 2012, **7**, 666-671.
157. J. N. Alumasa, A. P. Gorka, L. B. Casabianca, E. Comstock, A. C. de Dios and P. D. Roepe, *J. Inorg. Biochem.*, 2011, **105**, 467-475.
158. S. N. Cohen, K. O. Phifer and K. L. Yielding, *Nature*, 1964, **202**, 805-806.
159. S. R. Cheruku, S. Maiti, A. Dorn, B. Scorneaux, A. K. Bhattacharjee, W. Y. Ellis and J. L. Vennerstrom, *J. Med. Chem.*, 2003, **2003**, 3166-3169.
160. K. Bachhawat, C. J. Thomas, N. Surolia and A. Surolia, *Biochim. Biophys. Acta*, 2000, **276**, 1075-1079.
161. S. R. Vippagunta, A. Dorn, R. G. Ridley and J. L. Vennerstrom, *Biochim. Biophys. Acta*, 2000, **1475**, 133-140.
162. T. J. Egan and K. K. Ncokazi, *J. Inorg. Biochem.*, 2004, **98**, 144-152.
163. M. P. Crespo, L. Tilley and N. Klonis, *J. Biol. Inorg. Chem.*, 2010, **15**, 1009-1022.

164. I. Constantinidis and J. D. Satterlee, *J. Am. Chem. Soc.*, 1988, **110**, 4391-4395.
165. H. M. Marques, K. Voster and T. J. Egan, *J. Inorg. Biochem.*, 1996, **64**, 7-23.
166. O. Q. Munro and H. M. Marques, *Inorg. Chem.*, 1996, **35**, 3752-3767.
167. S. Moreau, B. Perly, C. Chachaty and C. Deleuze, *Biochim. Biophys. Acta*, 1985, **840**, 107-116.
168. L. B. Casabianca, D. An, J. K. Natarajan, J. N. Alumasa, P. D. Roepe, C. Wolf and A. C. de Dios, *Inorg. Chem.*, 2008, **47**, 6077-6081.
169. L. B. Casabianca, J. B. Kallgren, J. K. Natarajan, J. N. Alumasa, P. D. Roepe, C. Wolf and A. C. De Dios, *J. Inorg. Biochem.*, 2009, **103**, 745-748.
170. K. F. Schwedhelm, M. Horstmann, J. H. Faber, Y. Reichert, M. Büchner, G. Bringmann and C. Faber, *Open Spec. J.*, 2008, **2**, 10-18.
171. G. T. Webster, D. McNaughton and B. R. Wood, *J. Phys. Chem. B*, 2009, **113**, 6910-6916.
172. A. Leed, K. DuBay, L. M. B. Ursos, D. Sears, A. C. de Dios and P. D. Roepe, *Biochemistry*, 2002, **41**, 10245-10255.
173. J. K. Natarajan, J. N. Alumasa, K. Yearick, K. Ekoue-Kovi, L. B. Casabianca, A. C. de Dios, C. Wolf and P. D. Roepe, *J. Med. Chem.*, 2008, **51**, 3466-3479.
174. A. C. de Dios, R. Tycko, L. M. B. Ursos and P. D. Roepe, *J. Phys. Chem. A*, 2003, **107**, 5821-5825.
175. C. Portela, C. M. M. Afonso, M. M. M. Pinto and M. J. Ramos, *Bioorg. Med. Chem.*, 2004, **12**, 3313-3321.
176. C. Portela, C. M. M. Afonso, M. M. M. Pinto and M. J. Ramos, *FEBS Lett.*, 2003, **547**, 217-222.
177. T. J. Egan, *J. Inorg. Biochem.*, 2006, **100**, 916-926.
178. D. F. Koenig, *Acta Cryst.*, 1965, **18**, 663-673.
179. M. J. Dascombe, M. G. B. Drew, H. Morris, P. Wilairat, S. Auparakkitanon, W. A. Moule, S. Alizadeh-Shekalgourabi, P. G. Evans, M. Lloyd, A. M. Dyas, P. Carr and F. M. D. Ismail, *J. Med. Chem.*, 2005, **48**, 5423-5436.
180. P. A. Adams, P. A. M. Berman, T. J. Egan, P. J. Marsh and J. Silver, *J. Inorg. Biochem.*, 1996, **63**, 69-77.
181. M. Asghari-Khiavi, J. Vongsvivut, I. Perepichka, A. Mechler, B. R. Wood, D. McNaughton and D. S. Bohle, *J. Inorg. Biochem.*, 2011, **105**, 1662-1669.
182. G. T. Webster, L. Tilley, S. Deed, D. McNaughton and B. R. Wood, *FEBS Lett.*, 2008, **582**, 1087-1092.
183. D. Kuter, K. Chibale and T. J. Egan, *J. Inorg. Biochem.*, 2011, **105**, 684-692.
184. H. Behzadi, M. D. Esrafili, J. Beheshtian, N. L. Hadipour and D. van der Spoel, *Chem. Phys. Lett.*, 2009, **476**, 196-200.
185. A. K. Bhattacharjee, *J. Mol. Struct.-Theochem*, 2001, **549**, 27-37.
186. L. B. Casabianca, C. M. Faller and A. C. de Dios, *J. Phys. Chem. A*, 2006, **110**, 234-240.
187. L. B. Casabianca and A. C. de Dios, *J. Phys. Chem. A*, 2006, **110**, 7787-7792.
188. T. Frosch, M. Schmitt, G. Bringmann, W. Kiefer and J. Popp, *J. Phys. Chem. B*, 2007, **111**, 1815-1822.
189. M. Gouterman, *J. Chem. Phys.*, 1959, **30**, 1139-1161.
190. M. Gouterman, *J. Mol. Spectrosc.*, 1961, **6**, 138-163.
191. A. J. McHugh, M. Gouterman and C. Weiss, *Theor. Chim. Acta*, 1972, **24**, 346-370.
192. A. M. Schaffer, M. Gouterman and E. R. Davidson, *Theor. Chim. Acta*, 1973, **30**, 9-30.
193. A. Ceulemans, W. Oldenhof, C. Görller-Walrand and L. G. Vanquickenborne, *J. Am. Chem. Soc.*, 1986, **108**, 1155-1163.
194. J. R. Platt, *J. Chem. Phys.*, 1949, **17**, 484-495.
195. W. T. Simpson, *J. Chem. Phys.*, 1949, **17**, 1218-1221.
196. M. W. Makinen and A. K. Churg, in *Iron Porphyrins. Part I. Physical Bioinorganic Chemistry Series*, Addison-Wesley, Reading, MA, 1983, vol 1, pp 141.

197. D. R. McMillin, in *Physical methods in bioinorganic chemistry: Spectroscopy and magnetism*, University Science Books, Sausalito, CA, 2000, pp 1.
198. R. W. Mason, in *A practical guide to magnetic circular dichroism spectroscopy*, Wiley, Hoboken, 2007.
199. J. Mack, M. J. Stillman and N. Kobayashi, *Coord. Chem. Rev.*, 2007, **251**, 429-453.
200. E. I. Solomon, E. G. Pavel, K. E. Loeb and C. Campochiaro, *Coord. Chem. Rev.*, 1995, **144**, 369-460.
201. P. J. Stephens, *Annu. Rev. Phys. Chem.*, 1974, **25**, 201-232.
202. S. B. Piepho and P. N. Schatz, in *Group Theory in Spectroscopy: With Applications to Magnetic Circular Dichroism*, John Wiley, New York, 1983.
203. P. J. Stephens, *J. Chem. Phys.*, 1970, **52**, 3489-3516.
204. P. J. Stephens, *Adv. Chem. Phys.*, 1976, **35**, 197-264.
205. J. Mack, Y. Asano, N. Kobayashi and M. J. Stillman, *J. Am. Chem. Soc.*, 2005, **127**, 17697-17711.
206. C. Hu, C. D. Sulok, F. Paulat, N. Lehnert, A. I. Twigg, M. P. Hendrich, C. E. Schulz and W. R. Scheidt, *J. Am. Chem. Soc.*, 2010, **132**, 3737-3750.
207. F. Paulat and N. Lehnert, *Inorg. Chem.*, 2008, **47**, 4963-4976.
208. N. Kobayashi and K. Nakai, *Chem. Commun.*, 2007, 4077-4092.
209. M. G. I. Galinato, T. Spolitak, D. P. Ballou and N. Lehnert, *Biochemistry*, 2011, **50**, 1053-1069.
210. R. Ortega, A. Carmona, I. Llorens and P. L. Solari, *J. Anal. At. Spectrom.*, 2012, **27**, 2054-2065.
211. A. Levina, R. S. Armstrong and P. A. Lay, *Coord. Chem. Rev.*, 2005, **249**, 141-160.
212. J. J. Rehr and R. C. Albers, *Rev. Mod. Phys.*, 2000, **72**, 621-654.
213. D. E. Sayers and E. A. Stern, *Phys. Rev. Lett.*, 1971, **27**, 1204-1207.
214. P. J. Merkling, R. Ayala, J. M. Martínez, R. R. Pappalardo and E. S. Marcos, *J. Chem. Phys.*, 2003, **119**, 6647-6654.
215. I. Ascone and R. Strange, *J. Synchrotron Rad.*, 2009, **16**, 413-421.
216. A. F. G. Slater, W. J. Swiggard, B. R. Orton, W. D. Flitter, D. E. Goldberg, A. Cerami and G. B. Henderson, *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 325-329.
217. M. S. Walczak, K. Lawniczak-Jablonska, A. Sienkiewicz, M. T. Klepka, L. Suárez, A. J. Kosar, M. J. Bellemare and D. S. Bohle, *J. Non-Cryst. Solids*, 2010, **356**, 1908-1913.
218. M. S. Walczak, K. Lawniczak-Jablonska, A. Wolska, A. Sienkiewicz, L. Suárez, A. Kosar and D. S. Bohle, *J. Phys. Chem. B*, 2011, **115**, 1145-1150.
219. A. Earnshaw, in *Introduction to magnetochemistry*, Academic Press, London, 1968.
220. J. Stanek and D. Dziedzic-Kocurek, *J. Mag. Mag. Mat.*, 2010, **322**, 999-1003.
221. S. S. Bhatnagar and K. N. Mathur, in *Physical principles and applications of magnetochemistry*, MacMillan and Co., London, 1935, pp 29-55.
222. D. R. Evans, *J. Chem. Soc.*, 1959, 2003-2005.
223. E. M. Schubert, *J. Chem. Educ.*, 1992, **69**, 62.
224. G. A. Bain and J. F. Berry, *J. Chem. Educ.*, 2008, **85**, 532-536.
225. J. Woolcock and A. Zafar, *J. Chem. Educ.*, 1992, **69**, A176-A179.
226. B. N. Figgis and R. S. Nyholm, *J. Chem. Soc.*, 1958, 4190-4191.
227. K. I. Ramachandran, G. Deepa and K. Namboori, in *Computational chemistry and molecular modeling*, Springer-Verlag, Heidelberg, 2008, pp 115-133.
228. E. G. Lewars, in *Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics*, Springer, New York, 2011, pp 445-449.
229. P. Hohenberg and W. Kohn, *Phys. Rev.* , 1964, **136**, B864-B871.
230. W. Kohn and L. J. Sham, *Phys. Rev. A*, 1980, **21**, 1561-1572.
231. M. E. Casida, *J. Mol. Struct.* , 2009, **914**, 3-18.
232. A. Dreuw and M. Head-Gordon, *Chem. Rev.*, 2005, **105**, 4009-4037.

233. M. A. L. Marques, C. A. Ullrich, F. Nogueira, A. Rubio, K. Burke and E. K. U. Gross, in *Time-dependent density functional theory*, Springer, Berlin Heidelberg, 2006.
234. E. Runge and E. K. U. Gross, *Phys. Rev. Lett.*, 1984, **52**, 997-100.
235. S. F. Sousa, P. A. Fernandes and M. J. Ramos, *J. Phys. Chem. A*, 2007, **111**, 10439-10452.
236. A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098-3100.
237. C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785-789.
238. J. P. Perdew, A. Ruzsinszky, J. Tao, V. N. Staroverov, G. E. Scuseria and G. Csonka, *J. Chem. Phys.*, 2005, **123**, 062201-062209.
239. J. C. Slater, *Phys. Rev.*, 1930, **36**, 57-64.
240. F. Jensen, in *Introduction to computational chemistry*, Wiley, Chichester, 1999, pp 171-172.
241. T. H. J. Dunning and P. J. Hay, in *Methods of Electronic Structure Theory*, Plenum Press, New York, 1977, vol 3, pp 1.
242. P. J. Hay and W. R. Wadt, *J. Chem. Phys.*, 1985, **82**, 270-283.
243. W. R. Wadt and P. J. Hay, *J. Chem. Phys.*, 1985, **82**, 284-298.
244. P. J. Hay and W. R. Wadt, *J. Chem. Phys.*, 1985, **82**, 299-310.
245. A. R. Leach, in *Molecular Modelling: principles and applications*, 2nd edn., Pearson Education Ltd, Harlow, 2001.
246. B. R. Brooks, C. L. Brooks III, A. D. MacKerell Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, H. F. I. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York and M. Karplus, *J. Comput. Chem.*, 2009, **30**, 1545-1614.
247. A. D. MacKerell Jr., D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Staub, M. Watanabe, J. Wiókiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586-3616.
248. K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. MacKerell Jr., *J. Comput. Chem.*, 2009, **31**, 671-690.
249. L. Verlet, *Phys. Rev.*, 1967, **159**, 98-103.
250. R. W. Hockney, *Methods Comput. Phys.*, 1970, **9**, 136-211.
251. W. C. Swope, H. C. Andersen, P. H. Berens and K. R. Wilson, *J. Chem. Phys.*, 1982, **76**, 637-649.
252. A. Hinchliffe, in *Molecular modelling for beginners*, John Wiley & Sons Ltd, West Sussex, 2005.
253. T. Frosch and J. Popp, *J. Mol. Struct.*, 2009, **924-926**, 301-308.
254. P. A. Adams, T. J. Egan, D. C. Ross, J. Silver and P. J. Marsh, *Biochem. J.*, 1996, **318**, 25-27.
255. P. Gans, A. Sabatini and A. Vacca, *Talanta*, 1996, **43**, 1739-1753.
256. P. W. Linder, L. R. Nassimbeni, A. Polson and A. L. Rodgers, *J. Chem. Educ.*, 1976, **53**, 330-332.
257. J. T. Edward, *J. Chem. Educ.*, 1970, **47**, 261-270.
258. H. M. Marques, O. Q. Munro and M. L. Crawcour, *Inorg. Chim. Acta*, 1992, **196**, 221-229.
259. J. B. Goodenough, in *Magnetism and the chemical bond*, Wiley, New York, 1963, pp 170-172.
260. W. R. Scheidt, B. Cheng, M. K. Safo, F. Cukiernik, J. C. Marchon and P. G. Debrunner, *J. Am. Chem. Soc.*, 1992, **114**, 4420-4421.
261. D. R. Evans, R. S. Mathur, K. Heerwegh, C. A. Reed and Z. Xie, *Angew. Chem., Int. Ed.*, 1997, **36**, 1335-1337.
262. S. K. Ghosh and S. P. Rath, *J. Am. Chem. Soc.*, 2010, **132**, 17983-17985.
263. Z. Chen, Z. Xu, L. Zhang, F. Yan and Z. Lin, *J. Phys. Chem. A*, 2001, **105**, 9710-9716.
264. D. F. Othmer and M. S. Thakar, *Ind. Eng. Chem.*, 1953, **45**, 589-593.
265. W. Hayduk and H. Laudie, *AICHe J.*, 1974, **20**, 611-615.

266. C. R. Wilke and P. Chang, *AICHe J.*, 1955, **1**, 264-270.
267. E. G. Scheibel, *Ind. Eng. Chem.*, 1954, **46**, 2007-2008.
268. K. E. Gustafson and R. M. Dickhut, *J. Chem. Eng. Data*, 1994, **39**, 281-285.
269. *Advanced Chemistry Development Inc., Toronto, ON, Canada, 2013*.
270. R. Niesner and A. Heintz, *J. Chem. Eng. Data*, 2000, **45**, 1121-1124.
271. C. Ercolani, M. Gardini, F. Monacelli, G. Pennesi and G. Rossi, *Inorg. Chem.*, 1983, **22**, 2584-2589.
272. G. W. Rayner Canham and A. B. P. Lever, *Inorg. Nucl. Chem. Lett.*, 1973, **9**, 513-517.
273. O. P. Charkin, N. M. Klimenko, P. T. Nguyen, D. O. Charkin, A. M. Mebel, S. H. Lin, Y.-S. Wang, S.-C. Wei and H.-C. Chang, *Chem. Phys. Lett.*, 2005, **415**, 362-369.
274. O. P. Charkin, N. M. Klimenko, D. O. Charkin, H.-C. Chang and S. H. Lin, *J. Phys. Chem. A*, 2007, **111**, 9207-9217.
275. T. Frosch, S. Koncarevic, L. Zedler, M. Schmitt, K. Schenzel, K. Becker and J. Popp, *J. Phys. Chem. B*, 2007, **111**, 11047-11056.
276. T. Frosch, S. Koncarevic, K. Becker and J. Popp, *Analyst*, 2009, **134**, 1126-1132.
277. N. Marom, A. Tkatchenko, L. Kronik, K. S. and L. Leiserowtiz, *Cryst. Growth Des.*, 2011, **11**, 3332-3341.
278. T. B. J. Pinter, E. L. Dodd, D. S. Bohle and M. J. Stillman, *Inorg. Chem.*, 2012, **51**, 3743-3753.
279. A. Paneque, J. Fernandez–Bertran, E. Reguera and H. Yee-Madeira, *Struct. Chem.*, 2003, **14**, 551-558.
280. C. Brénard, P. Kowalewski, J. C. Merlin and S. Moreau, *J. Raman Spectrosc.*, 1992, **23**, 325-333.
281. C. Brénard, J. J. Girerd, P. Kowalewski, J. C. Merlin and S. Moreau, *Appl. Spectrosc.*, 1993, **47**, 1837-1842.
282. B. R. Wood, S. J. Langford, B. M. Cooke, J. Lim, F. K. Glenister, M. Duriska, J. K. Unthank and D. McNaughton, *J. Am. Chem. Soc.*, 2004, **126**, 9233-9239.
283. T. Frosch, B. Küstner, S. Schlücker, A. Szeghalmi, M. Schmitt, W. Kiefer and J. Popp, *J. Raman Spectrosc.*, 2004, **35**, 819-821.
284. Gaussian 09, A. 01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.
285. G. Schaftenaar and J. H. Noordik, *J. Comput.-Aided Mol. Des.*, 2000, **14**, 123-134.
286. *Avogadro: An open-source molecular builder and visualization tool. Version 1.0.3; http://avogadro.openmolecules.net/*.
287. A.-R. Allouche, *J. Comput. Chem.*, 2011, **32**, 174-182.
288. S. I. Gorelsky, *SWizard program; University of Ottawa: Ottawa*, 2010, http://www.sg-chem.net/.
289. S. I. Gorelsky and A. B. P. Lever, *J. Organomet. Chem.*, 2001, **635**, 187-196.
290. E. A. Schmitt, L. Noodleman, E. J. Baerends and D. N. Hendrickson, *J. Am. Chem. Soc.*, 1992, **114**, 6109-6119.
291. G. B. Bacskay, *Chem. Phys.*, 1981, **61**, 385-404.
292. C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158-6169.

293. C. E. Check, T. O. Faust, J. M. Bailey, B. J. Wright, T. M. Gilbert and L. S. Sunderlin, *J. Phys. Chem. A*, 2001, **105**, 8111-8116.
294. A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B*, 2009, **113**, 6378-6396.
295. J. R. Platt, *J. Chem. Phys.*, 1949, **17**, 484.
296. W. Moffitt, *J. Chem. Phys.*, 1954, **22**, 320.
297. J. P. Perdew, *Phys. Rev. B*, 1986, **33**, 8822-8824.
298. Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215-241.
299. N. C. Handy and A. Cohen, *J. Mol. Phys.*, 2001, **99**, 403-412.
300. J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865-3868.
301. J. W. Pavlik, A. Barabanshikov, A. G. Oliver, E. E. Alp, S. W., J. Zhao, T. Sage and W. R. Scheidt, *Angew. Chem., Int. Ed.*, 2010, **49**, 4400-4404.
302. M. Swart, *J. Chem. Theory Comput.*, 2008, **4**, 2057-2066.
303. M. Swart, A. R. Groenhof, A. W. Ehlers and K. Lammertsma, *J. Phys. Chem. A*, 2004, **108**, 5479-5483.
304. D. S. Bohle, P. G. Debrunner, P. A. Jordan, S. K. Madsen and C. E. Schulz, *J. Am. Chem. Soc.*, 1998, **120**, 8255-8256.
305. C.-Y. Yeh, C. J. Chang and D. G. Nocera, *J. Am. Chem. Soc.*, 2001, **123**, 1513-1514.
306. K. K. Ncokazi and T. J. Egan, *Anal. Biochem.*, 2005, **338**, 306-319.
307. A. V. Pandey, S. K. Joshi, B. L. Tekwani and V. S. Chauhan, *Anal. Biochem.*, 1999, **268**, 159-161.
308. A. K. Tripathi, S. I. Khan, L. A. Walker and B. L. Tekwani, *Anal. Biochem.*, 2004, **325**, 85-91.
309. G. H. Loew, in *Iron Porphyrins. Part 1*, Addison-Wesley, New York, 1983, vol 1, pp 1.
310. G. A. Peralta, M. Seth and T. Ziegler, *Inorg. Chem.*, 2007, **46**, 9111-9125.
311. T. Yanai, D. Tew and N. C. Handy, *Chem. Phys. Lett.*, 2004, **393**, 51-57.
312. O. A. Vydrov and G. E. Scuseria, *J. Chem. Phys.*, 2006, **125**, 234109.
313. O. A. Vydrov, J. Heyd, A. V. Krukau and G. E. Scuseria, *J. Chem. Phys.*, 2006, **125**, 074106.
314. O. A. Vydrov, G. E. Scuseria and J. P. Perdew, *J. Chem. Phys.*, 2007, **126**, 154109.
315. J. Wan, Y. Ren, J. Wu and X. Xu, *J. Phys. Chem. A*, 2004, **108**, 9453-9460.
316. Y. Zhu, S. Zhou, Y. Kan and Z. Su, *Int. J. Quantum Chem.*, 2007, **107**, 1614-1623.
317. V. N. Nemykin, J. G. Olsen, E. Perera and P. Basu, *Inorg. Chem.*, 2006, **45**, 3557-3568.
318. C. Adamo and V. Barone, *Chem. Phys. Lett.*, 1999, **314**, 152-157.
319. A. Dreuw and M. Head-Gordon, *J. Am. Chem. Soc.*, 2004, **126**, 4007-4016.
320. M. J. G. Peach, P. Benfield, T. Helgaker and D. J. Tozer, *J. Chem. Phys.*, 2008, **128**, 044118.
321. J. B. Foresman, M. Head-Gordon, J. A. Pople and M. J. Frisch, *J. Phys. Chem.*, 1992, **96**, 135-149.
322. B. Cheng, M. K. Safo, R. D. Orosz, C. A. Reed, P. G. Debrunner and W. R. Scheidt, *Inorg. Chem.*, 1994, **33**, 1319-1324.
323. J. A. Sigman, B. C. Kwok and Y. Lu, *J. Am. Chem. Soc.*, 2000, **122**, 8192-8196.
324. C. Fufezan, J. Zhang and M. R. Gunner, *Proteins*, 2008, **78**, 690-704.
325. M. Devereux and M. Meuwly, *Biophys. J.*, 2009, **96**, 4363-4375.
326. U. F. Röhrig, A. Grosdidier, V. Zoete and O. Michielin, *J. Comput. Chem.*, 2009, **30**, 2305-2315.
327. F. Pietra, *Chem. Biodivers.*, 2012, **9**, 606-614.
328. F. Autenrieth, E. Tajkhorshid, J. Baudry and Z. Luthey-Schulten, *J. Comput. Chem.*, 2004, **25**, 1613-1622.
329. J. Henriques, P. J. Costa, M. J. Calhorda and M. Machuqueiro, *J. Phys. Chem. B*, 2013, **117**, 70-82.
330. X. de Hatten, Z. Cournia, I. Huc, J. C. Smith and N. Metzler-Nolte, *Chem. Eur. J.*, 2007, **13**, 8139-8152.
331. R. P. Matthews, G. A. Venter and K. J. Naidoo, *J. Phys. Chem. B*, 2011, **115**, 1045-1055.
332. K. J. Naidoo and M. Kuttel, *J. Comput. Chem.*, 2001, **22**, 445-456.
333. K. J. Naidoo, G. Klatt, K. R. Koch and D. J. Robinson, *Inorg. Chem.*, 2002, **41**, 1845-1849.

334. R. B. Best, G. E. Jackson and K. J. Naidoo, *J. Phys. Chem. B*, 2001, **105**, 4742-4751.
335. M. P. Allen and D. J. Tildesley, in *Computer Simulation of Liquids*, Oxford University Press, New York, 1987, pp 183-184.
336. B. Ravel and M. Newville, *J. Synchrotron Rad.*, 2005, **12**, 537-541.
337. A. L. Ankudinov, A. I. Nesvizhskii and J. J. Rehr, *Phys. Rev. B*, 2003, **67**, 115120.
338. M. Newville, *J. Synchrotron Rad.*, 2001, **8**, 322-324.
339. E. A. Stern, *Phys. Rev. B*, 1993, **48**, 9825-9827.
340. E. A. Hudson, P. G. Allen and L. J. Terminello, *Phys. Rev. B*, 1996, **54**, 156-165.
341. M. A. Karolewski, R. G. Cavell, R. A. Gordon, C. J. Glover, M. Cheah and M. C. Ridgway, *J. Synchrotron Rad.* , 2013, **20**, 555-566.
342. K. Provost, E. C. Beret, D. Bouvet Muller, A. Michalowicz and E. Sánchez Marcos, *J. Chem. Phys.*, 2013, **138**, 084303.
343. T. R. Cuya Guizado, S. R. W. Louro and C. Anteneodo, *J. Chem. Phys.*, 2011, **134**, 055103.
344. A. Laaksonen, P. Stilbs and R. E. Wasylishen, *J. Chem. Phys.*, 1998, **108**, 455-468.
345. W. R. Scheidt and Y. J. Lee, *Struct. Bond*, 1987, **64**, 1-70.
346. S. K. Ghosh, R. Patra and S. P. Rath, *Inorg. Chem.*, 2008, **47**, 10196-10198.
347. A. C. de Dios, L. B. Casabianca, A. Kosar and P. D. Roepe, *Inorg. Chem.*, 2004, **43**, 8078-8084.
348. V. A. Otelo, A. C. Sant'Ana, D. L. A. de Faria and C. M. S. Menezes, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 250-254.
349. K. Vanommeslaeghe and A. D. MacKerell Jr., *J. Chem. Inf. Model.*, 2012, **52**, 3144-3154.
350. K. Vanommeslaeghe, E. P. Raman and A. D. MacKerell Jr., *J. Chem. Inf. Model.*, 2012, **52**, 3155-3168.
351. *ParamChem interface: https://www.paramchem.org*, Accessed 19 December 2013.
352. F. Partal, M. Fernández-Gómez, J. J. López-González, A. Navarro and G. J. Kearley, *Chem. Phys.*, 2000, **261**, 239-247.
353. A. Topaçli and S. Bayari, *Spectrochim. Acta A*, 1999, **55**, 1389-1394.
354. R. E. Wilde, *Am. J. Phys.*, 1964, **32**, 45-52.
355. P. Pulay, G. Fogarasi, F. Pang and J. E. Boggs, *J. Am. Chem. Soc.*, 1979, **101**, 2550-2560.
356. I. Solomon, *Phys. Rev.*, 1955, **99**, 559-565.
357. I. Solomon and N. Bloembergen, *J. Chem. Phys.*, 1956, **25**, 261-266.
358. L. Banci, M. Piccioli and A. Scozzafava, *Coord. Chem. Rev.*, 1992, **120**, 1-28.
359. I. Bertini, C. Luchinat and A. Rosato, *Prog. Biophys. Mol. Biol.*, 1996, **66**, 43-80.
360. F. Dubar, T. J. Egan, B. Pradines, D. Kuter, K. K. Ncokazi, D. Forge, J.-F. Paul, C. Pierrot, H. Kalamou, J. Khalife, E. Buisine, C. Rogier, H. Vezin, I. Forfar, C. Slomianny, X. Trivelli, S. Kapishnikov, L. Leiserowtiz, D. Dive and C. Biot, *ACS Chem. Biol.*, 2011, **6**, 275-287.
361. G. Otting, *Ann. Rev. Biophys.*, 2010, **39**, 387-405.
362. I. Bertini, C. Luchinat, G. Parigi and R. Pierattelli, *Dalton Trans.*, 2008, 3761-3908.
363. I. Bertini, C. Luchinat and G. Parigi, *Prog. Nucl. Mag. Res. Sp.*, 2002, **40**, 249-273.
364. G. A. Jeffrey, in *An introduction to hydrogen bonding*, Oxford University Press, Oxford, 1997.
365. T. Steiner, *Angew. Chem., Int. Ed.*, 2002, **41**, 48-76.
366. L. Brammer, E. A. Bruton and P. Sherwood, *Cryst. Growth Des.*, 2001, **1**, 277-290.
367. J.-A. van den Berg and K. R. Seddon, *Cryst. Growth Des.*, 2003, **3**, 643-661.
368. A. Kovács and Z. Varga, *Coord. Chem. Rev.*, 2006, **250**, 710-727.
369. C. L. D. Gibb, E. D. Stevens and B. C. Gibb, *J. Am. Chem. Soc.*, 2001, **123**, 5849-5850.
370. Z. R. Laughrey, C. L. D. Gibb, T. Senechal and B. C. Gibb, *Chem. Eur. J.*, 2003, **9**, 130-139.
371. MarvinSketch, 5.5.1.0, ChemAxon Ltd, Budapest, 2011.
372. C. A. Hunter and J. K. M. Sanders, *J. Am. Chem. Soc.*, 1990, **112**, 5525-5534.
373. C. A. Hunter, K. R. Lawson, J. Perkins and C. J. Urch, *J. Chem. Soc. Perkin Trans.*, 2001, **2**, 651-699.

374. S. R. Vippagunta, A. Dorn, H. Matile, A. K. Bhattacharjee, J. M. Karle, W. Y. Ellis, R. G. Ridley and J. L. Vennerstrom, *J. Med. Chem.*, 1999, **42**, 4630-4639.
375. D. O'Hagan, *Chem. Soc. Rev.*, 2008, **37**, 308-319.
376. D. Dziedzic-Kocurek, H. J. Byrne, A. Świderski and J. Stanek, *Acta Physica Polonica A*, 2009, **115**, 552-555.

APPENDICES

Appendix A

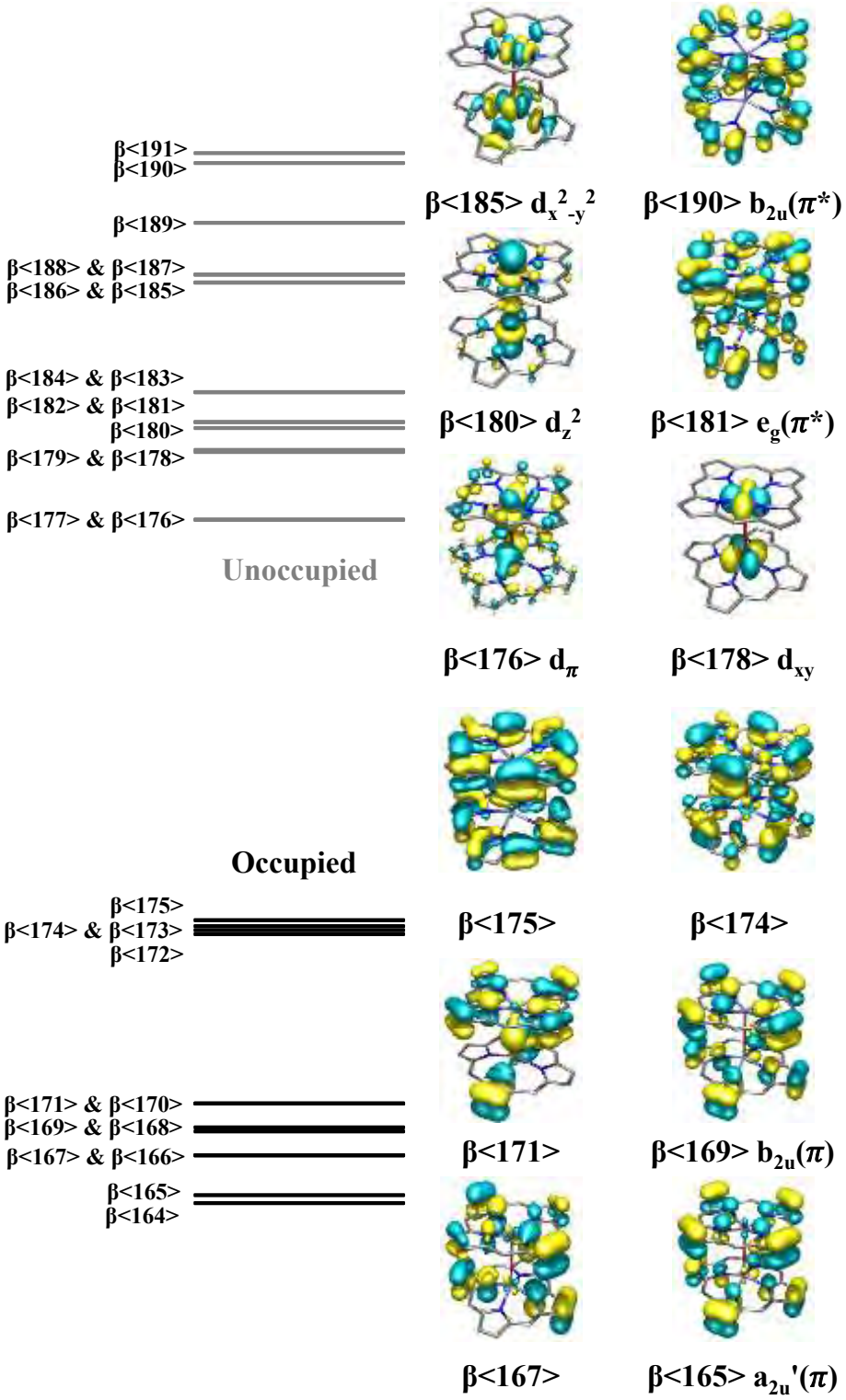


Figure A1. β-Spin MO energy diagram of high-spin [μ-oxo ferriporphine]^0 (left) and selected MOs (right).

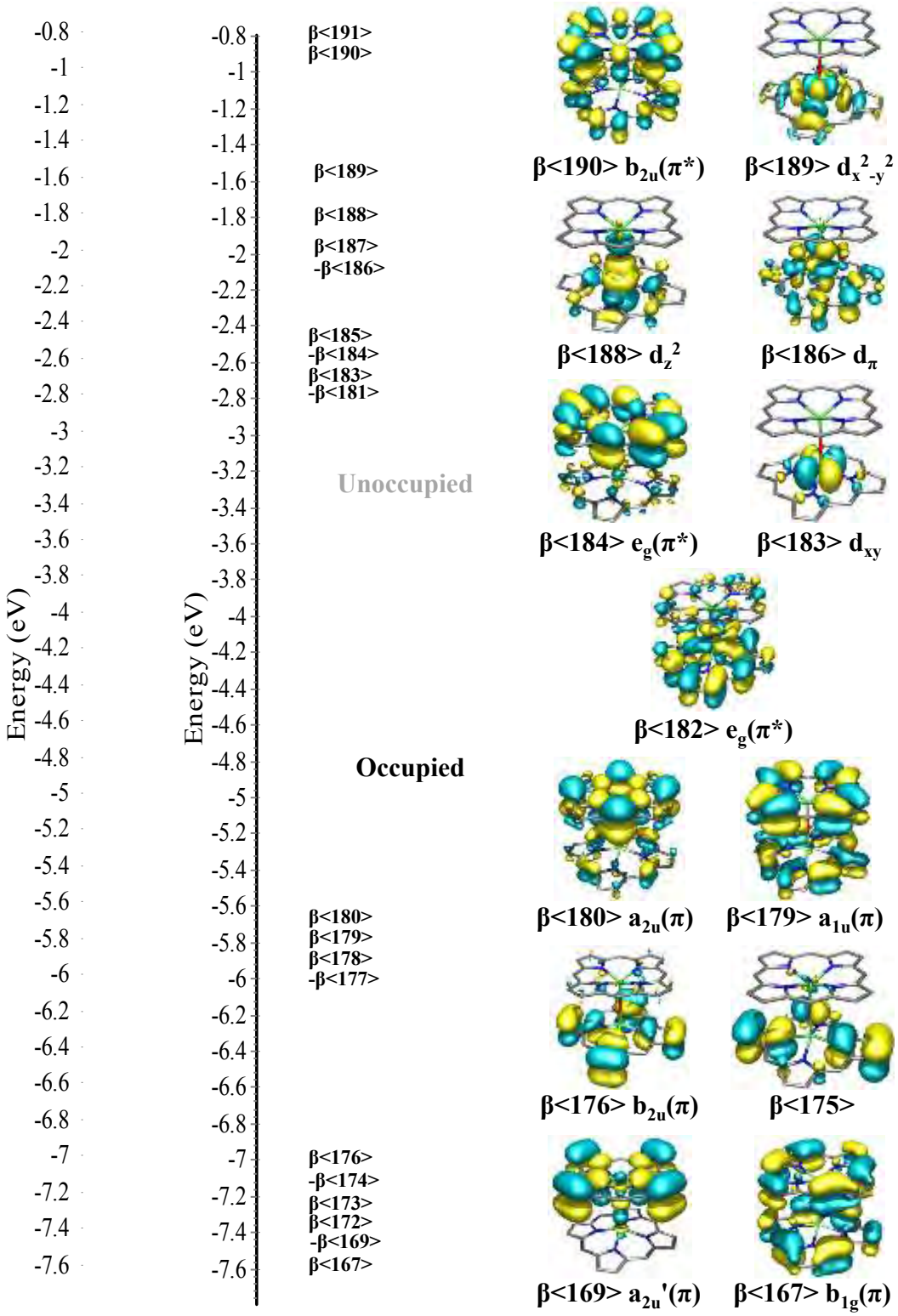


Figure A2. β-Spin MO energy diagram of antiferromagnetically coupled [μ-oxo ferriporphine]^{0} (left) and selected MOs (right).

Table A1.Compositions of β-spin MOs of paramagnetic [μ-oxo ferriporphine]^{0}. [a]

Orbital	Label [b]	Energy (eV)	Composition (%)				
			Fe		Porphyrin		O^{2-}
			1	2	1	2	
β<191>	$b_{2u}(\pi^*)$	-0.830	0	0	50	50	0
β<190>	$b_{2u}(\pi^*)$	-0.989	0	0	50	50	0
β<189>	d_z^2	-1.285	27	28	12	12	20
β<188>	d_π	-1.624	29	29	15	14	13
β<187>	d_π	-1.625	29	29	14	14	13
β<186>	$d_{x^2-y^2}$	-1.676	11	71	11	37	0
β<185>	$d_{x^2-y^2}$	-1.679	40	11	37	11	0
β<184>	$e_g(\pi^*)$	-2.387	17 [d]	17 [d]	33	34	0
β<183>	$e_g(\pi^*)$	-2.388	17 [d]	17 [d]	33	34	0
β<182>	$e_g(\pi^*)$	-2.581	12 [d]	12 [d]	37	37	1
β<181>	$e_g(\pi^*)$	-2.581	12 [d]	12 [d]	37	37	1
β<180>	d_z^2	-2.612	31	30	18	18	3
β<179>	d_{xy}	-2.758	48	40	6	6	0
β<178>	d_{xy}	-2.769	35	42	12	12	0
β<177>	d_π	-3.213	21	21	29	29	0
β<176> LUMO	d_π	-3.217	21	21	29	29	0
β<175> HOMO	– [c]	-5.812	1	1	49	49	0
β<174>	– [c]	-5.843	1	1	47	50	0
β<173>	– [c]	-5.865	1	1	50	47	0
β<172>	– [c]	-5.898	1	1	49	49	0
β<171>	– [c]	-6.993	2	2	38	39	18
β<170>	– [c]	-6.993	2	2	38	39	18
β<169>	$b_{2u}(\pi)$	-7.147	0	0	49	51	0
β<168>	$b_{2u}(\pi)$	-7.171	0	0	51	49	0
β<167>	– [c]	-7.332	2	2	48	48	0
β<166>	– [c]	-7.333	2	2	48	48	0
β<165>	$a_{2u}'(\pi)$	-7.589	1	1	48	49	0

[a] All orbitals below the dotted horizontal line are filled; [b] symmetry labels refer to related orbitals in monomeric porphine which has idealized D_{4h} symmetry; [c] no corresponding porphine orbital can be identified; [d] Fe p_x and p_y.

Table A2.Compositions of β-spin MOs of antiferromagnetically coupled [μ-oxo ferriporphine]^{0}. [a, b]

Orbital	Label [c]	Energy (eV)	Composition (%)				
			Fe		Porphyrin		O^{2-}
			1	2	1	2	
β<191>	$b_{2u}(\pi^*)$	-0.823	0	0	58	42	0
β<190>	$b_{2u}(\pi^*)$	-0.908	0	0	38	62	0
β<189>	$d_{x^2-y^2}$	-1.578	52	0	48	1	0
β<188>	d_{z^2}	-1.806	56	3	16	12	14
β<187>	d_{π}	-2.057	29	2	34	33	2
β<186>	d_{π}	-2.060	29	2	34	33	2
β<185>	$e_g(\pi^*)$	-2.583	2	23 [e]	22	53	0
β<184>	$e_g(\pi^*)$	-2.583	2	23 [e]	22	54	0
β<183>	d_{xy}	-2.629	79	0	10	12	0
β<182>	$e_g(\pi^*)$	-2.729	29 [e]	7	44	19	1
β<181> LUMO	$e_g(\pi^*)$	-2.731	29 [e]	7	44	19	1
β<180> HOMO	$a_{2u}(\pi)$	-5.734	0	3	39	57	1
β<179>	$a_{1u}(\pi)$	-5.787	1	0	33	66	0
β<178>	$a_{2u}(\pi)$	-5.862	2	1	84	13	0
β<177>	$a_{1u}(\pi)$	-5.877	1	0	48	51	0
β<176>	$b_{2u}(\pi)$	-7.153	0	0	91	8	0
β<175>	– [d]	-7.162	3	1	82	4	10
β<174>	– [d]	-7.163	3	1	82	4	10
β<173>	$b_{2u}(\pi)$	-7.213	0	8	6	86	0
β<172>	– [d]	-7.307	0	6	4	90	0
β<171>	– [d]	-7.308	0	6	4	90	0
β<170>	– [d]	-7.377	0	12	6	82	0
β<169>	$a_{2u}'(\pi)$	-7.435	8	2	83	1	5
β<168>	$a_{2u}'(\pi)$	-7.611	1	2	13	83	1

[a] All orbitals below the dotted horizontal line are filled; [b] β-spin MOs are equivalent to α-spin MOs; [c] symmetry labels refer to related orbitals in monomeric porphine which has idealized D_{4h} symmetry; [d] no corresponding porphine orbital can be identified; [e] Fe p_x and p_y.

Appendix B

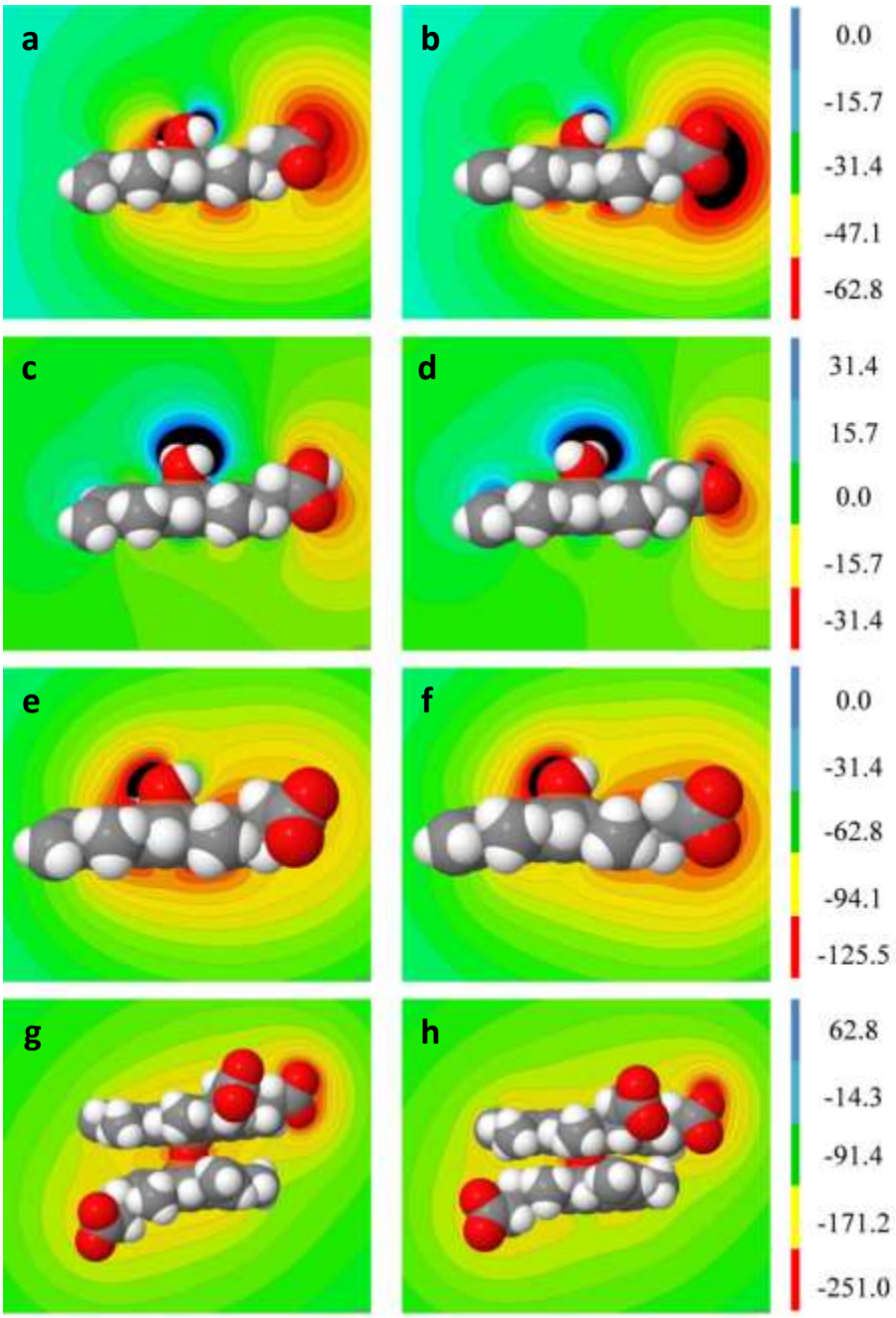


Figure B1. QM- (left) and MM-generated (right) ESPs of ferrihaem species. (**a, b**) $[H_2O\text{-ferrihaem}]^{1-}$; (**c, d**) $[H_2O\text{-ferrihaem}]^{0}$; (**e, f**) $[HO\text{-ferrihaem}]^{2-}$; (**g, h**) [μ-oxo ferrihaem]$^{4-}$. MM ESPs were generated for MM-minimized structures. ESPs were generated using the Jmol software package. Energies displayed are in kcal mol^{-1}.

Table B1. WIEs (kcal mol^{-1}) for monomeric ferrihaem species calculated using optimized atomic charges in CHARMM and using DFT (OPBE/6-31G*). MM-minimised geometries were used to calculate interaction energies.

		Water Interaction Energy					
Interaction Type [a]		**$[H_2O$-ferrihaem$]^{0}$** [b]		**$[H_2O$-ferrihaem$]^{1-}$** [c]		**$[HO$-Ferrihaem$]^{2-}$** [c]	
		QM	**CHARMM**	**QM**	**CHARMM**	**QM**	**CHARMM**
1	COO^---- H-OH	-7.41	-5.94	-10.42	-9.69	-16.30	-13.49
2	COOH --- OH_2	-7.64	NA [d]		-		-
3	C=O --- H-OH	-4.85	-2.35		-		-
4	H_{meta} ---OH_2	*Spin Contaminated*		-3.00	-4.07	*Spin Contaminated*	
5	H_{meta} ---OH_2	-2.95	-2.56	-0.71	-0.86	*Spin Contaminated*	
6	H_{meta} ---OH_2	-2.98	-3.52	-0.77	-1.82	0.94	0.63
7	H_{meta} ---OH_2	-1.30	-1.20	0.74	0.56	2.11	2.38
8	H_{vinyl} --- OH_2	-2.34	-2.17	-0.75	-0.72	-0.75	1.21
9	H_{vinyl} --- OH_2	-1.81	-1.53	-0.17	-0.17	0.97	1.33
10	O-H_{ax} --- OH_2	-11.45	-9.43	-6.87	-5.81	-1.95	-1.32
11	O_{ax} --- H-OH	-2.48	0.40	-2.55	-0.33	-8.47	-9.00
12	N --- H-OH	-2.41	-1.85	-2.87	-4.92	-3.25	-4.30
13	Fe --- OH_2	-3.58	-3.44	-2.55	-0.33	2.77	0.72
	MUE	*1.00*		*0.96*		*1.11*	

[a] See figure 5.3 for numbering; [b] QM interaction energy scaled by 1.12*1.16=1.30; [c] QM interaction energy scaled by 1.12; [d] Not available – intermolecular hydrogen bonding precluded measurement.

Table B2. Statistics and refined parameters determined from fitting the solid-state μ-oxo dimer EXAFS spectrum. [a, b]

Refinement Statistics			
N_{ind}		20.73	
N_{var}		14	
$\chi(r_i)^2$		1.11	
R-factor		0.009	
ε_k		0.0012	
Fitted parameters	*N*	ΔR	σ^2
Fe-O_{oxo}	1	0.044(8)	0.0023(6)
Fe-N	4	-0.021(5)	0.0037(9)
Fe-C_α	8	-0.026(8)	0.005(1)
Fe-C_β	8	0.10(2)	0.004(4)
Fe-C_{meso}	4	0.03(4)	0.004(4)
Fe-Fe	1	0.0(2)	0.01(2)

[a] Error given in parenthesis; [b] see Section 5.2.3 for explanation of parameters.

Table B3. Statistics and refined parameters determined from fitting the solution μ-oxo dimer EXAFS spectrum. [a, b]

Statistics and fitted parameters	Solvated [c]	No solvation [d]
N_{ind}	21.25	21.25
N_{var}	12	7
$\chi(r_i)^2$	1.26	11.25
R-factor	0.005	0.069
ε_k	0.0015	0.0015
α	–0.012(2)	–0.004(4)
α_{oxo}	0.029(7)	0.05(1)
α_{Fe}	–0.05(1)	–0.02(5)
$\alpha(1)_{wat}$ [e]	–0.174(3)	– [g]
$\alpha(2)_{wat}$ [f]	0.072(4)	– [g]
β	1.3(4)	0.7(6)
β_{oxo}	2.5(8)	2(1)

[a] Error given in parenthesis; [b] see Section 5.2.3 for explanation of parameters; [c] Fitted using the MD data including five water molecules at positions identified from the SDF; [d] Fitted using the MD data without solvent water molecules; [e] water molecule positioned on the unligated face of a given ferrihaem molecule; [f] water molecules surrounding the bridging oxide ligand; [g] not applicable.

Appendix C

Figure C1. Internal coordinates of *N*-methyl-4-aminopyridinium used in describing vibrations in terms of symmetry coordinates (see Table C1). Bonds (**a**), angles (**b**) and dihedrals (**c**) are coloured in red, blue and green respectively. In (b) $\beta_i = \phi_i - \phi_i'$ and in (c) γ_i refers to improper dihedrals. Double bonds have been removed in (b) and (c) for clarity.

Table C1. Symmetry coordinates used to describe vibrations of *N*-methyl-4-aminopyridinium.[a, b]

$\mathbf{V_1}$	$r_{11} + r_{12}$	$\mathbf{V_{16}}$	r_{19}	$\mathbf{V_{31}}$	$\gamma_3 + \gamma_4$
$\mathbf{V_2}$	$r_{13} + r_{14}$	$\mathbf{V_{17}}$	r_{10}	$\mathbf{V_{32}}$	γ_5
$\mathbf{V_3}$	r_8	$\mathbf{V_{18}}$	$r_2 - r_3$	$\mathbf{V_{33}}$	$\tau_7 - \tau_8$
$\mathbf{V_4}$	r_1	$\mathbf{V_{19}}$	$r_4 - r_5$	$\mathbf{V_{34}}$	$\tau_2 - \tau_4 + \tau_6 - \tau_5 + \tau_3 - \tau_1$
$\mathbf{V_5}$	$r_2 + r_3$	$\mathbf{V_{20}}$	$r_6 - r_7$	$\mathbf{V_{35}}$	$\tau_2 - \tau_6 + \tau_5 - \tau_1$
$\mathbf{V_6}$	$r_4 + r_5$	$\mathbf{V_{21}}$	β_1	$\mathbf{V_{36}}$	$-\tau_2 + 2\tau_4 - \tau_6 - \tau_5 + 2\tau_3 - \tau_1$
$\mathbf{V_7}$	$r_6 + r_7$	$\mathbf{V_{22}}$	β_7	$\mathbf{V_{37}}$	r_{15}
$\mathbf{V_8}$	$\alpha_1 - \alpha_2 + \alpha_4 - \alpha_6 + \alpha_5 - \alpha_3$	$\mathbf{V_{23}}$	β_6	$\mathbf{V_{38}}$	r_{16}
$\mathbf{V_9}$	$2\alpha_1 - \alpha_2 - \alpha_4 + 2\alpha_6 - \alpha_5 - \alpha_3$	$\mathbf{V_{24}}$	$\beta_2 + \beta_3$	$\mathbf{V_{39}}$	r_{17}
$\mathbf{V_{10}}$	$\alpha_2 - \alpha_4 + \alpha_5 - \alpha_3$	$\mathbf{V_{25}}$	$\beta_4 + \beta_5$	$\mathbf{V_{40}}$	$\alpha_8 + \alpha_9 + \alpha_{10} - \beta_8 - \beta_9 - \beta_{10}$
$\mathbf{V_{11}}$	$2\alpha_7 - \varphi_7 - \varphi_7'$	$\mathbf{V_{26}}$	$\gamma_1 - \gamma_2$	$\mathbf{V_{41}}$	$2\alpha_8 - \alpha_9 - \alpha_{10}$
$\mathbf{V_{12}}$	$\beta_2 - \beta_3$	$\mathbf{V_{27}}$	$\gamma_3 - \gamma_4$	$\mathbf{V_{42}}$	$\alpha_9 - \alpha_{10}$
$\mathbf{V_{13}}$	$\beta_4 - \beta_5$	$\mathbf{V_{28}}$	γ_0	$\mathbf{V_{43}}$	$2\beta_8 - \beta_9 - \beta_{10}$
$\mathbf{V_{14}}$	$r_{11} - r_{12}$	$\mathbf{V_{29}}$	$\tau_7 + \tau_8$	$\mathbf{V_{44}}$	$\beta_9 - \beta_{10}$
$\mathbf{V_{15}}$	$r_{13} - r_{14}$	$\mathbf{V_{30}}$	$\gamma_1 + \gamma_2$	$\mathbf{V_{45}}$	$\tau_9 + \tau_{10} + \tau_{11}$

[a] see Fig. C1 for internal coordinate labelling; [b] $\beta_i = \phi_i - \phi_i'$.

Table C2. Statistics and refined parameters determined from fits to EXAFS spectra of the CQ-ferrihaem complex in π-stacked and docked conformations. [a]

Parameter	Fitted value		
	π-Stacked [d]	**Docked** [e]	**Docked** [f]
N_{ind}	15.79	16.22	16.22
N_{var}	12	13	7
$\chi(r_i)^2$	6.80	0.77	3.14
R-factor	0.089	0.008	0.100
ε_k	0.004	0.004	0.004
S_0^2	0.7(4)	0.8(1)	0.9(2)
ΔE_0	5(5)	3(1)	2(2)
α	–0.02(2)	–0.004(4)	–0.013(8)
α_{oxo}	0.04(6)	–0.03(1)	–0.04(2)
α_{Fe}	–0.01(9)	–0.05(3)	–0.06(3)
α_{CQ}	–0.0(4)	0.08(4)	– [g]
$\alpha(1)_{wat}$ [b]	– [g]	–0.279(9)	– [g]
$\alpha(2)_{wat}$ [c]	–0.01(2)	–0.23(1)	– [g]
β	1(3)	1(1)	1(1)
β_{oxo}	2(8)	1.5(8)	1.70(1)
β_{Fe}	1(12)	1.5(8)	1.70(1)
β_{CQ}	0.005(2)[h]	0.3(2)	– [g]

[a] Errors in parenthesis, see section 5.2.2 for explanation of parameters; [b] Water molecule situated on unligated face; [c] Water molecules situated around μ-oxo bridge; [d] Fitted using the MD data including three water molecules at positions identified from the SDF and N1, C2 and C8A atoms of CQ; [e] Fitted using the MD data including solvent water molecules and CQ N1, C2 and C8A atoms at positions identified from the SDF; [f] Fitted using the MD data excluding solvent water molecules and CQ N1, C2 and C8A atoms; [g] Not applicable; [h] value corresponds to σ^2_{CQ}.