Australian National University

Research School of Chemistry

Doctorate of Philosophy

Computational Investigation of the Oxygen Evolving Complex of Photosystem II and Related Models *via* Density Functional Theory

Doctoral Dissertation of:

Richard Norman Leslie Terrett

Chair of Supervisory Panel:

Prof. Robert Stranger

Advisors:

Prof. Ronald J. Pace

Dr Simon Petrie

Supervisor of the Doctoral Program:

Prof. Gottfried Otting

April 2017

Declarations

Except where specific reference is made to other sources, the work presented in this dissertation is the work of the author. It has not been submitted, in whole or in part, for any other degree.

_____

Richard Norman Leslie Terrett

Declarations regarding individual publications

The following citations comprising this thesis by compilation are valid and the contributions of the candidate to the research and authorship of each are accurately enumerated in their accompanying introductions.

I. Terrett, R.; Petrie, S., Pace, R.J.; Stranger, R.; *What does the Sr-substituted 2.1 Å resolution crystal structure of photosystem II reveal about the water oxidation mechanism?*; Chemical Communications 50 (**2014**) pp. 3187–3190 doi: 10.1039/C3CC49324E (**In Press**)

Richard Norman Leslie Terrett	Prof. Rob Stranger (senior author)

II. Terrett, R.; Frankcombe, T., Pace, R.; Stranger, R.; *Effect of concomitant oxidation and deprotonation of hydrated Mn centres in rationalising the FTIR difference silence of D1-Asp170 in Photosystem II*; Journal of Inorganic Biochemistry 155 (**2016**) pp. 101–104 doi: 10.1016/j.jinorgbio.2015.11.023 (**In Press**)

Richard Norman Leslie Terrett

Prof. Rob Stranger

(senior author)

III. Terrett, R.; Petrie, S.; Stranger, R.; Pace, R.J.; *What computational chemistry and magnetic resonance reveal concerning the oxygen evolving centre in Photosystem II*; Journal of Inorganic Biochemistry 162 (**2016**) pp. 178–189; doi: 10.1016/j.jinorgbio.2016.04.009 (**In Press**)

Richard Norman Leslie Terrett	Prof. Ron J. Pace (senior author)

IV. Terrett, R.; Stranger, R.; Frankcombe, T.; Pace, R.J.; *Vibrational Intensities in the Mobile Block Hessian Approximation*; 19 (**2017**) pp. 6654–6664; doi: 10.1039/C6CP07498G (**In Press**)

Richard Norman Leslie Terrett

Prof. Rob Stranger

(senior author)

Preface

This thesis consists of a compilation of publications relating to the computational investigation of the Mn_4CaO_5 cofactor of the Photosystem II (PSII) holoenzyme, which is implicated in the light-driven four-electron oxidation of water ($2H_2O \longrightarrow 4e^- + O_2 + 4H^+$). This reaction constitutes the first step of photosynthetic metabolism in the chloroplasts of plants and cyanobacteria.

Photosynthetic water oxidation is of great practical interest as it provides a concrete example of facile hydrogen and oxygen production through a photo-assisted process. Whereas approximately 90% of hydrogen production in the world is presently achieved through catalytic reforming of fossil fuels (with a minority achieved through brine electrolysis), adapting insights from photosynthetic water oxidation is anticipated to furnish significantly more economic and sustainable routes to the production of hydrogen. This hydrogen may either be used directly as a fuel or as an intermediate product in the synthesis of value-added products.

Over the past decades, substantial progress has been made in understanding the structural nature of the photosynthetic apparatus as it pertains to water oxidation. This progress is in large part due to the refinement of preparative protocols, X-Ray Diffractometry (XRD), site-directed mutagenesis, and the coming of age of X-ray Free Electron Laser (XFEL)

diffractometry in the investigation of PSII. Nevertheless, key mechanistic and electronic details of water oxidation still remain highly contentious. Elucidation of these details is complicated by the fact that the active site of Photosystem II exists in four natural metastable oxidation states, as well as putative unnatural forms that are plausibly induced during experimental investigation.

Experimental investigation of PSII exists in symbiosis with competing theoretical justifications for observed spectroscopic, metabolic and diffractometric behaviour. Density Functional Theory (DFT) is an important tool in the computational chemist's arsenal, permitting investigation of the electronic structure of complex models of bio-inorganic cofactors whilst maintaining good computational efficiency and scalability on extant computational hardware. The size and fidelity of computational models of PSII have necessarily been limited by the rate of increase in central processing unit (CPU) speed and memory capacity. Nevertheless, the explosive increase in available computational resources and efficiency has resulted in a flourishing of increasingly sophisticated computational models of components and features of the photosynthetic apparatus. Integration of these models into a coherent and consistent picture of the photosynthetic mechanism is a colossal multidisciplinary and multilateral undertaking. Even the incomplete understanding of biological photosynthesis obtained thus far has led to the development of a large

number of functional mimics of the Photosystem II OEC. In turn, the wide variety of morphologies and chemistries of these water oxidation catalysts provides insight into the irreducible features of photosynthetic systems.

Acknowledgements

This project would have been impossible or at the very least much more onerous in the absence of a number of individuals and organisations who lent technical, financial, and emotional support to this undertaking. In the first instance, my supervisory panel (composed of Prof. Rob Stranger, Prof. Ron Pace and Dr Simon Petrie) have furnished me with expert guidance and supervision, have kept me on track and have demonstrated a considerable degree of perserverence and understanding in accomodating my program, for which they should be well-credited.

My parents (Les and Vikki Terrett) have contributed considerable financial and emotional support without which this endeavour would have been impossible, whilst my girlfriend, Dr Erin Walsh, has been my constant companion and muse throughout this program. Her family (Kathy and Bernard Walsh) have also been a wellspring of hospitality and encouragement. Erin's coworkers at the Research School of Psychology (Dr Yiyun *'Cloudy'* Shou and Prof. Michael Smithson) have further offered stimulating and friendly discussion during my interactions with them.

A number of other academics at RSC should be particularly distinguished for technical advice and companionship indispensable to my activities, including Dr Germán Cavigliasso, Prof. Michael Collins, Dr Terry Frankcombe, Dr Andrew Gilbert, Prof. Elmars Krausz, Prof. Ray

Withers, and Dr Lasse Noren. Additionally, a number of PhD, honours and undergraduate research students who have passed through and each contributed to the accumulated formal and informal wisdom and camaraderie of the group should be acknowledged, including but not limited to Dr Gemma Christian, Dr Phillip Gatt, Dr Erandi Kulasekara, Dr Fazira Razak, Dr Jinasena Hewage, Dr Somayeh Rastegar, Mahesh Kodikara, Bernard Baituti, Zöe Buxton, Amy Kendrick, Kassetra von Nessi, and Maki Maeda. I would like to thank the RSC IT department (Dr Christopher Delfs, Dr Graeme Lindsell, and Hamish McArthur–Onslow) and members of the RSC front office, including Gavin Perri and Allison Scott, for keeping everything running smoothly. The NCI National Facility staff—particularly Dr Rika Kobayashi—should be commended for their preternatural commitment to the upkeep and uptime of the XE, AC, Vayu and Raijin supercomputers.

I would furthermore like to thank the developers of the Amsterdam Density Functional (ADF) software, who have always been willing to provide expert support for a well-crafted DFT code. Computational chemistry software is a recherché field, however ADF comports itself with the *Principle of Least Astonishment* much more often than many other packages I have used.

Finally, this research was supported variously by an Australian Postgraduate Award (furnished by the Australian Federal Government) an

ANU PhD Scholarship (furnished by the Australian National University), the Adrien Albert Honours Prize, the RSC Supplementary Scholarship, and the A.S. Sargeson Supplementary Scholarship (furnished by the Australian National University Research School of Chemistry).

Richard Norman Leslie Terrett

Canberra

April 2017

Colophon

This document was typeset using the PDFLaTeX typesetting system. It is based on modifications of a template graciously provided by Federico Maggi, which is in turn based on the `memoir` class created by Peter Wilson.

The document is typeset using the Johannes Kepler fonts by Christophe Caignaert.

Illustrations appearing in this document were variously created using the Inkscape vector graphics suite, the Blender 3D graphics suite and Cycles rendering engine, the PyMOL chemical visualisation suite created by Warren DeLano, and the R project for statistical computing. HDRi maps for illumination of 3D graphics were graciously provided by Alejo Grigera.

The bibliography of this dissertation is presented in the IEEE style, using the BibLaTeX package. If you are reading a digital edition of this dissertation, you should be able to click in-text citations to reach their respective entries in the bibliography, as well as click digital object identifier (DOI) entries to conveniently access them online, where applicable.

Abstract

The first step of photosynthetic metabolism effects the facile oxidation of water to dioxygen and hydrogen cations. This is achieved through an incompletely-understood process of light-driven four-electron oxidation at the Mn_4CaO_5 cofactor of the Oxygen-Evolving Complex (OEC) of the Photosystem II (PSII) holoenzymatic complex in photosynthetic autotrophs. Biomimesis of this reaction—*artificial photosynthesis*—may offer energy-efficient routes to industrial hydrogen generation and value-added derivatives, with implications for solar energy fixation. This thesis consists of a compilation of four publications relating to Density Functional Theory (DFT) studies of structural and spectroscopic aspects of the Oxygen-Evolving Complex (OEC) of PSII. These publications consist of research resolving the basis of structural anomalies in metal-substituted PSII, combinatoric simulation of difference spectra corresponding to proton-coupled oxido-reduction scenarios of PSII models, simulation of the hyperfine and superexchange magnetic interactions in PSII models, and the development of a methodology for obtaining vibrational intensities in the Mobile Block Hessian (MBH) approximation, with applications to accelerated modeling of the vibrational structure of complex models of PSII, as well as other large molecules. These publications are presented alongside explanatory introductions and preceded by a general survey of the state of the art of

photosynthesis research, context for the relevance of this research, and methodological discussion. Concluding remarks are also presented.

Contents

List of Figures

List of Tables

Acronyms

ADF Amsterdam Density Functional. 27, 118, 119, 121, 141

ASIC Application Specific Integrated Circuit. 67

CE Common Era. xxv, 40

CI Configuration Interaction. 72

COSMO Conductor-Like Solvent MOdel. 82

CPU central processing unit. vii, 62, 66, 68

CUDA Compute Unified Device Architecture. 67

DAG directed acyclic graph. 104

DFT Density Functional Theory. vii, xiii, 60, 62, 71–74, 86, 161, 163

DOI digital object identifier. xii

DSSC dye-sensitised solar cell. 28–33

ECP Effective Core Potential. 80

EEZ exclusive economic zone. 44

ENDOR Electron-Nuclear Double Resonance. 27

EPR Electron Paramagnetic Resonance. 10, 13, 14, 27, 116, 119, 121, 122

ESR Electron Spin Resonance. 13

EXAFS Extended X-ray Absorption Fine Structure. 10–12, 25, 164

FEA Finite Element Analysis. 69

FPGA Field Programmable Gate Array. 68

FRET Förster Resonance Energy Transfer. 3

FTIR Fourier transform infrared. 100, 101, 104, 142, 163, 164

GGA Generalised Gradient Approximation. 75, 119

GPGPU General Purpose processing on Graphical Processing Units. 67, 68

HDvV Heisenberg–Dirac–van-Vleck. 116

HTE High Temperature Electrolysis. 52

IPCC Intergovernmental Panel on Climate Change. 40

LDA Local Density Approximation. 74, 119

LHC Light-Harvesting Complex. 3

MBH Mobile Block Hessian. xiii, 69, 85, 138, 141–143, 164

MD Molecular Dynamics. 62, 65–68, 86

ML multiline. 14

MM Molecular Mechanics. 65, 67, 86

OEC Oxygen-Evolving Complex. xiii, xvi, xvii, xxiv, 3, 4, 6–20, 22–24, 28, 88, 89, 91–93, 100–102, 108, 116–119, 142, 143, 159, 161–165

OPV organophotovoltaic. 29, 32

PCM Polarisable Continuum Model. 82

PES potential energy (hyper)surface. 71, 84

PHVA Partial Hessian Vibrational Analysis. 139

PSII Photosystem II. vi, vii, xiii, xxiv, 3–13, 15, 17, 18, 20–24, 33, 52, 88–92, 100–102, 108, 116–119, 142, 143, 159, 161–165

PW plane-wave. 78

QM Quantum Mechanics. 66, 69, 86

SES Solvent-Excluding Surface. 82

STP standard temperature and pressure. 51

VSEPR Valence Shell Electron Pair Repulsion. 66

VTK Visualisation Toolkit. 121

WOC Water-Oxidising Complex. 6

XC eXchange–Correlation. 73

XFEL X-ray Free Electron Laser. vi, 7, 12, 13

XRD X-Ray Diffractometry. vi, 5, 11, 12, 14

ZORA Zeroth Order Regular Approximation. 118

Photosynthesis, Natural and Artificial 1

Stand a little out of my sunshine

Diogenes of Sinope

1.1 What is Photosynthesis?

Schmitt and coworkers describe plant metabolism at its base as the *'capture of solar radiation and transformation into Gibbs free energy'*. [1] This sentiment expresses the underlying truth that the Earth lies in a thermodynamic gradient between the Sun and the cold vacuum of deep space, and that the ultimate provenance of almost all biochemical energy on Earth is the fixation of solar energy in the form of thermodynamically unstable chemical species. The metabolisms of heterotrophic life flow

on from herbivory of these autotrophs. There are notable exceptions to photosynthesis as an ultimate source of biological energy[a] however they are distinguished by their provinciality and transience, whereas stellar evolution is nothing if not ponderous.

The evolution of photosynthetic life on Earth precipitated epochal changes in Earth's biogeochemistry, including the considerable reconstitution of the ancient atmosphere. [12] It is photosynthesis that is responsible for almost all free dioxygen on the planet, and the generation of this oxygen by the emergence of photosynthetic organisms approximately 2.4 Ga before present is correlated with considerable changes in the palaeogeochemical record.[b] Notably, increasing oxygen concentrations in the atmosphere and oceans precipitated a mass extinction of anaerobic life. This transition in Earth's palaeobiogeochemical history is termed the *'oxygen catastrophe'*, and was also accompanied by a dramatic drop in the atmospheric concentration of CO_2 and H_2. [14] Nevertheless, the oxygenation of the atmosphere paved the way for aerobic respiration and the wide variety of complex lifeforms it supports.

[a] Including the symbiotic chemosynthetic ecosystems that have evolved to exploit the heat and chemistry of underwater volcanic features including so-called *'black smokers'* and recently discovered and quite chemically distinct *'Lost City'*-type features. [2], [3] There now exists evidence that reduced hydrocarbons supporting chemosynthetic microbiota at these sites may be in part produced abiotically *via* geological serpentinisation and Fischer–Tropsch pathways, and that chemosynthetic microbiota may extend quite some distance into the suboceanic mantle. [4]–[8] Additionally, submarine features involving the percolation of lower temperature hydrocarbons through the seabed such as so-called *'cold seeps'* and *'asphalt volcanoes'* are known to support ecosystems. [9]–[11]

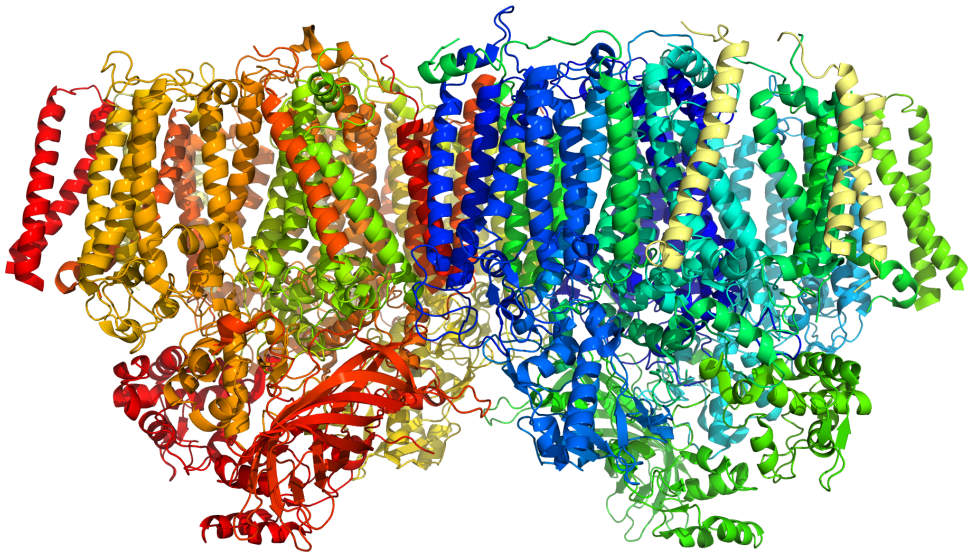
[b] Although see [13] for a recent presentation of evidence of nontrivial atmospheric oxygenation processes as early as 3 Ga b.p.

1.2 Natural Photosynthesis

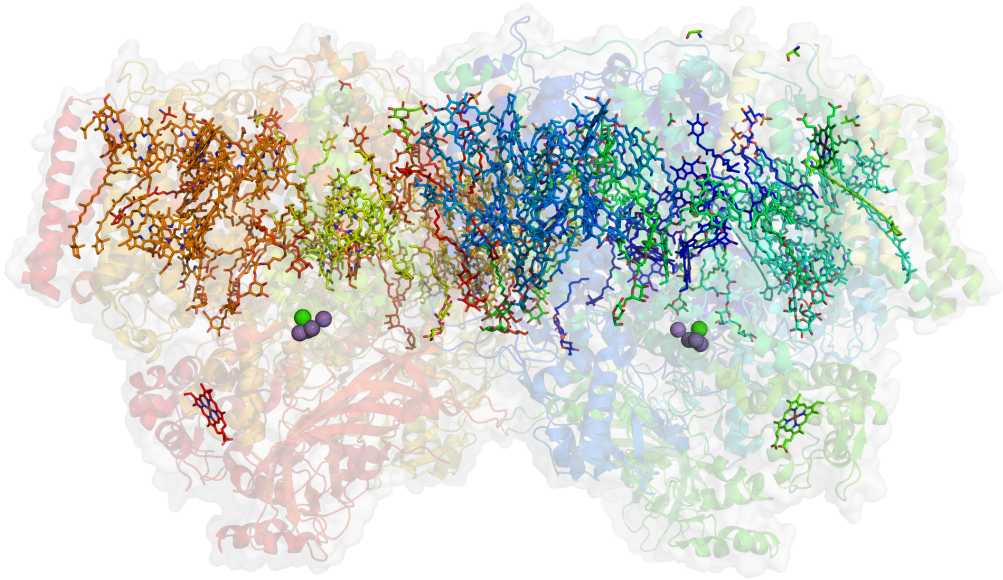
1.2.1 Photosystem II

Photosystem II (PSII) is a transmembrane holoenzyme (*i.e.* a supramolecular complex of multiple associated proteins) which transfixes the thylakoid membrane. The thylakoid (from Greek θύλακος, *'sac' + -oid*) is contained within the chloroplast, which is itself situated inside a photosynthetic cell. The schematic relationship of these components is illustrated in Figure 1.2. PSII is composed of 17 transmembrane subunits held together by supramolecular interactions, as well as a collection of extrinsic and integrated cofactors. [15] PSII is implicated in the photooxidation of water in plant metabolism. In this process, four oxidising equivalents are accumulated by the oxidation of manganese ions within the Oxygen-Evolving Complex (OEC), a bioinorganic cofactor within PSII. These oxidising equivalents are ultimately generated by a Förster Resonance Energy Transfer (FRET) process originating in Light-Harvesting Complex (LHC) proteins which are embedded proximate to PSII in the membrane, potentially supplemented by supramembrane phycobilisomes and other peripheral antenna units supplementing the light-harvesting efficiency of the gross photosynthetic apparatus. [16]

PSII is dimeric and possesses a quasi-C_2 symmetry element normal



a)



b)

Figure 1.1: Dimeric PSII holoenzyme, 300 pm structure of Loll, *et al.*, [17] viewed perpendicular to thylakoid normal. **a)** Polypeptide structure of PSII. Note the visually distinctive parallel α helices, which interface with and transfix the thylakoid membrane. **b)** Non-peptide cofactors of PSII, including chlorophylls, carotenoids, and the OEC Mn and Ca ions (represented as spheres) in the center-left and right of the figure. The OECs are related by an approximate C_2 rotation element, vertically perpendicular to the page normal. (PDB accession code 2AXT.)

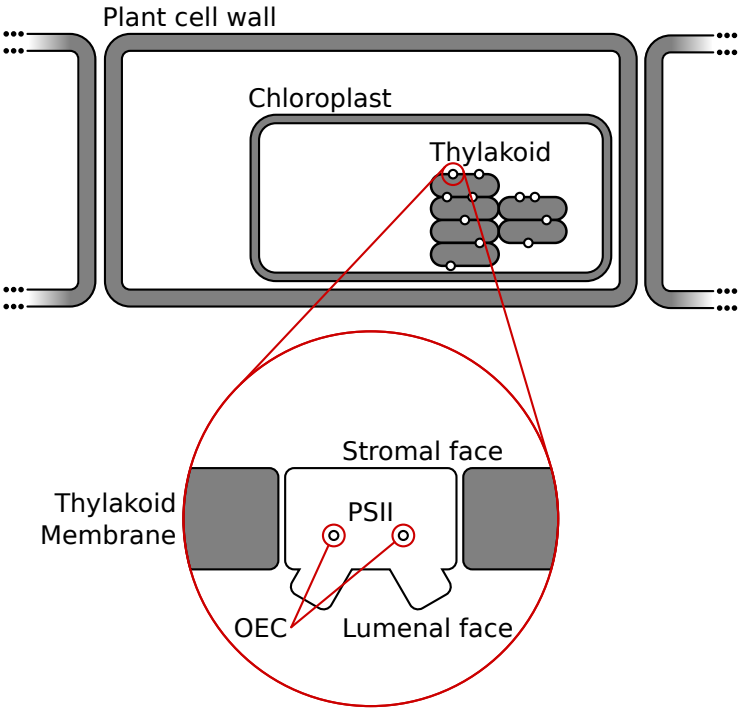


FIGURE 1.2: Schematic representation of the locale of PSII within a plant cell. The PSII dimer is a transmembrane protein that transfixes the thylakoid membrane. The thylakoid is located inside the chloroplast, which is in turn located within the plant cell.

to the phospholipid bilayer in which it is embedded. Differential hydrophobicity of the surface residues of PSII along this axis stabilise the transfixation of PSII to the thyalakoid membrane. Concretely, the holoenzyme possesses a transverse region exposing a large number of alpha helices perpendicular to the thylakoid surface, which interface with detergent cofactors and the membrane.[c]

In crystallographic structures of PSII, the visually apparent association of water molecules with the lumenal and stromal surfaces of the

[c]Blankenship, R.E. Ch. 6 offers a good survey of the gross morphology of PSII and its situation within the plant cell, although it was published without the benefit of high-resolution XRD structures that are now available. [18]

holoenzyme, and absence of peripheral waters in the transmembrane region permits facile identification of *water channels* in the structure. These water channels are networks of water molecules and associated cofactors which provide a route for diffusive transport of water molecules and protons to and from the active site of PSII, the Oxygen-Evolving Complex (OEC). [15]

1.2.2 The Oxygen Evolving Complex

The Oxygen-Evolving Complex (OEC), also known as the Water-Oxidising Complex (WOC), is a cofactor coordinated by the D1 and CP43 subunits of PSII. It has been identified as the site where storage of four oxidising equivalents and catalytic oxidation of water takes place, effecting the net reaction given in Equation 1.1.

$$2\,H_2O \longrightarrow 4\,e^- + O_2 + 4\,H^+ \tag{1.1}$$

The OEC is now known to be comprised of a Mn_4CaO_5 cluster where oxidising equivalents are stored by successive oxidation of Mn ions. The Mn ions of this cluster are coordinated by carboxylate moieties donated by residues of the D1 and CP43 subunits as well as a imidazolyl nitrogen provided by a histidine residue of D1.

The gross morphology of the OEC is characterised as a distorted heterometallic cubane comprised of three Mn ions ($Mn_{1,2,3}$) and a calcium ion bound by capping *oxo* ligands, and a further pendant Mn (Mn_4) connected to the cubane *via* Mn_3 by a further μ^2-*oxo* bridge as well as weak interactions with the O_5 atom, which is of contentious identity and connectivity. A recent crystallographic geometry of the PSII OEC obtained by Umena, *et al*. is illustrated in Figure 1.3. [15]

The proton concentration gradient generated across the thylakoid membrane by water oxidation powers membrane-transfixing ATPases, representing the first step in the so-called photosynthetic Z-Scheme.[d] [19]

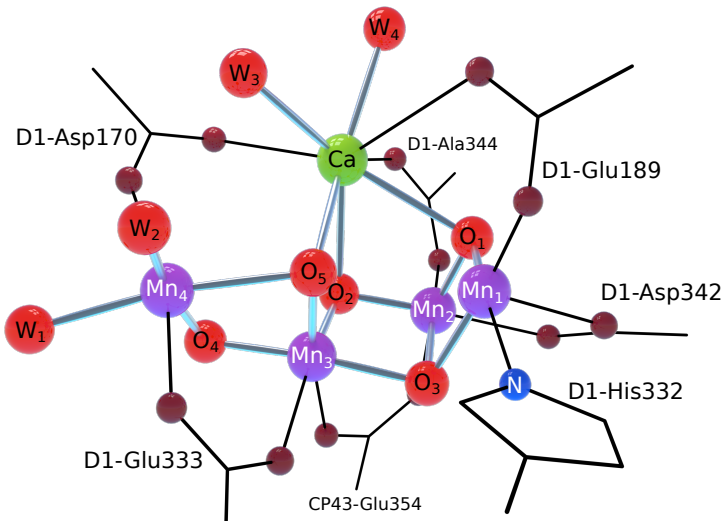


FIGURE 1.3: Illustration of the OEC and coordinating amino acids in monomer 'A' of PSII, adapted from the 190 pm-resolution XFEL crystallographic geometry of Umena, *et al*. Inorganic cofactor has been visually highlighted. (PDB accession code `4ub6/a`).

[d] PSII is so-named as it was the second photosystem to be discovered.

1.2.3 The Joliot–Kok Cycle

Studies originating in the 1960s involving short-duration illumination of dark-adapted chloroplasts reveal that PSII water photo-oxidation is a four-photon process which cycles the OEC through four oxidation state configurations, termed *S-states*. This cycle is known as the Kok or Kok–Joliot cycle, after Pierre Joliot, who experimentally resolved the cycle, and Bessel Kok, who offered a kinetic theory for the periodicity of the cycle. [20], [21] The cycle is illustrated in Figure 1.4. The existence of intermediate photoexcited states of the OEC that are built upon by further excitation may be inferred from the Emerson–Arnold experiments [22], [23] which evidenced tuneability in oxygen yield *via* dark period duration, and the observation of Allen and Franck that oxygen evolution is suppressed in dark-adapted PSII in single-flash scenarios, [24] a phenomenon which is now understood to be the result of regression of the OEC to the S_1 state from states S_{2-3}. The oxidation state distribution within the OEC corresponding to each *S*-state remains contentious, with two competing models—the low and high-oxidation state paradigms—existing, separated by two a two-electron redox. [25] Due to the inability to directly and unambiguously probe metal oxidation states with currently existing experimental tools, these paradigms are likely to stimulate heated debate for the forseeable future.

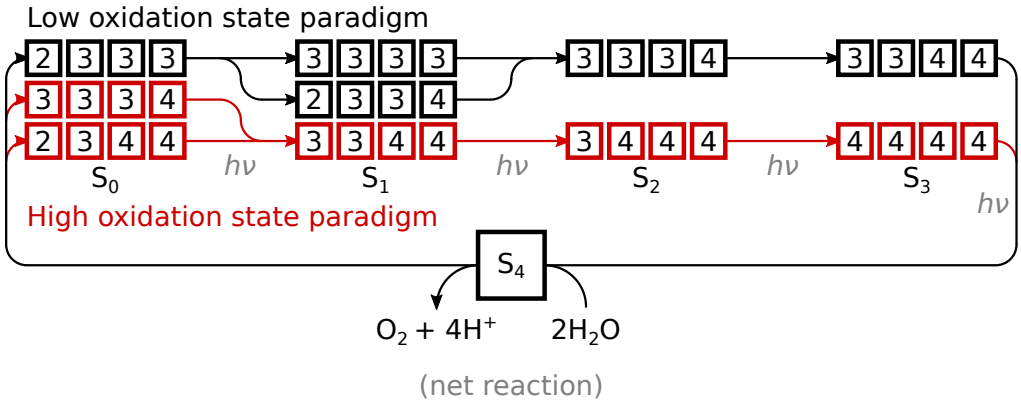


FIGURE 1.4: Simplified representation of the two predominant paradigms for oxidation state progression in PSII, based on the work of Krewald, *et al*. [25]

It was inferred that oxygen evolution occurs only in the transition $S_4 \rightarrow S_0$, though massed chloroplastic PSII rapidly desynchronises due to double-excitations, missed excitations and spontaneous state regression. [20] Nevertheless, the S_0 and S_1 intermediates are kinetically stable and the S_2 and S_3 states have lifetimes on the order of seconds to minutes. The S_4 state spontaneously oxidises water and reverts to the S_0 state on a sub-millisecond timescale. [26] completing the cycle. On the basis of free-energy analysis, Krishtalik proposed that the oxidation of water at the active site proceeds as a rate-determining two-electron redox to form peroxide, followed by two rapid oxidations effecting free dioxygen. [27] The difficulty in reliabily eliciting S-state transitions across entire populations of PSII OECs has motivated the application of Markov models in disentangling spectroscopic signatures arising from contamination of a predominating S-state by other S-states (*e.g.* [28], [29], also see [30]

for application of this insight in OEC Electron Paramagnetic Resonance (EPR).)

1.2.4 Characterisation of OEC structure

Characterisation of the PSII OEC by diffractometric methods has been an area of intense interest. Prior to a complete understanding of the general connectivity of the OEC, important diagnostic information on the relationship between metal ions in the OEC was obtained *via* Extended X-ray Absorption Fine Structure (EXAFS) and EPR, and these techniques remain valuable adjuncts in OEC structural and mechanistic characterisation. [31]

1.2.4.1 X-ray absorption methods

X-ray absorption techniques such as Extended X-ray Absorption Fine Structure (EXAFS) spectroscopy can furnish useful information regarding bioinorganic cofactors within proteins. Briefly, EXAFS measures perturbations at a core electron absorption edge due to photoelectron resonance with nearby atoms. EXAFS is capable of obtaining indirect geometric information regarding such cofactors, whilst placing few burdens on experimentalists in terms of sample preparation.[e] [32] As EXAFS only

[e] Although Cinco, *et al*. cite contamination of samples by exogenous ions of the same type as being probed by an EXAFS experiment can represent a confound to analysis (*vide infra*), whilst Yano, *et al*. cite the presence of endogenous Fe elsewhere in PSII as an interferent.

obtains the distances between atoms and nearby photoelectron-scattering nuclei, EXAFS data may be necessarily topologically ambiguous, however the technique can be an important adjunct to other analyses.

Prior to the production of high-resolution X-ray structures capable of identifying the connectivity of the OEC, Cinco, *et al*. established the presence of a Ca/Mn cluster within the OEC *via* Ca K-edge EXAFS. [31] Concretely, the data indicated a 340 pm Ca···Mn vector. At this point, it was known that Ca was an essential ion for photosynthetic metabolism, however it was not known how Ca related geometrically to Mn. Cinco, *et al*. also obtained Sr K-edge spectra consistent with Sr substitution for Ca, establishing the mode of poisoning of photosynthetic activity in Sr-substituted PSII prior to the availability of Sr-substituted XRD data (see Chapter 4). These data indicated the presence of oxo-bridges of some form between the Mn ions and the heterometal.

Yano, *et al*. obtained higher-resolution EXAFS measurements of the OEC in 2005 to supplement the incomplete XRD picture of that time. [33] At this point, loci of electron density corresponding to the OECs had been identified, however the resolution (Yano, *et al*. cite 320 pm to 380 pm) fell short of that required to unambiguously resolve the OEC connectivity, and potential X-ray photodamage further reduced confidence in the diagnostic quality of XRD results. [33]

1.2.4.2 X-ray photodamage of the OEC

A potential confound in X-ray studies of PSII is the possible photoreduction of the OEC by X-ray irradiation. This was identified by Suga, *et al*. as a plausible cause of dissimilarities in OEC internuclear distances between their femtosecond X-ray diffraction (§1.2.4.3) and EXAFS data. In this scenario, discrepancies are generated by photoreduction of Mn ions, generating putative *S*-states (see §1.2.3) below S_0. [34] Contrariwise, Yano, *et al*. indicated that EXAFS produces less photodamage than non-pulse XRD due to the lower specific energy of the EXAFS technique. [33]

1.2.4.3 Crystallographic characterisation

In 2011, Umena, *et al*. produced an X-ray crystal structure based on single-crystal synchrotron irradiation at a resolution of 190 pm, permitting unambiguous identification of the atoms within the OEC and furnishing interatomic distances with a low degree of measurement uncertainty. [15] This structure was supplemented with an improved 195 pm structure by means of X-ray Free Electron Laser (XFEL) irradiation in 2015. [34] This technique uses an ultra-short ($t \leq 10$ fs) X-ray pulse of extreme specific brilliance (delivering $1.4\,\mathrm{MGy}\cdot\mu\mathrm{m}^{-2}$ per pulse[f]).

Whilst possessing a lower absolute resolution factor compared to the

[f] On the basis of the experimental section of Suga, *et al*., XFEL pulses in this experiment had a specific power of approximately 4.2 GW.

precedent 190 pm structure, the 195 pm structure is considered superior to its predecessor structure on the grounds of the mitigation of the presence of photodamage in the structure (see §1.2.4.2). Whilst catastrophic (and sometimes non-local) damage to the crystal occured in this experiment, the extremely short duration of irradiation produced a non-time-averaged *'snapshot'* of the crystal structure. This structure therefore lacks structural relaxation that would occur due to the photoreduction of Mn ions. This principle is termed *'diffraction before destruction'*, and has been explored theoretically by Jurek, *et al*., wherein XFEL irradiation of molecular clusters is seen to produce almost complete ionisation, eliciting isotropic Coulombic explosion over a timeframe of tens of femtoseconds. [35] In practice, Suga, *et al*. found that XFEL irradiation of prepared PSII crystals produced non-local macroscopic damage in some instances, necessitating careful selection of sufficiently spatially separated irradiation sites[g] on multiple crystals. [34] Nevertheless, XFEL diffractometry should be considered the gold-standard of crystallographic characterisation of PSII.

1.2.4.4 Electron Paramagnetic Resonance

Electron Paramagnetic Resonance (EPR)/Electron Spin Resonance (ESR) has proven to be a powerful tool in characterising the OEC, as it permits

[g] 7650 irradiations of 336 crystals in total.

the inference of information regarding Mn oxidation state distributions inside the OEC. In cryogenic EPR analysis of the S_2 state, the OEC possesses a complex Mn nuclear hyperfine structured feature (commonly referred to as the multiline (ML) signal) centered at $g \approx 1.98$, as well as a broad feature at $g \approx 4.1$. [36] The presence of the multiline (ML) feature was used to argue for a probable $Mn^{III}Mn^{IV}$ or $Mn^{II}Mn^{III}$ dimer in the OEC before the topology of the OEC was established by XRD. [37] The presence of 4 Mn ions in the OEC had been previously established by stoichiometric studies ([38] and references therein). The observed ML signal is consistent with both the high and low oxidation state paradigms later established, which both contain potential $Mn^{III}Mn^{IV}$ dimers (Figure 1.4).

Nuclear hyperfine anisotropy resolved within the OEC EPR signals has important diagnostic value with respect to the coordination environment of Mn centres in the OEC, as the orientations and aspect ratios of the hyperfine tensors are correlated with oxidation state and coordination environment. As the anisotropic hyperfine response underdetermines molecular geometry and electronic structure, a large number of credible models accounting for the EPR response can be proposed. Nevertheless, the computational simulation of EPR spectra of candidate models can be effectively used to test hypotheses regarding the oxidation states, magnetic couplings and topology of the OEC (Chapter 6).

1.2.5 Mn_4CaO_5 cofactor topology

Broadly, current understanding of the OEC oxo-metal cofactor indicates the presence of an oxo-manganese heterocubane, with one metal vertex substituted by a calcium ion and labelled manganese ions Mn_1, Mn_2 and Mn_3 constituting the remaining three vertices. Labelled oxygen atoms O_1, O_2, O_3 and O_5 comprise the μ^3-*oxo* vertices of this cubane. The fourth manganese—labelled Mn_4—is connected to Mn_3 *via* putative μ^2-*oxo* ligand O_4 and *via* O_5, which is of contentious bridging character, but which is reasonably interpreted as a di-bridged ion pendant to the heterocubane. The geometry of the cluster is illustrated schematically in Figure 1.5, absent coordinating histidyl or carboxylate ligands and interstitial waters. Additionally, four putative *μ-aqua* ligands labelled W_1, W_2, W_3 and W_4 coordinate the complex, with two coordinating Mn_4 and completing the octahedral coordination environment of this ion, and an additional closely associated with the Ca ion. Crystallographic geometries of the PSII OEC are presented in Figure 1.6, based on the 1.95 Å XFEL structure of Suga, *et al*. [34]

1.2.6 Mn_4CaO_5 cofactor ligation

The Mn_4CaO_5 cofactor is ligated by a number of carboxylate and imidozyl moieties donated by histidine, aspartate, glutamine and alanine residues.

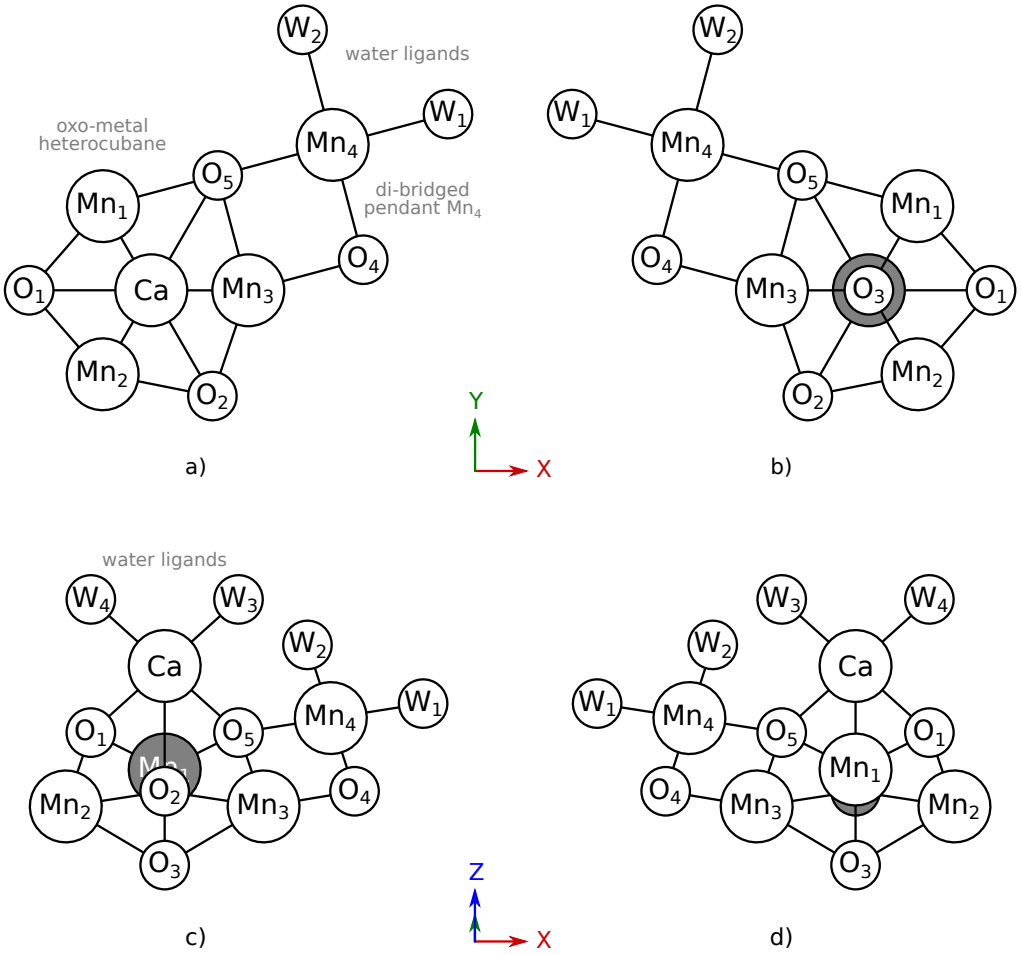


Figure 1.5: Illustration schematically representing the OEC Mn_4CaO_5 cofactor, and establishing the conventions used to define orientations of the cluster. **a**) and **b**) Projections collinear with Ca/O_3 axis with W_3 and W_4 removed, from *above* and *below*, respectively. **c**) and **d**) Projections approximately perpendicular to Ca/O_3 axis, from the *back* and *front*, respectively. The Ca/O_3 axis has been designated the Z-axis on account of it being approximately collinear with the normal plane of the thylakoid membrane *in vivo*.

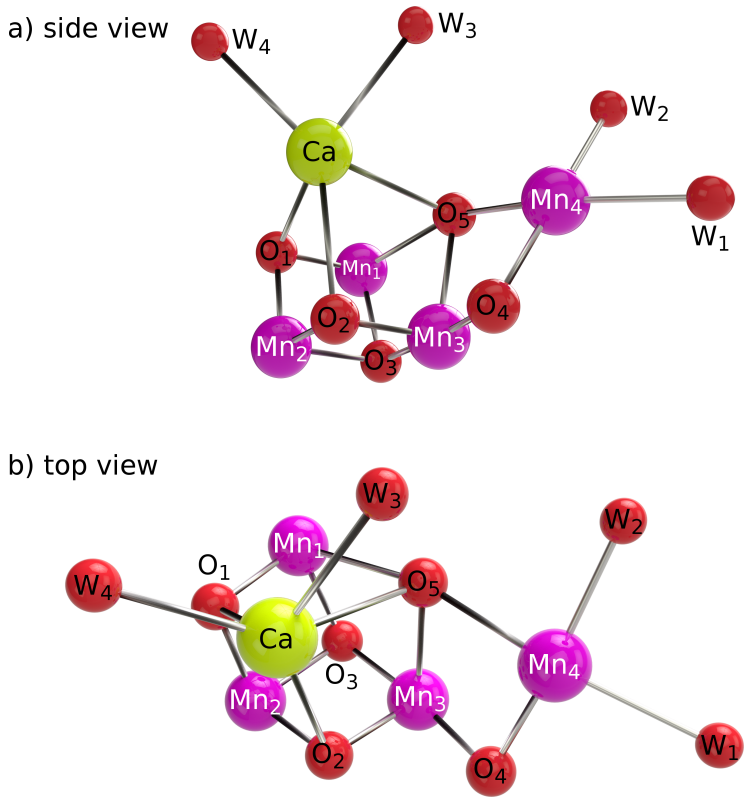


FIGURE 1.6: Crystallographic geometries of the PSII OEC, based on the X-ray structure provided by Suga, *et al.* [34]

This coordination is illustrated schematically in Figure 1.7.

1.2.7 Mn_4CaO_5 spin network

Each OEC cofactor contains 4 manganese nuclei in Kramers–Anderson superexchange contact *via* multiple μ^n-*oxo* bridges, admitting a nontrivial Heisenberg Hamiltonian and rich magnetic structure that evolves over the course of the *S*-state cycle and is of important diagnostic value. Briefly, Kramers–Anderson superexchange is the magnetic interaction of ions that are bridged by a non-magnetic but spin-polarisable ligand.

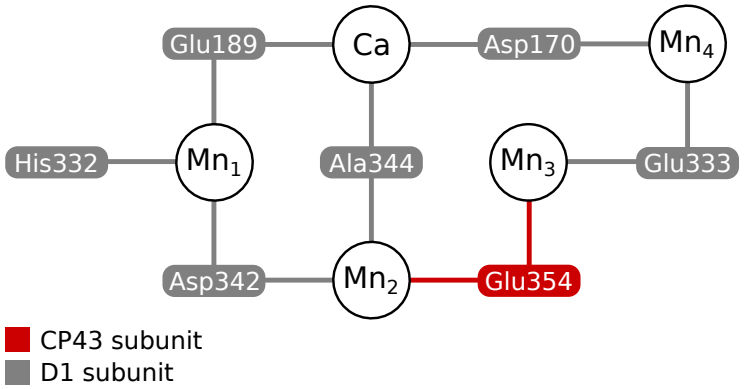


Figure 1.7: Schematic representation of amino acids of D1 and CP43 polypeptides ligating the PSII OEC. All ligation occurs through sidechains except for D1-Ala344, which is the C-terminus of the D1 subunit and ligates *via* the α-carboxylate of the amino acid.

[39] The character and degree of the superexchange-mediated spin-spin interaction is in part dependent on the precise geometry of the bridging ligand with respect to the metal ions, however it is typically several orders of magnitude larger than that which would be predicted for unbridged cations. [40]

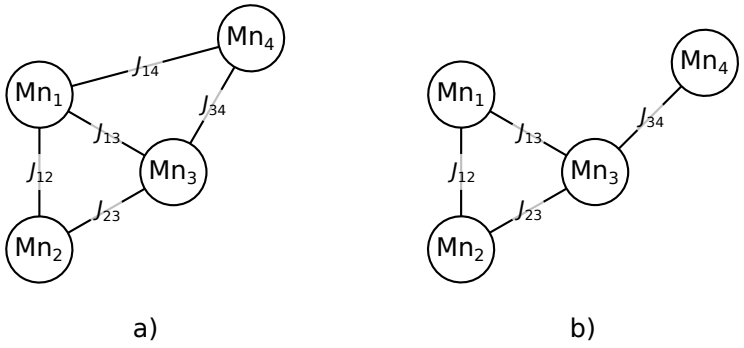


Figure 1.8: Representation of the magnetic coupling graph of the PSII OEC, subsuming multiple potential superexchange pathways into J_{ij} couplings.

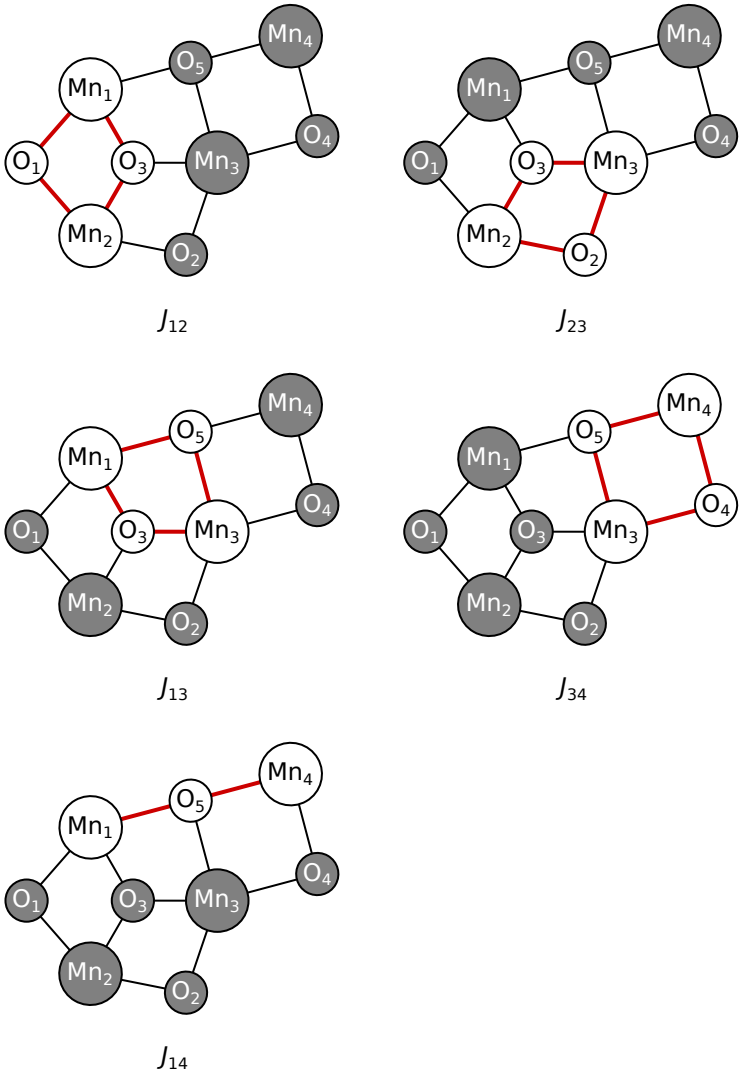


FIGURE 1.9: Schematic representation of contributory Kramers–Anderson superexchange pathways to spin-spin interactions in the OEC Mn_4CaO_5 cofactor.

1.2.8 Morphological differences between extant OEC structures

Whilst being broadly consistent in gross structure with the 190 pm structure, the 195 pm PSII structure of Suga, *et al*. possesses a distinct coordination geometry of O_5 with respect to the earlier structure. [34] Suga, *et al*. note that the O_5 position is somewhat biased towards Mn_3 and Mn_4 compared to the 190 pm structure, wherein O_5 is more evenly associated with Mn_1, Mn_3, Mn_4 and Ca. The authors interpret this as potentially indicating preferential photoreduction of Mn_4 in the 190 pm structure.[h] The authors suggest that this provides some data on the relative susceptibility of the Mn ions to reduction (specifically, Mn_1 with respect to Mn_4). They indicate that this may be due to the relative solvent accessibility of Mn_4 (which has two binding sites for *aqua* ligands) with respect to Mn_1, which is coordinatively saturated with *oxo*-caps or peptide termini.

Suga, *et al*. point out commonality with computational and theoretical results that produce preferential association of O_5 with either Mn_1 or Mn_4. [34]

[h] This would suggest that while photodamage may be *stochastic* (albeit weighted by the X-ray cross-sections of individual atoms), the oxidative flexibility of the Mn ions act as a sort of attractor that could reasonably be expected to produce coherent photodamage across independent OECs over the timescale of synchrotron irradiation.

1.3 Artificial Photosynthesis

Artificial photosynthesis is the replication of key aspects of plant photosynthesis in artificial systems. This definition usually implies the harvesting of solar photons to generate a redox potential catalysing a chemical reaction germane to plant physiology, such as the oxidation of water the reduction of protons, or the fixation of carbon dioxide associated with plant respiration.

Artificial photosynthesis systems (*artificial photosystems*) are typically most appealing if they are chemically simple, robust to photodamage and other forms of degradation (potentially by recourse to self–repairing and self-assembling behaviour) and are manufacturable with cheap, earth-abundant materials. Earth-abundance refers to the abundance and facile provenance of a material at or near the Earth's surface. These desirable properties of an artificial photosystem are most succinctly embodied in the metaphor of the *'artificial leaf'*, a term popularised by Nocera, *et al*.. [41]

A candidate biomimetic catalyst will be synthetically straightforward and robust with respect to a wide variety of environmental conditions without requiring the delicate homœostasis of the intracellular environment that preserves the activity of PSII, nor the extensive self-repair

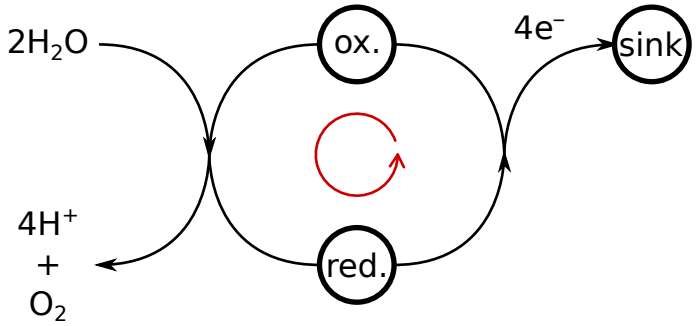


Figure 1.10: A generic water splitting catalytic cycle. **red.** and **ox.** represent reduced and oxidised forms of the catalytic species, related by four-electron oxidation by an electron **sink** and four-electron reduction by the oxidation of water. In the case of biological photosynthetic water oxidation, the oxidation of the catalyst occurs in a step-wise rather than concerted process.

framework required to recycle photodamaged subunits; neutron scattering studies [42] provide evidence of systematic conformational flexibility in the PSII framework, indicative of maintenance activity beyond simple immobilisation of the OEC. This would appear to recapitulate the rapid photodamage, ejection, and *de novo* biosynthesis and reinsertion cycle of the D1 subunit of PSII. [43]

Integral to the definition of a catalyst is the self–regeneration of the catalytic species. Ideally (neglecting poisoning reactions) catalytic yield does not rely on the relative stoichiometry of the catalytic species and reactants, even if a sacrificial reducing agent, photoelectrode or external electrical bias is required to regenerate the catalyst (Figure 1.10).

1.4 Synthetic Water Oxidation Catalysts with OEC-like Motifs

This section will introduce and briefly discuss two exemplary synthetic water oxidation catalysts that possess noted structural similarity with the PSII OEC. In each of the cases discussed, the irreducible water oxidising component appears to be the presence of a redox-active polynuclear oxo-bridged motif.

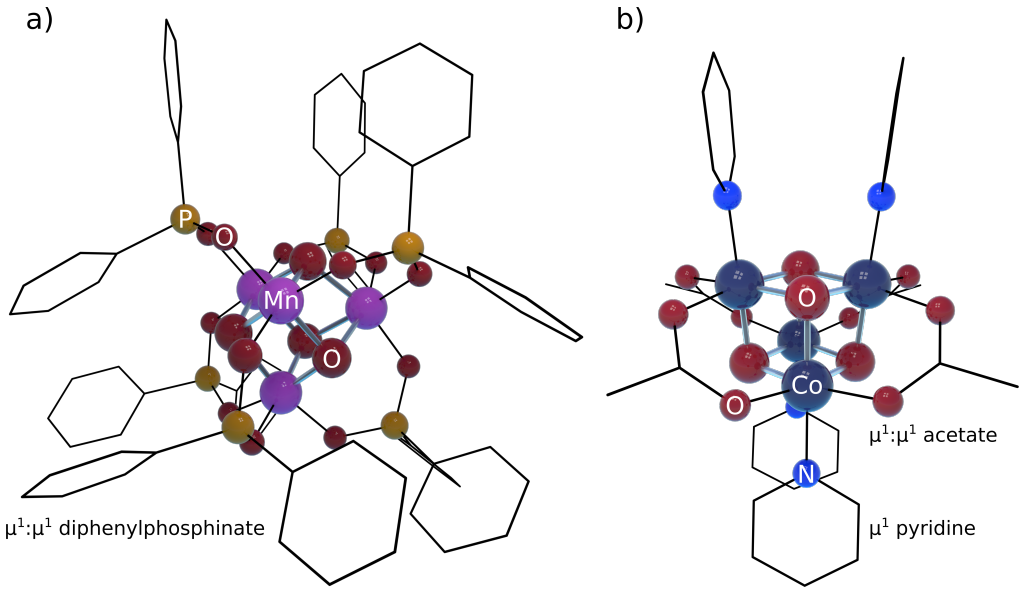


FIGURE 1.11: Representation of Mn and Co oxo-metal cubanes, produced from crystallographic coordinates published by Ruettinger, *et al.* and Chakrabarty, *et al.* [44], [45] The embedded cubane motif has been visually emphasised. **a**) A prototypical manganese oxo-metal cubane of the form $Mn_4O_4L_6$. L: $\eta^1{:}\eta^1$ bridging ligands, *e.g.* Ph_2POO^-. **b**) A prototypical cobalt oxo-metal cubane of the form $Co_4O_4L_4L'_4$. L: $\eta^1{:}\eta^1$ bridging ligands, *e.g.* CH_3COO^-, L′: monodentate ligands, *e.g.* NC_5H_5.

1.4.1 The oxo-manganese heterocubane of Dismukes and coworkers

Henceforth referred to as manganese cubium complexes, oxo-manganese complexes of the form $Mn_4O_4(R_2PO_2)_6$ bear an oxo-manganese heterocubane motif (*i.e.* a cubane with alternating Mn and μ^3-*oxo* caps.[i] The coordination sphere of each Mn ion is completed by $\eta^1 : \eta^1$-*bis(aryl)-phosphinato* ligation. This complex is known crystallographically in both neutral and cationic forms, and a number of aryl ring functionalisations have been described. [46]

The bis(*p*-methoxyphenyl)phosphinate-ligated form of this complex is known to photocatalytically oxidise water when immobilised in a porous nanostructured Nafion[j] substrate, at potentials as low as 0.8 V. [46] Whilst strong analogy has been drawn between the oxo-cubane morphology of this species and the OEC of PSII, it is contentious as to whether the complex as crystallographically characterised is the catalytically active form. To wit, Hocking, *et al*. have advanced the hypothesis—which they support with X-ray scattering evidence—that the complex decomposes into a Birnessite-like phase *in situ*.

Hocking, *et al*. note that the electro-oxidation products of both the $[Mn_4O_4L_6]^+$ cubane and the di-μ_2-oxo bridged dinuclear OEC model

[i] This motif may be considered as interpenetrating Mn_4 and O_4 tetrahedra.
[j] A cation exchange polymer.

$[(bipy)_2MnO_2Mn(bipy)_2]^{3+}$ generate the same EXAFS spectrum, and that furthermore this spectra is nearly superimposable with the EXAFS spectra of birnessite, todorokite, and to a lesser extent other manganese oxides, but is not a good fit with the EXAFS spectrum of the unmodified cubane. They draw parallels between the putative *in situ* catalyst and the amorphous cobalt–phosphate (Co–Pi) water oxidation catalyst of Lutterman, *et al*. [47], [48]

Hocking, *et al*. also provide experimental evidence that the apparent water oxidising activity of $[Mn^{III/IV}(\mu\text{–}O)_2(terpy)_2(H_2O)_2]^{3+}$ is actually a consequence of a similar decomposition *in situ*, which has important implications with respect to ostensible correlations between heterogeneous structural motifs and catalytic water oxidation. [49]

Birnessite is a manganese oxide mineraloid comprised of oxo-manganese laminae with ideal stoichiometry MnO_2, however in practice naturally occuring birnessite contains nontrivial concentrations of water, non-manganese counterions and other interferents. Birnessite rapidly accumulates crystallographic defects in the form of oxygen vacancies and proton-compensated manganese vacancies (also known as Rütschi defects), with various permissible scenarios for functionalisation of these defected sites. [50] The laminar structure of birnessite also offers a wide variety of edge characteristics that may be of relevance to catalytic activity.

1.4.2 Oxo-Cobalt Cubanes

Functionalised oxo-cobalt heterocubanes comprised of a Co_4O_4 oxo-metal cubane core coordinated by four bridging carboxylates and four monodentate O or N donors (*e.g.* Figure 1.11b) were identified by the group of Das and coworkers as catalytic oxidants of a wide range of organic species, such α-pinene, *p*-xylene, aromatic and aliphatic alcohols, and ethylbenzene. [45], [51]–[55]

Complexes of this form consist of a heterocubane core with alternating Co and O vertices. Four faces of the cubane are bridged by carboxylate ligands coordinating pairs of cobalt ions in an $\eta^1 : \eta^1$ fashion *via* the acetate oxygens, whilst the remaining two faces (on opposing sides of the cubane) are capped with μ^1-coordinating ligands such as pyridine or methoxide.

Thereafter, McCool, *et al.* identified the complex $Co_4O_4(CH_3COO)_4(py)_4$ as having putative catalytic activity as a water oxidant. Cyclic voltammetry suggests that the catalytic species in the water oxidation cycle exists as the cationic form. [56] Catalytic water splitting activity has been observed in this complex using a $[Ru(bipy)_3]^{3+/2+}$ couple as a photosensitiser and persulfate as a sacrificial oxidant. A similar cycle has been identified for a polyoxotungstate with an adamantanoid oxo-ruthenium core complex by Natali, *et al.* [57] Berardi, *et al.* have identified strongly divergent

quantum efficiencies for the $[Ru(bipy)_3]^{3+/2+}/[S_2O_8]^{2-}$ photocatalytic cycle wherein the *para*- substituent of the bipyridine ligands is replaced with other moieties such as OCH_3, CH_3, ^tBu, Br, $COOCH_3$, and CN. [58]

It is presumed that the neutral complex has an implicit electron configuration of $4 \times d^6$ arising from four Co^{III} ions, which in the low-spin octahedral splitting regime gives rise to an obligate $S = 0$ electronic configuration. One-electron oxidation of the molecule presumably results in a $d^5, 3 \times d^6$ configuration, whilst the 4-electron reduction of the cubane implicitly necessary to dissociate electroneutral dioxygen from the molecule suggests a final d electron configuration of $d^6, 3 \times d^7$ and a $Co^{II}{}_3Co^{III}$ valence configuration as the most parsimonious solution.

McAlpin, *et al*. have performed EPR and Electron-Nuclear Double Resonance (ENDOR) simulation of the cationic form of this cubane, indicating that the unpaired spin density arising from the cationic structure is completely delocalised across the metal centres and μ^3-*oxo* caps of the molecule. Experimental ENDOR characterisation resolves a $S = \frac{1}{2}$ ground state, with a high degree of consistency between experimental and theoretical EPR and ENDOR results. [59] The result of complete delocalisation of the unpaired spin in the cationic complex obtained by McAlpin, *et al*. are consistent with unpublished exploratory results we independently obtained on the complex at a BP86-D3/TZP-ZORA level of theory in Amsterdam Density Functional (ADF). This delocalisation

would appear to account for the lack of crystallographic distortion arising from differential oxidation.

This complex bears a great deal of structural similarity with the aforementioned oxo-manganese complex (§1.4.1), being comprised of an oxo-cobalt cubane with alternating metal and μ^3-*oxo* vertices. McCool, *et al*. draw explicit analogy between cubane motifs of the form M_4O_4 where M is a 3*d* metal, and the Mn_4CaO_5 cofactor of the OEC.

1.5 Integrated Artificial Photosystems

Integration of a water oxidising catalyst into a functional device or process is a prerequisite for viability as an industrial source of hydrogen. Brimblecombe, *et al*. have recently realised an integrated package that generates hydrogen gas from water without the application of external bias, inspired in part by the dye-sensitised solar cells (DSSCs) of Grätzel and others. [60] The prior art of Grätzel, *et al*. in this domain must be understood to derive a complete understanding of integrated artifical photosystems and their likely future development trajectories.

1.5.1 Dye-sensitised Solar Cells (DSSCs) / (Grätzel) cells

Commonly referred to as Grätzel cells, DSSCs are a subset of organophotovoltaic (OPV) cells. Such cells use photoexcitation of dye molecules to promote electrons into the conduction band of supported TiO_2 (*titania*) nanoparticles. [61] These electrons diffuse into a transparent anode that supports an electrical load. Electrons returning from the load are supplied to a cathode at the rear of the cell. Electron recombination with the oxidised dye molecules is prevented by competitive reduction *via* a net reaction $3I^- \longrightarrow I_3^- + 2e^-$, which is a relatively rapid process. [62] The process is schematically illustrated in Figure 1.12. A number of alternative redox couples including $Co^{III}(dbbip)_2 \rightleftharpoons Co^{II}(dbbip)_2$ [63] and $TEMPO^+ \rightleftharpoons$ TEMPO [64] have been developed, whilst the development of viable ionic liquid [65] or solid state hole transport [66], [67] substrates represents an important goal in making DSSC cells more physically robust.

The microstructural properties of the TiO_2 substrate are of great importance in the enhancement of DSSC performance, with a large number of structural parameters offering a number of dimensions of tuneability. Enhancement of the surface area of mesoporous TiO_2 increases the number of available sites for dye binding, whilst advances in the preparation

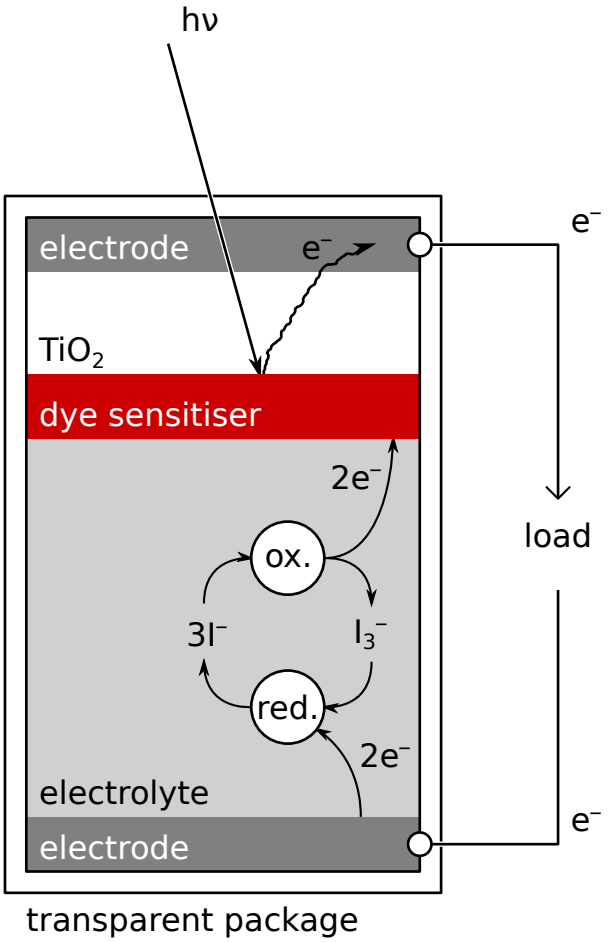


Figure 1.12: Schematic representation of a dye-sensitised solar cell. In a conventional DSSC, dye molecules are adsorbed to the surfaces of individual TiO_2 nanoparticles, contrarily represented here by discrete layers.

of heterogeneous networks of semiconductor phases are anticipated to give rise to new DSSC modalities or dramatically increased solar-electrical conversion efficiency. [68]

Such cells bear arguable analogy to photosynthetic light harvesting insomuch as they mobilise electrons *via* photoexcitation of a chromophore, where conventional doped semiconductor solar cells operate on the basis

of direct photoexcitation of semiconductor valence electrons. Electron transport in the plant photosystems is achieved through multiple Chl and carotene units in Van der Waals contact (for which there is evidence of excitonic resonance energy transfer), [69] whereas in dye-sensitised solar cells the dye molecules are directly adsorbed to TiO_2 nanoparticles.

The use of a dye sensitiser carries two principal benefits over traditional doped semiconductor solar cells: there is no cut-off illuminance behaviour due to slow recombination dynamics in comparison to the $3I^-/I_3^-$ couple (as such, the DSSC cells continue to operate under low light conditions) and the dye absorption may be tuned to cover a much larger effective spectral range than the $TiO_{2(s)}$ semiconductor bandgap (typically 3.2 eV, [68] corresponding to a photon wavelength of ≈ 388 nm) would permit, [70] with research continuing into near-infrared sensitivity.

Additionally, recent research in dye design has led to the development of a number of all-organic dye molecules, [71] eschewing the use of Ru^{II} based sensitisers, with clear implications for cost and environmental friendliness. All-organic sensitiser species follow the $\mathbf{D} \rightarrow \pi \rightarrow \mathbf{A}$ structural principle. $n \rightarrow \pi^*$ electron promotion from the photoexcited electron-rich donor/antenna moiety **D** allows charge transfer through the conjugated π system to acceptor group **A**, which is generally covalently bound to the semiconductor substrate. All-organic sensitiser-based solutions are slightly less efficient than ruthenium-based dyes, however they

have the potential to be significantly more cost efficient per Watt.

The use of a dye-sensitised microcrystalline phase as opposed to a continuous amorphous or crystalline semiconductor phase facilitates the construction of flexible DSSCs, as well as continuous fabrication techniques. A criterion that may inform the desirability of a light harvesting technology is the ability to manufacture the technology in a continuous fashion, for instance by continuous extrusion or printing onto a substrate. This manufacture methodology was pursued by the now defunct Konarka technologies, who were able to achieve 'roll-to-roll' manufacture of flexible organic photovoltaic (OPV) solar cells by repurposing existing facilities. [72] The low specific power achieved by these devices is offset by their very low dollar and energy cost of manufacture.

1.5.2 DSSC water photolysis

Water scission may be afforded by replacing the $I_3^- \rightleftharpoons 3I^-$ couple of a dye-sensitised solar cell with oxidation of water on a suitable dye-adsorbed substrate affording $2e^-$ per H_2O molecule. In the system [73] of Mallouk, *et al.*, the Ru^{II}-based sensitiser/linker units connected to TiO_2 nanoparticles also covalently bind $IrO_2 \cdot nH_2O$ nanoparticles that act as water oxidation substrates, with concomitant proton reduction at a platinum cathode.

Whilst demonstrably effective, this method of scission is hampered by very low quantum yield ($\approx 1\%$) and rapid degradation of the Ru^{III}-bearing oxidised dye in aqueous media (< 20 turnovers). [73] Quantum yield is chiefly harmed by connective topology of the two types of nanoparticle and linking dye molecule: a dye molecule must bridge both a TiO_2 nanoparticle and an $IrO_2 \cdot nH_2O$ nanoparticle for its excitation to actually contribute to the photocurrent, whereas a dye molecule is 90–95% more likely to only be bound to $IrO_2 \cdot nH_2O$. [73]

1.5.3 Tandem water-splitting cells

Directly inspired by the combined functionality of Photosystems I and II, coupled DSSCs have been augmented to carry out photolytic water scission. In this system, water oxidation occurs at a photoexcited WO_3 substrate, wherein valence band holes oxidise water in a process essentially analogous to the action of PSII. [70], [74] Conduction band electrons from the WO_3 substrate replenish the valence band of a dye-sensitised $TiO_{2(s)}$ substrate, replacing the $I_3^- \rightleftharpoons 3I^-$ couple. Finally, conduction band electrons from TiO_2 reduce H^+ to H_2, catalysing the net reaction $2H_2O \longrightarrow O_2 + 2H_2$. Grätzel, *et al.*, have drawn strong analogy between the coupled WO_3 water oxidation and TiO_2 proton reduction systems and the 'Z-scheme' model of photosystem electrochemistry.

Similarly, Brimblecombe, *et al*. have recently reported the construction of a dye-sensitised photoelectrochemical cell that carries out both water oxidation and proton reduction without external bias (*i.e*. no voltage need be applied). [60] In effect, this cell design combines the dye-sensitised TiO_2 concept of Grätzel and coworkers with the putative photoactive Nafion–immobilised oxo-manganese cubane.

The Wider Context 2

The idea of the future being different from the present is so repugnant to our conventional modes of thought and behavior that we, most of us, offer a great resistance to acting on it in practice.

John Maynard Keynes

2.1 Photosynthesis in the Human World

2.1.1 Ecological Relevance

It is now uncontentious that global mean temperatures have increased since the industrial revolution, concomitant with the widespread combustion of fossil fuels and human activities and land use patterns associated with industrialisation. There now also exists a strong evidentiary and theoretical case for a causal relationship between human greenhouse gas emissions and this increase in temperature. [75] Anthropogenic perturbations to the global mean temperature are governed by and govern a richly interconnected set of factors including but not limited to greenhouse gas emissions (§2.1.1.1), planetary albedo, hydrological and cryospheric cycling, atmospheric aerosols and cloud cover modulation, and oceanic pH and dissoved gases. [76]–[82]

Factors which drive global mean temperature are known a *'radiative forcings'* and may be both positive or negative in direction. [83] The second-order effects of this mean temperature increase on the planetary climate are termed *climate change*. The effects of climate change may include rapid (in geological terms) changes in the spatial extent of climate regions, with attendant negative effects on land and ocean ecosystem stability. [84], [85] Climate change may also induce sea level rise due to the melting of glaciers, sea ice and ice sheets, and changes in extremal

weather event patterns. [86]–[89]

The increase in global temperatures over the past hundred years is amply illustrated by the Australian and global mean temperature anomalies, depicted in Figures 2.1 and 2.2, respectively. These figures depict the difference (*'anomaly'*) in mean annual temperature with respect to a reference mean taken over a particular 30-year period, illustrating a clear (albeit non-monotonic) increase in the annual mean temperature anomaly with respect to their corresponding reference windows. Notably, no part of Australia has exhibited a decrease in mean temperature with respect to a 1910 reference. [90]

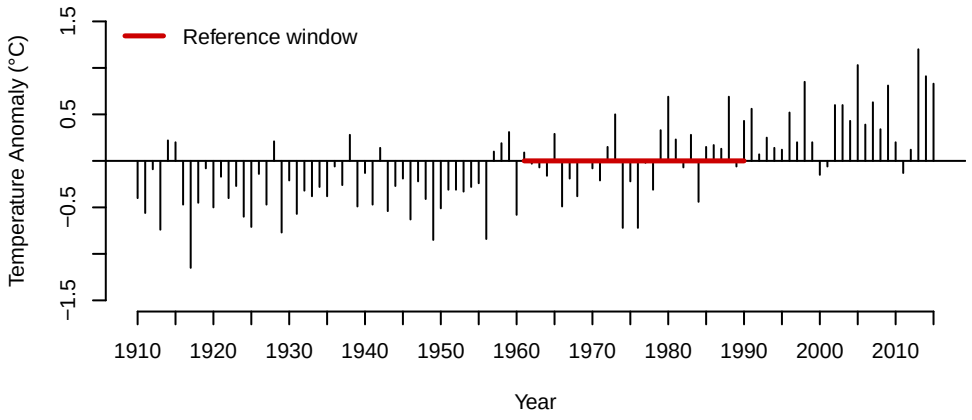


Figure 2.1: Trends in Australian annual mean temperature anomaly between 1910 and 2015, inclusive, with respect to reference interval between 1961 and 1990. Plot produced with data provided by the Australian Government Bureau of Meteorology. [91]

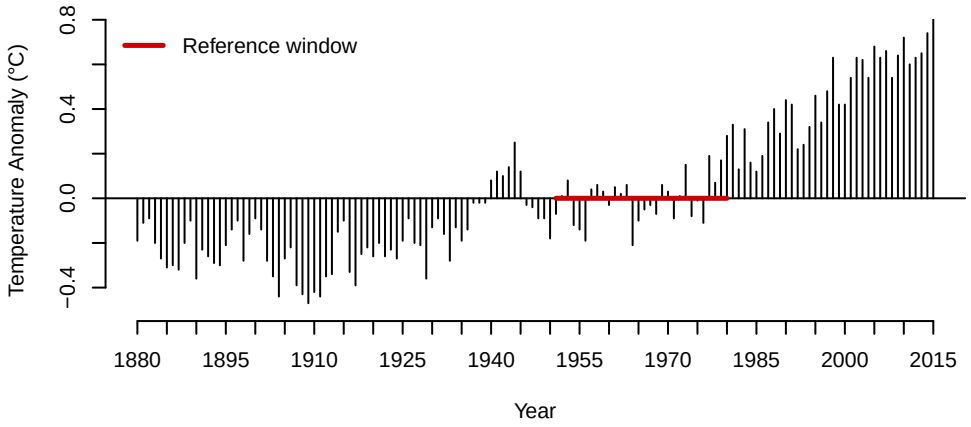


Figure 2.2: Trends in global annual mean land/ocean temperature anomaly between 1910 and 2015, inclusive, with respect to reference interval between 1951 and 1980. Plot produced with data provided by the United States National Aeronautic and Space Administration Goddard Institute for Space Studies. [92], [93]

2.1.1.1 The Greenhouse Effect

The *greenhouse effect* refers to the tendency for infrared-active molecules (*i.e.* molecules possessing a fixed or inducible electric dipole moment[a]) to absorb and scatter reflected solar radiation. [94] This effect traps some portion of solar radiation that would otherwise be reflected directly back into space, creating a relative increase in temperature (Figure 2.3). Molecules that produce positive radiative forcings due to their participation in a greenhouse effect are known as *'greenhouse gases'*. While the natural greenhouse effect is responsible for Earth's ability to support liquid water, it is the case that a runaway greenhouse effect owing to a majority CO_2

[a]This includes species with permanent dipole moments such as H_2O as well as species that acquire a dipole moment due to vibrational motion such as CO_2, but does not include homodiatomic species such as N_2, nor monatomic species such as Ar.

atmosphere is also responsible for the intense surface temperature of the planet Venus. [95]

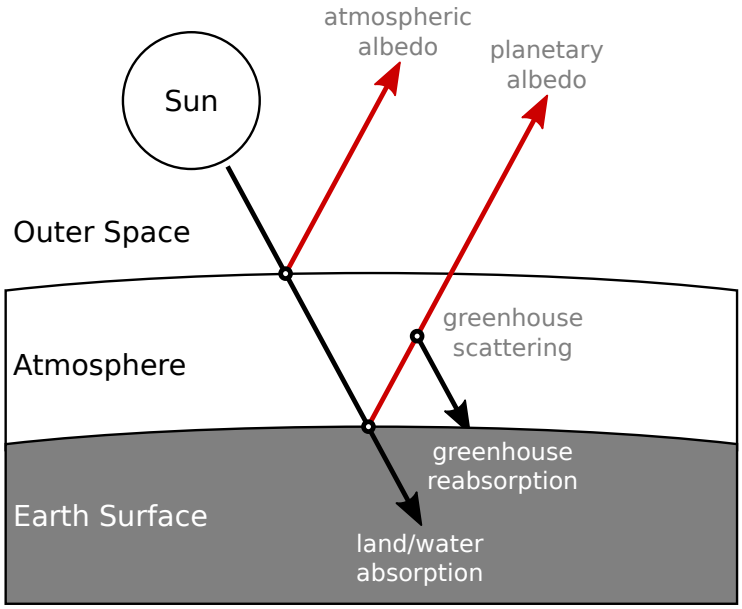


FIGURE 2.3: Schematic representation of the greenhouse effect on radiation reflection, absorption, re-emission and scattering. Planetary and atmospheric albedo is strongly dependent on the specific optical properties of regolith, plant life, cloud cover, atmospheric particulates, snow, and other factors. Re-emission of solar radiation can be in the form of prompt reflection as well as thermal re-emission.

Whilst the Earth's carbon cycle involves large fluxes of carbon dioxide, it is not until recently in Earth's biogeochemical history that a large source of excess CO_2 has appeared in the form of a combustion byproduct of fossil fuels, as well as other anthropogenic 'process emissions' such as cement manufacture.[b] [97] These excesses contribute a large positive

[b] Although see [96] with respect to unintentional sequestration of CO_2 by cement carbonation.

radiative forcing that accounts in climate models for much of the aforementioned observed mean temperature increase. Global atmospheric CO_2 concentrations, measured historically through ice core data and contemporarily through direct atmospheric sampling, have increased sharply in the wake of human industrialisation (Figure 2.4). The Intergovernmental Panel on Climate Change (IPCC) nominates the dramatic increase in atmospheric concentration of CO_2 since the industrial revolution as the largest contributor to radiative forcing. [98]

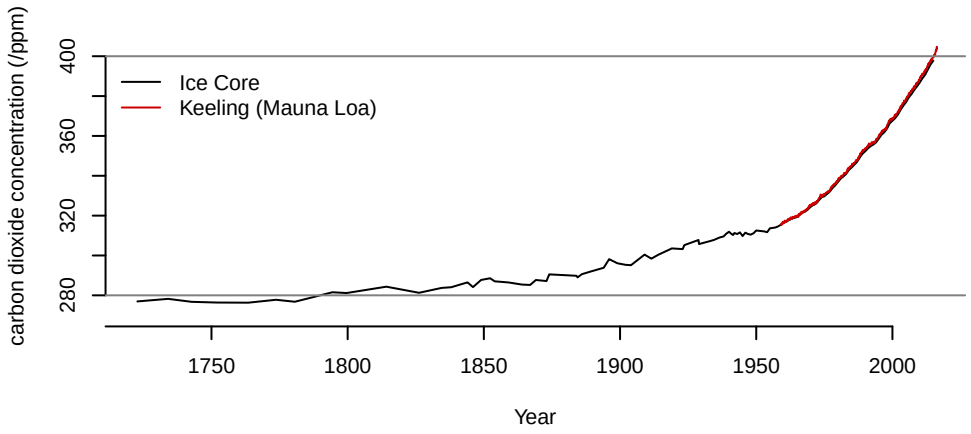


Figure 2.4: Trends in historic atmospheric CO_2 concentration between approximately 13 CE to 2015 inclusive, derived from ice core and atmospheric sampling. [99], [100] The 280 ppm pre-industrial CO_2 level and 400 ppm milestone are indicated in grey. Deseasonalised Keeling curve of *in situ* atmospheric CO_2 data between 1959 and 2015 is overlaid in red.

2.1.2 Economic and Geopolitical Relevance

2.1.2.1 Climate and Complex Societies

The United States Department of Defense has identified climate change-related effects (*'[...] recurring conditions of flooding, drought, and higher temperatures [...]'*, *'[m]ore frequent and/or more severe extreme weather events [...]'*, *'[s]ea level rise [...]'* and *'[d]ecreases in Arctic ice cover[...]'*) as potential hazards to international security due to the second-order effects of these phenomena on the stability of complex societies and economies. [101] This judgment comports with the opinion of the United States National Intelligence Council, which identifies the confluence of food, water, and energy security concerns in the context of climate change as a *'megatrend'* with *'[...] far-reaching effects on global development [...]'*, and notes the particular sensitivity of developing economies to climate change-related stressors. [102] The World Health Organisation has predicted significant excess mortality and morbidity due to a variety of public health issues exacerbated by climate change between 2030–2050—particularly in the developing world—even when public health improvement projections are considered. [103] It is anticipated that so-called *'urban micro-climates'*/*'heat islands'* are likely to exacerbate the local effects of climate change in many cities. [104]

2.1.2.2 Fossil Fuels and Geopolitical Stability

The price of oil is highly responsive to supply, demand, discovery of economically viable petrochemical deposits, technological development, and sociopolitical contingencies. With this in mind, access to affordable oil is a *sine qua non* of energy security, and thereby economic prosperity and national security. In Merrigan's 1975 survey [105] of projections of energy usage in the United States of America for the year 2000, the author pointed out the inherently gloomy prospect of forecasting 25 years into the future, where even the marked nonlinearities of supply, demand and population are clearly subordinated to unknowable geopolitical crises. To illustrate this, Merrigan points out the 1973–74 energy crisis (induced by embargoes on the United States in response to its support for Israel in the aftermath of the 1973 Arab–Israeli war, leading U.S. oil prices to quadruple), [106] however he could not have anticipated the much larger crash in U.S. oil imports in 1979, nor the veritable phase shift in U.S. domestic oil production in the mid-2000s, nor the collection of smaller perturbations in oil imports that can be attributed to various wars in the Middle East (Figure 2.5). Regardless of these crises, net petrochemical imports in the United States approximately doubled between 1975 and 2005, representing an average increase in imports of over 87 kL per day.

The past century has borne out a pervasive pattern of international

chicanery, malfeasance and even warfare on the part of many nations in their respective quests for energy security. The case has been made that World War I was in fact substantially a war for control over access to Mesopotamian oil required to power European navies. [107] Energy security has been explicated as a *casus belli* in U.S. foreign policy. [108] The notion that the 2003 invasion of Iraq by the United States was a *'war for oil'* is too ubiquitous to attribute. Ironically, the 1979 oil crisis was precipitated by the Iranian Revolution, overthrowing a government that was installed by the United States for reasons of oil access. This was done in part to forestall Soviet involvement in the same nation, for the same reasons. [109], [110]

Desire for access to Kuwaiti oil was a direct cause of the Iraqi invasion of Kuwait in 1991, accompanied by the undercutting of the Iraqi economy by Kuwaiti exports and the allegation that Kuwait had improperly removed U.S. $2.4B (approximately U.S. $4.4B in 2016) [111] of oil from the contentious Rumaila oilfield straddling the Kuwait–Iraq border, potentially with the assistance of directional drilling technology. [112] The ignition of Kuwaiti oilfields and the destruction of wellheads during the Iraqi retreat from Kuwait precipitated an ecological disaster that was identified by Carl Sagan as having the potential to trigger climatic effects akin to nuclear winter. [113] This is not to single out the Middle East as an energy security arena; the calculus of energy *realpolitik* is an international

phenomenon and multipolar exclusive economic zone (EEZ) claims in the South China Sea and Timor Sea—driven in large part by suboceanic petrochemical wealth—represent a continuing source of tension and intrigue. [114]–[116]

It is an uncontroversial and even mundane observation that nations are willing to violate national sovereignty and go to war over access to fossil fuels. Therefore, energy independence takes on an additional moral dimension directly measurable in the destruction of nations.

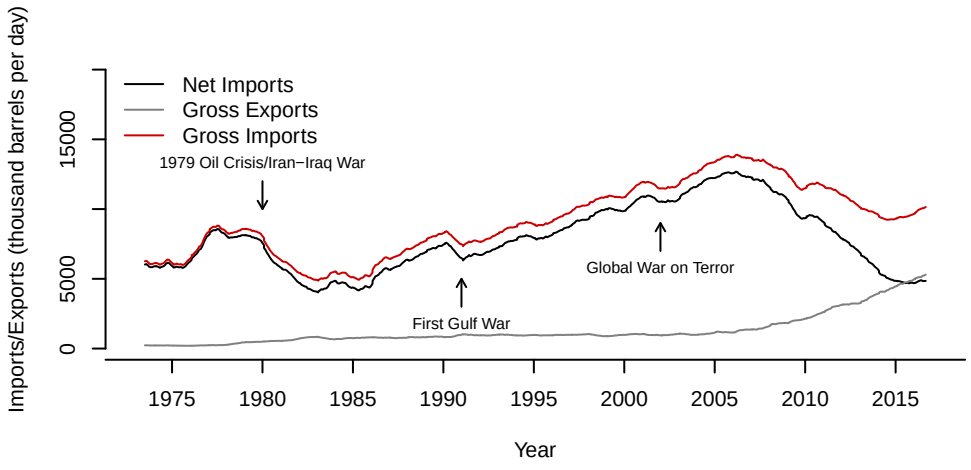


Figure 2.5: Trends in oil import and export in the United States of America between 1972 and February 2017. Data has been deseasonalised by means of a rolling average filter with a 12-month window. Data obtained from the United States Energy Information Agency. [117]

2.1.3 Renewable and Non-Renewable Energy and The Environmental Burden of Extraction

Oil, coal, and natural gas are considered to be non-renewable resources as they are generated on geological timescales by metamorphism of buried biological matter. Similarly, radionuclides relevant to fission power generation are remnants of supernova nucleosynthesis and daughter isotopes thereof and are therefore not renewable. This stands in contrast with a number of other energy resources which are considered renewable because of their short regeneration timescales (*e.g.* hydroelectric, biomass, bioethanol), because they are functionally unlimited under current usage patterns (*e.g.* wave power, geothermal, wind, and speculatively $^{2}H/^{3}H$ nuclear fusion [118]), or because human activity cannot actually affect supply (solar).[c] Indeed, oil and coal deposits represent the fixation of ancient solar energy and human industrial activity represents a massive discharge of this stored energy (see for instance the perspective of Schramski, *et al*. [120]).

The spectre of petrochemical depletion is a source of considerable concern, with Miller and Sorrel suggesting that there exists a *'growing consensus that the era of cheap oil has passed'* [121] Scholarly discussion

[c]The Sun will continue to produce solar radiation until well after the Earth is incapable of supporting liquid water due to stellar evolution, approximately 1 billion years hence. [119]

of the concept now known as *'peak oil'* was first undertaken by Hubbert within the context of domestic U.S. oil production, [122] however the dynamic of resource discovery and exploitation that Hubbert describes is generalisable across the globe and with respect to other mineralogical resources, under the assumption that these resources are functionally non-renewable. The trends of progressive resource recovery under the conditions of diminishing returns have a characteristic profile known as the Hubbert curve, represented in Figure 2.6. Localised depletion curves for national oil reserves ([123] and references therein) have approximated the Hubbert trend and it is a reasonable inference that global production will recapitulate historical behaviour.

Miller and Sorrel indicate that progressive petrochemical depletion—manifesting as unacceptable deficits in cheap energy supply—will motivate *'discovery and enhanced recovery of conventional oil [...] and the development of* 'non-conventional' *resources such as oil sands'*. [121] Expansion of petrochemical exploration as well as enhanced recovery and recovery of non-conventional sources have troubling environmental implications due to the capacity of these methods to intrude into natural and human environments in unexpected and potentially harmful ways. Emblematic of these concerns is the practice of hydraulic fracturing (*'fracking'*), an enhanced recovery technique originally developed in the 19th century that has gained popularity in recent years for exploitation of

hard-to-get oil and gas deposits, notably in the United States.[d] Fracking has been implicated in groundwater contamination and induced seismicity. [124], [125] Similarly, the exploitation of oil sands has raised serious environmental and health concerns due to a combination of chemical disruption, intrinsic greenhouse gas emission,[e] and the ecological footprint of extraction facilities. [126], [127] Concerning expansion of petrochemical discovery, interest in Arctic oil and natural gas has sharply spiked in the last decade due to the inference of large, unexplored suboceanic reserves. [128] Development of this region poses serious questions for the regional environment. Antarctica is also known to possess (difficult to obtain) fossil fuel reserves, however it is protected by the Madrid Protocol from development until at least 2048. [129], [130]

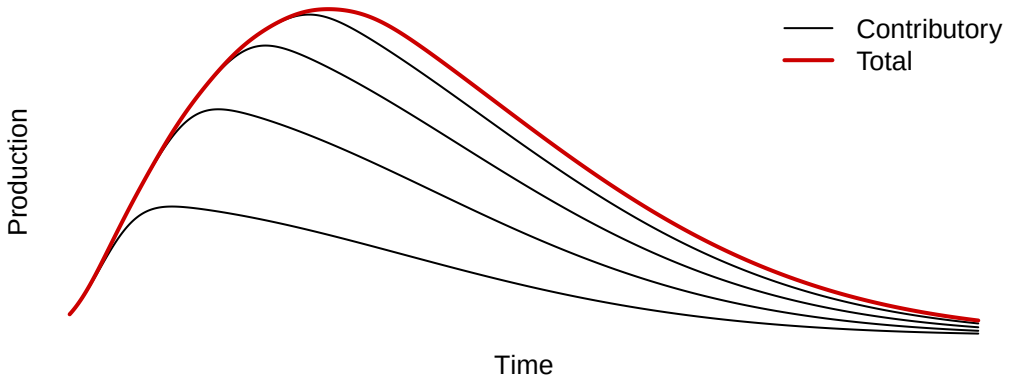


FIGURE 2.6: Prototypical Hubbert curve, comprised of a series of sequential discoveries represented by skew normal distributions, decreasing in size by 20% each time. After Bentley, *et al*. [123]

[d]The United States Environmental Protection Agency nominates approximately 275,000 hydraulically fractured wells drilled between 2000–2013. [124]

[e]Although high uncertainties exist in the total greenhouse impact of oil sand extraction with respect to conventional sources, see [126].

2.2 Hydrogen Economies

As pointed out by Romm in reference to Bockris, the notion of a hydrogen economy is distinguishable into two distinct and largely non-overlapping components: the generation of hydrogen in a renewable fashion, and its storage, transport and eventual combustion. [131] These two components comprise the two halves of a conservative and at least potentially carbon-neutral economic cycle, however either can be developed and operated independently of the other. Anecdotally, there appears to be a widespread belief that most of the world's hydrogen is produced by photovoltaic-driven electrolysis of water (*i.e. 'brine electrolysis'/'alkaline electrolysis'*). Whilst this particular modality is at least physically plausible for the generation of hydrogen at scale (recall the Downs process for production of chlorine, or the electrolysis of molten fluxed alumina to produce aluminium, at massive energetic expense), [132] it is not the case that much hydrogen is at present generated this way. Rather, most of the world's hydrogen production comes from the steam reforming of fossil methane, through a process that necessarily produces oxides of carbon. Therefore, most of the world's hydrogen production is neither carbon neutral, nor renewable.

2.3 Inadequacy of contemporary hydrogen generation methods as bases for a hydrogen economy

The two principal industrial methods of hydrogen generation are catalytic decomposition of petrochemicals (the steam reforming process) and brine electrolysis. These methods are inadequate for supporting a hydrogen fuel economy within the context presented previously.

2.3.1 Steam reforming

This process [133] uses a heterogeneous, nickel-promoted cycle to produce hydrogen gas from a hydrocarbon (generally methane, however the process is adaptable to a wide variety of olefins) through reaction with steam (Equations 2.1, 2.2). The initial reaction generates H_2 and CO, whilst subsequent reaction with steam generates further H_2 and oxidises CO to CO_2. The reaction requires considerable gage pressure and temperature (≈1025 K to 1075 K, 0.3 MPa to 2.5 MPa) to drive the oxidation.

$$CH_{4(g)} + H_2O_{(g)} \xrightarrow[\text{Ni cat.}]{\text{T, P}} 3\,H_{2(g)} + CO_{2(g)} \tag{2.1}$$

$$CO_{(g)} + H_2O_{(g)} \xrightarrow[(2)\ 465-485\,K]{(1)\ 625\,K} H_{2(g)} + CO_{2(g)} \qquad (2.2)$$

Reforming can also be performed as a partial oxidation with stoichiometrically-controlled O_2. This process (Equation 2.3) yields CO which can be likewise be oxidised *via* the water-gas shift reaction to yield further CO_2 and H_2.

$$2CH_{4(g)} + O_{2(g)} \longrightarrow 4H_{2(g)} + 2CO_{(g)} \qquad (2.3)$$

Ultimately, all forms of the steam reforming process result in the generation of CO_2 or CO, both of which are implicated in radiative forcing. [134] Furthermore, this route generally requires a non-renewable petrochemical feedstock. Improvement of steam reforming technology is considered an intermediate step to a hydrogen economy. It is estimated that 45% of global industrial hydrogen generation is *via* steam methane reforming, whilst a variety of other methods including coal and biomass gasification or pyrolysis effectively reduce to steam reforming of cracked olefins. [135] It is estimated that 95% of all hydrogen gas production in the United States can be ultimately attributed to steam reforming. [136]

The generation of hydrogen through *in situ* steam or heterogeneous catalytic reforming of methanol and other alcohols has been identified as a potential method for fuelling hydrogen-powered motor vehicles [137] as

well as for on-demand generation of fuel hydrogen down to microfluidic scales. [138] Despite the very high mass energy density [139] of molecular hydrogen ($140\,\mathrm{MJ \cdot kg^{-1}}$ / $280\,\mathrm{kJ \cdot mol^{-1}}$), this is a non-ideal medium for energy storage. Methanol is significantly more convenient to handle than hydrogen as it is a liquid at standard temperature and pressure (STP) and thus does not require cryogenic or high pressure storage. [140] Moreover, molecular hydrogen is known to cause several forms of metal degradation through lattice percolation, further complicating storage and handling. A variety of methods for the storage of molecular H_2 are being investigated; these include clathration, [141] physisorption onto graphene or fullerenes [142], [143] and adsorption to (or percolation into) a suitable metal matrix.[f] Additionally, facile release of H_2 from hydrides is being explored. [146]

2.3.2 Direct Electrolysis

Whilst direct electrolysis of water represents a relatively minor source of industrial H_2, it is noteworthy in that its fundamental electrochemical reactions (Equations 2.4, 2.5) do not evolve CO_2. Under mild conditions,

[f] Palladium is capable of reversibly sequestering large quantities of hydrogen, however economic considerations preclude its use in automotive fuel storage. [144] Nevertheless, economic mining of platinum group metals such as palladium from certain classes of near-Earth objects that have not undergone gravitational fractionation in the same way that Earth has may prove technically and economically feasible in the intermediate future. [145]

this reaction demands a redox potential of ≈1.2 V, however the anode reaction (Equation 2.5) proceeds very slowly at this potential. At 573 K the reaction proceeds at 1.0 V, [147] whilst electrolysis of steam at temperatures as high as 1273 K has been proposed as a highly efficient electrolysis regime (so called High Temperature Electrolysis (HTE)). In this modality, the water heating source is typically a nuclear reactor, which Yildiz, *et al*. indicate *'necessitate materials development and intricate design requirements for integrated plant layouts'*. [148] Such technologies are related to the speculative nuclear-assisted generation of hydrogen through chemical cycles at very high temperatures (§2.4.2). Sigurvinsson, *et al*. also demonstrate a credible hypothetical cycle exploiting geothermal energy in concert with an advanced heat exchanger system. [149]

At potentials above ≈1.9 V, oxidation of chloride becomes a competitive reaction. Whilst this is an intended end product of brine electrolysis (Eq. 2.6), chlorine may constitute an unwanted environmental pollutant in water scission processes. Conversely, the performance envelope of PSII is approximately invariant with regards to overpotential ($\approx 1000\,\mathrm{kat}\cdot\mathrm{mol}^{-1}$). [150]

$$4\,H^{+}_{(aq)} + 4\,e^{-} \longrightarrow 2\,H_{2(g)} \tag{2.4}$$

$$2\,H_2O_{(l)} \longrightarrow 4\,H_{(aq)}{}^{+} + O_{2(g)} + 4\,e^{-} \tag{2.5}$$

This reaction is thermodynamically disfavoured, and must be driven with an electric potential. The source of this electricity ultimately determines the greenhouse potential and sustainability of this process. Alternatively, hydrogen may be generated through brine electrolysis [151] (evolving H_2 and Cl_2) as a secondary product of the chlor-alkali and mercury amalgam cell processes, with the net reaction given in Equation 2.6. The mercury amalgam process is not environmentally friendly due to its cycling of large amounts of Na/Hg amalgam, which inevitably escapes the cell in some quantity over time (*e.g.* the Ontario Minimata disease tragedy of 1970 [152]).

$$2\,NaCl_{(aq)} + H_2O_{(l)} \xrightarrow[2.9\text{-}4.2\,V]{} H_{2(g)} + NaOH_{(aq)} + Cl_{2(g)} \tag{2.6}$$

Assuming that solutes do not act as interferents in a photolytic cycle, water represents a functionally unlimited feedstock, as the product of hydrogen combustion is water.

2.4 Speculative modalities

There exist a number of intriguing water splitting modalities that have not been developed in an economically meaningful sense. In the case of

high-temperature thermolytic, electrolytic or electrochemical processes, these methods pose extraordinary process engineering challenges. Metal oxidation processes are more likely to be of relevance to *in situ* hydrogen demands.

2.4.1 Thermolysis

At temperatures above approximately 2000 K at 100 kPa water vapour becomes thermodynamically unstable, and at approximately 3500 K the equilibrium molar fraction of $H_{2(g)}$ is maximised (at around 0.2). [153] Therefore, in this (extreme) temperature regime, direct thermolysis of water vapor becomes feasible. However, thermolytically generating and separating H_2 from a superheated mixture of O_2, hydroxide, and atomic oxygen and hydrogen poses engineering and safety challenges that have left the field relatively fallow. [154]

2.4.2 Nuclear/Electrochemical Water Splitting

There exist a number of speculative hydrogen generation techniques that leverage the high temperature of a nuclear reactor in either an electrolytic, thermolytic, or electrochemical cycle. These techniques are not commerical sources of hydrogen production, and the coupling of water scission reactions to the operation of a very high temperature nuclear reactor

poses extraordinary process and materials engineering challenges.

2.4.2.1 Westinghouse Sulfur Cycle

The Westinghouse cycle is water oxidation cycle that achieves hydrogen production at voltages as low as 170 mV. This process was developed in the 1970s by the Westinghouse Corporation. It consists of two steps, enumerated in Equations 2.7 and 2.8. In the first step, sulfuric acid is decomposed into water, oxygen, and sulfur dioxide. In the second step, sulfur dioxide assists electrochemical water oxidation in an electrolyser, regenerating sulfuric acid and producing dihydrogen. [155], [156]

$$H_2SO_{4(g)} \xrightarrow[cat.]{>1000\,K} H_2O_{(g)} + SO_{2(g)} + \frac{1}{2}O_{2(g)} \tag{2.7}$$

$$2\,H_2O_{(l)} + SO_{2(aq)} \xrightarrow[0.17\,V]{} H_{2(g)} + H_2SO_{4(aq)} \tag{2.8}$$

The Westinghouse cycle requires high temperatures and pressures for efficient reduction of intermediate SO_3, cited by Farbman as up to approximately 1140 K and 2 MPa. Therefore, the Westinghouse cycle was envisioned as being likely coupled to a high-temperature nuclear fission reaction, with the attendant process and materials engineering challenges this entails. [156] Jomard, *et al*. cite the oil shocks of the 1970s and 1980s (discussed in §2.1.2) as temporary impetus for research in this

area. [157] They further indicate that interest in the Westinghouse cycle has been recently reignited in light of contemporary energy security and global warming concerns, although the cycle does not appear to have been commercially implemented to date.

2.4.2.2 Sulfur-Iodine cycle

The Sulfur-Iodine cycle is a thermochemical hydrogen production cycle that involves reaction of water with I_2 and SO_2 to generate HI and H_2SO_4, which are then thermalised to produce H_2 and O_2, respectively. The essential chemical reactions of the cycle are given in Equations 2.9–2.11. Similarly to the Westinghouse cycle, the high temperatures required for thermolysis of HI and H_2SO_4 are intended to be provided by a nuclear reactor. [158]

$$2\,H_2O_{(l)} + SO_{2(g)} + I_{2(l)} \longrightarrow H_2SO_{4(aq)} + 2\,HI_{(aq)} \tag{2.9}$$

$$2\,HI_{(aq)} \xrightarrow[\approx 775\,K]{} I_{2(aq)} + H_{2(g)} \tag{2.10}$$

$$H_2SO_{4(aq)} \xrightarrow[\approx 1125\,K]{} + H_2O_{(g)} + SO_{2(g)} + \frac{1}{2}O_{2(g)} \tag{2.11}$$

2.4.3 Metal oxidation

Hydrogen may be produced by the oxidation of a metal by water. For instance, the reaction of sodium with water is given in Equation 2.12.

$$2\,Na_{(s/l)} + 2\,H_2O_{(l)} \longrightarrow 2\,NaOH_{(aq)} + H_{2(g)} \tag{2.12}$$

Reactions of this sort have been proposed for *in situ* generation of hydrogen from water. [159] The requirement for a reduced, non-passivated and sufficiently reactive metal such as powdered aluminium generally implies an energetically expensive metal refinement process, which represents the fixed energy source. With reference to earlier discussion of nuclear-assisted hydrogen generation modalities, the high-temperature reaction of water with the zirconium alloy Zircaloy within nuclear reactors is an unintended and potentially dangerous source of explosive hydrogen generation, constituting the failure mode of the Fukushima Daiichi power plant biological shields in 2011, as well as being responsible for an explosion during the Three Mile Island accident of 1979. [160], [161]

2.5 Conclusion

Combustion of fossil fuels is a leading contributor to anthropogenic climate change, the impacts of which are likely to be substantially disruptive to society and its environment in the near and extended future. Concomitantly, dwindling petrochemical reserves are both national security assets and liabilities, and the increasing attractiveness of extracting nonconventional sources poses ecological concerns in and of itself.

The construction of a hydrogen economy is a plausible route away from dependence on fossil fuels and thereby long-term energy security and reduction in greenhouse gas emissions. Nevertheless, contemporary hydrogen generation methods are not fit for purpose on account of either requiring fossil fuels as feedstocks or requiring an external energy source to effect water splitting. Artificial photosynthesis, as introduced in Chapter 1, has the potential to yield intrinsically solar-powered water splitting cycles, thereby achieving hydrogen production in a carbon-neutral capacity.

Computational Chemistry 3

> Mathematics would certainly have not come into existence if one had known from the beginning that there was in nature no exactly straight line, no actual circle, no absolute magnitude.
>
> *Friedrich Wilhelm Nietzsche*

3.1 Introduction

This chapter aims to give a brief overview of computational chemistry, the chief subdivisions of the field, and the situation of methods relevant

to the publications presented in this thesis within the field as a whole. Particular focus will be directed to Density Functional Theory (DFT), key principles and vocabulary of the theory, and the features of contemporary practical computational chemistry implementations.

Computational chemistry is the discipline of solving chemical problems by means of computational algorithms. Multiple distinct subdisciplines exist, although a wide degree of overlap can exist between these disciplines due to the desirability of coupling differing levels of computational abstraction together for a variety of real-world problems. A principal distinction between computational methods in chemistry is whether or not explicit treatment of electronic structure is undertaken. This is to say that computational chemistry can be divided into classical models, which neglect explicit treatment of electronic structure, and quantum mechanical (QM) models, which explicitly solve approximations of the Schrödinger wave equation. An abridged schematic of how computational chemistry methodologies relate to each other is provided in Figure 3.1.

The Schrödinger equation $\hat{H}\Psi = E\Psi$, whilst exactly describing the energy of the stationary states of a quantum-mechanical system, is formally insoluble for the majority of molecular species that a chemist may be interested in, on account of those problems being *many-body problems*. [162] Specifically, analytic (or exactly convergent) solutions to the Schrödinger

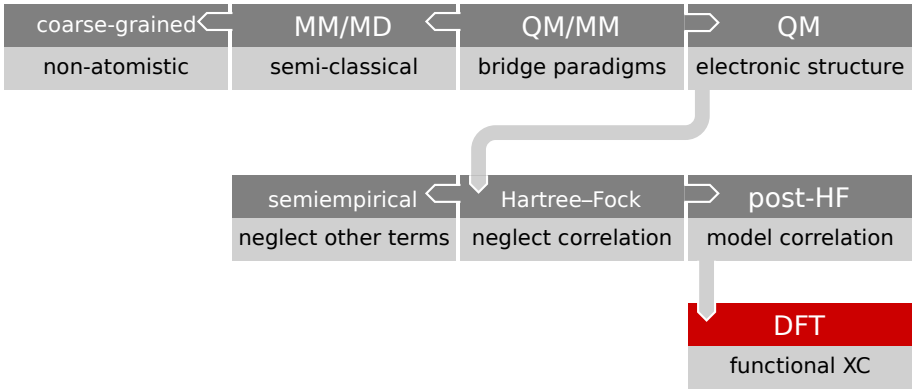


FIGURE 3.1: An abridged overview of relationships between classes of computational chemistry methods. Whilst DFT is technically a post-Hartree–Fock method, it is unusual in that it approximates the exchange and correlation interactions as a functional of the electron density.

equation do not exist for more than a handful of species containing more than two particles, and those solutions that are known are limited to species for which analytic knowledge of electronic properties is of niche value (*e.g.* H_2^+, [163] He^0, and isoelectronic homologues ([164] and references therein)), and in which certain symmetries can be exploited to reduce the problems in question to effective two-body problems. This intractability necessitates the development of approximate numerical solutions to the Schrödinger equation. At present many such methods exist, offering a hierarchy of relative accuracy and computational expense to theoreticians. Electronic structure methods exist which are suitable for the time-efficient simulation of systems containing tens of electrons to those containing a few hundred atoms.

Classical models, on the other hand, typically neglect explicit modeling of the electronic structure of a molecule in favour of heuristics that are significantly simpler to compute. These methods permit the simulation of much larger structures, typically consisting of thousands to millions of atoms. Time-dependent classical models, known as Molecular Dynamics (MD), may be used to model the time evolution and ensemble dynamics of proteins and other large molecular or supramolecular systems under some thermostat. [165] There also exist electronic structure methods amenable to calculating the time-evolution of molecular structure (notably, the Car–Parrinello method, which integrates DFT with MD [166]) however these methods are of niche interest at present.

3.2 Practical Computational Considerations

3.2.1 Processor Speed and Memory

The ability to perform electronic structure and other chemical calculations is strongly limited by the clock speed of computer processors. CPU transistor counts have increased in an approximately exponential fashion since the 1970s, [167] however CPU clock speeds have stagnated since around the year 2000, and transistor count alone is not a good proxy for computational power. Nevertheless, parallel computing has

allowed aggregate performance of supercomputer facilities to grow in an approximately exponential fashion, as illustrated in Figure 3.2. Modern supercomputer facilities typically possess tens of thousands of processors and act as elastic computing resources for a large number of concurrent users. Therefore, it is non-trivial to assess the size of chemical systems that can be currently simulated at any given level of theory on extant hardware, as this is contingent on the resources allocated to particular problems out of a typically much larger pool of computational resources.

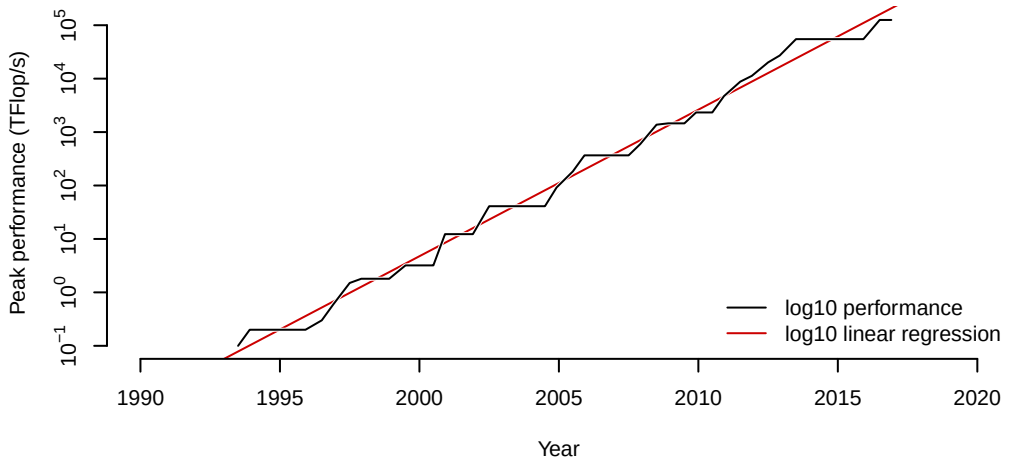


FIGURE 3.2: Illustration of peak processing power of #1 ranked supercomputer over time. Note the logarithmic scale. Derived from data obtained from TOP500 supercomputer index. 1 TFlop/s equals one trillion floating point operations performed per second. [168]

Additionally, computational chemistry in general places extraordinary demands on memory capacity and speed, with even small electronic structure calculations typically requiring several gigabytes of RAM. The price of RAM has followed a negative exponential trend since the 1960s,

permitting the economic construction of workstations containing tens of gigabytes of RAM and supercomputers containing terabytes to petabytes of RAM (Figure 3.3).

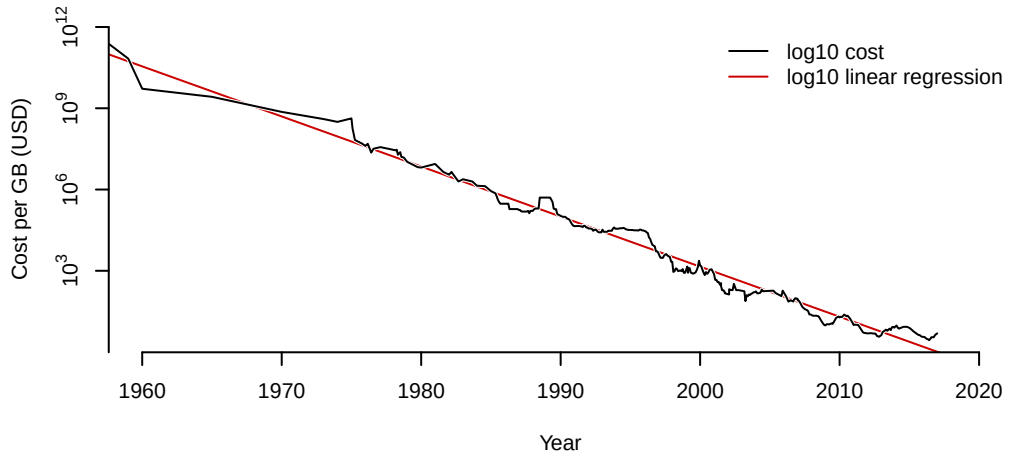


Figure 3.3: Prices of RAM over time, based on data provided by J.C. McCallum. [169] Prices are not adjusted for inflation. A spurious (out of order) datapoint was elided for December 1988.

3.2.2 Parallelism

The development of efficiently parallelised algorithms has been a considerable boon to computational chemistry. Parallelism, in this context, refers to the subdivision of a computational problem into smaller parts that can be computed simultaneously on a computer with more than one processor, a computer with hyperthreading capability,[a] or multiple discrete computers connected by some form of network, such as a local

[a] Hyperthreading refers to the ability of a single processor running a thread to allocate unused portions of the core to a second thread dynamically.

area network[b] or a wide area network such as the Internet (a generic cluster topology is illustrated in Figure 3.4). By dividing a computational problem into multiple subtasks, that problem can be completed in a shorter time. As described in §3.2.1, most contemporary supercomputers are composed of a cluster of discrete computers (termed *'compute nodes'*), each capable of multi-core parallelism. Whilst parallel algorithms must necessarily benefit over serial algorithms to justify their existence, the provision of additional processors to a parallel task typically demonstrates diminishing returns informed by Amdahl's Law, due to the existence of rate limiting steps in the algorithm that are inamenable to parallelism. [171] Nevertheless, in the absence of parallel computing, the majority of problems currently of consequence in computational chemistry would be infeasible.

3.3 Molecular Mechanics / Molecular Dynamics

Molecular Mechanics (MM) and Molecular Dynamics (MD) describe the structure and time-evolution, respectively, of molecules and supramolecular structures. MM/MD methods typically imply an eschewal of rigorous treatment of electronic structure at a quantum-mechanical level. Rather,

[b] Comprising a so-called Beowulf cluster. [170]

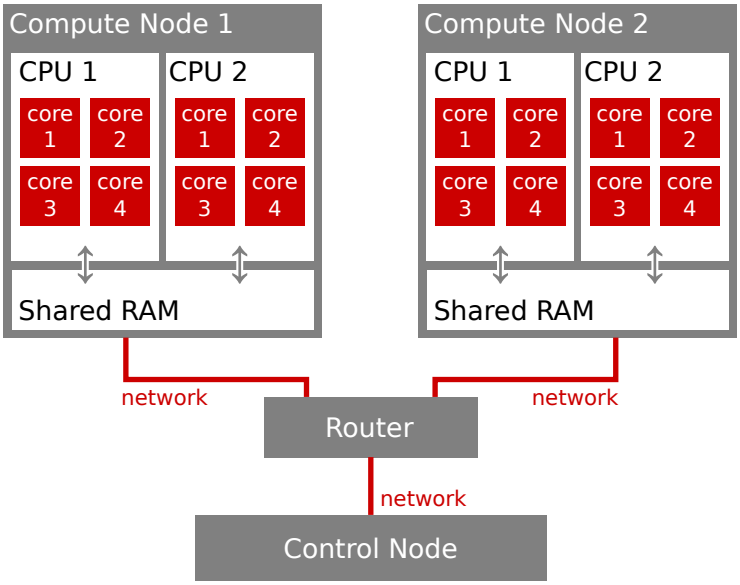


Figure 3.4: A schematic illustration of a compute cluster operating over a local area network. The cluster is comprised of 2 compute nodes, each possessing 2 CPU packages containing 4 cores each.

these methods rely on Newtonian dynamics and structural heuristics such as Valence Shell Electron Pair Repulsion (VSEPR) theory to approximate the expected behaviour of molecules. However, the relative speed with which molecular structures and forces can typically be evaluated in MD calculations with respect to Quantum Mechanics (QM) methods enables the evaluation of long-term non-equilibrium behaviour of chemical systems, such as membrane dynamics, diffusive processes, peptide folding, and ensemble dynamics thereof (e.g. [172]). The facile nature of parallelism of molecular dynamics calculations permits deployment of MD codes in a massively parallel, heterogeneous and distributed fashion. Perhaps the best known example of this paradigm is the Folding@Home

project of Pande and coworkers. Folding@Home leverages the computational resources of up to 100,000 personal computers worldwide in the form of a screensaver or background client that performs individual trajectories of ensemble dynamics simulations of protein folding. The aggregate computing power of Folding@Home (20.9 PFLOPS) is almost two-thirds of that of *Tianhe-2* (33.9 PFLOPS), the world's peak supercomputing facility as of 2015. [173], [174]

In addition to being amenable to distributed computing deployment, Molecular Mechanics (MM)/MD problems have been historically reasonably easy to parallelise on General Purpose processing on Graphical Processing Units (GPGPU) architectures such as Nvidia Compute Unified Device Architecture (CUDA) [175] and OpenCL. These technologies leverage the massively parallel vector processing architecture of modern graphical processing units and other similar coprocessors (such as Intel's Xeon Phi accelerator) to confer considerable acceleration of execution of amenable problems. [176] Development of electronic structure method implementations on GPGPU architectures has lagged considerably behind innovations in accelerated MM/MD.

Application Specific Integrated Circuits (ASICs) exist which implement molecular dynamics algorithms in hardware. An example of this is the ANTON and ANTON 2 architectures, which reflect the spatial relationships of MD simulations in the topology of the processor itself. [177]

Additionally, molecular dynamics algorithms have been successfully implemented in Field Programmable Gate Arrays (FPGAs). Particularly, Gu and Herbordt have successfully implemented multigrid-based computation of non-covalent MD interactions to FPGA hardware, delivering nearly an order of magnitude speedup with respect to a modern 2.8 GHz CPU despite the much lower clock frequency used (75 MHz). [178] Such hardware implementations—whilst requiring application-specific hardware and expressing hard-coded algorithms—can potentially provide considerable acceleration of calculations and dramatic improvements in energy efficiency per cycle. [179] Hardware-based implementations of quantum-mechanical levels of theory have lagged considerably behind either GPGPU or FPGA implementations of Newtonian mechanics/dynamics, however some progress has been made towards the implementation of density functional theory methodologies in hardware. Particularly, Wielgosz and coworkers have described a FPGA implementration of the exchange-correlation potential calculation that confers significant advantages in compute time and energy efficiency with respect to conventional implementations. [180], [181] however commercial products incorporating these innovations are not forthcoming as of 2017.

3.4 Coarse-Grained Dynamics

Coarse-grained dynamics refers to molecular and supramolecular dynamics wherein multiple atoms (potentially large chunks of a molecule or entire molecules) are replaced with rigid or soft bodies that approximate the bulk characteristics of the atoms they subsume.[c] Such simulations are positioned to straddle the space between atomistic models of matter and the Finite Element Analysis (FEA) or fluid mechanic models that characterise macroscopic materials. The consolidation of multiple atoms into blocks with reduced or frozen internal degrees of freedom is a notion that shares similarity with the Mobile Block Hessian (MBH) formalism (Chapter 7) and allied methods. These techniques freeze 'uninteresting' blocks of atoms to simplify QM calculations of vibrational structure and therefore infrared absorption spectra.

3.5 Electronic Structure Methods

Electronic structure methods are those methods in computational chemistry which seek to model the structure and properties of atoms, molecules and materials at the level of nuclear-nuclear, electron-nuclear and electron-electron interactions. Electronic structure methods typically

[c] See for instance the work of the NIH Theoretical and Computational Biophysics Group for representative research, [182]

reduce to variations on determining the electronic wavefunction of a particular molecule. The quantum-mechanical description demanded by electronic structure methods typically means that electronic structure calculations are orders of magnitude more computationally expensive (both in terms of time and memory requirements) than methods leveraging classical mechanics, on an atom-for-atom basis.

3.5.1 The Variational Principle

The variational principle in quantum mechanics is the principle that in a quantum chemical formalism, an arbitrary trial wavefunction Ψ_{trial} is guaranteed to have an associated energy $E_{\text{trial}} \geq E_0$, where E_0 is the true ground state wavefunction. [183] Therefore, given multiple trial wavefunctions, the wavefunction with the lower associated energy is a closer approximant of the true ground-state wavefunction and energy. This convergent character therefore makes numerical solutions to the Schrödinger equation amenable to systematic gradient descent towards some minimum. Density functional theory is a variational theory, as proven by the second Hohenberg–Kohn theorem. [184]

3.5.2 Geometry Optimisation, Transits and Transition State searches

The iterative approach to solving the electronic energy of a particular chemical system can be coupled with a gradient descent of the potential energy (hyper)surface (PES) associated with internuclear distances and other geometric parameters in a molecule. Minimisation of internuclear forces in a molecule represents the location of minima on the PES, whilst the location of first-order saddle points on the PES represent transition states between minima. Systematic variation of a particular geometric parameter (for instance, the dihedral angle of a rotable bond) is referred to as a *'transit'*, and can be used to probe the shape of low-dimensional sections through the PES. These concepts are illustrated in Figure 3.5.

3.5.3 Density Functional Theory

DFT is a mean field theory of electronic structure that approximates the Schrödinger wave equation. Perdew, *et al*. describe the method as *'[...] now central to most electronic structure calculations in quantum chemistry and condensed matter physics[...]'*. [185] Rather than being a many-body solution of n interacting electrons, DFT formulates n one-electron wave-functions interacting with a mean electron density. Dynamical interactions between electrons in the form of electron exchange and correlation

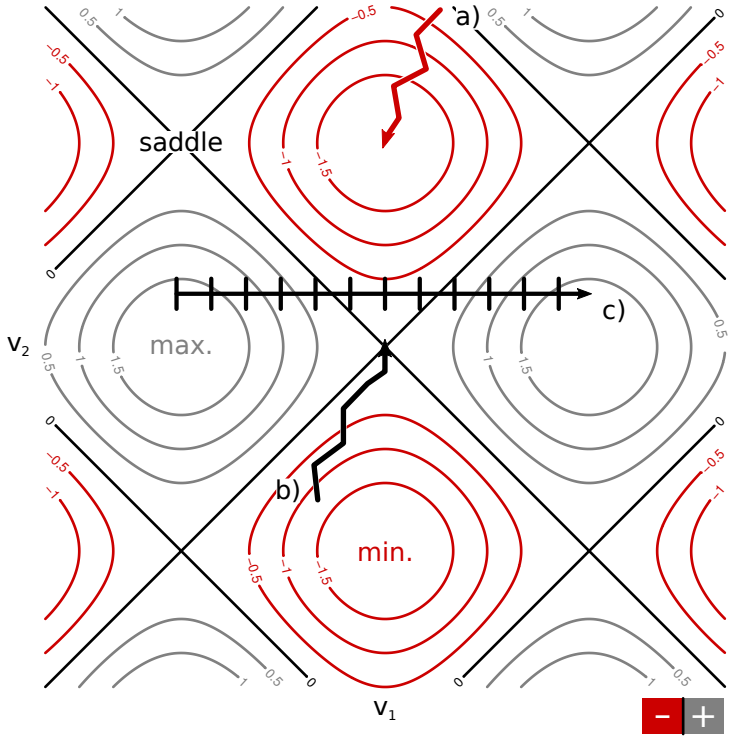


Figure 3.5: Generic illustration of a periodic potential energy surface as a function of two variables v_1 and v_2. **a**) Gradient descent pathway on potential energy surface associated with geometry optimisation. **b**) Gradient ascent pathway on potential energy surface to locate first-order saddle point, *i.e.* a transition state search. **c**) Evaluation of 1D section through potential energy surface, representing a linear transit.

are modelled as a *functional* of the electron density. This distinguishes DFT from other post-Hartree–Fock models of electronic structure such as Configuration Interaction (CI), which typically obtain correlation energies through the explicit evaluation of large numbers of configuration state functions corresponding to permutative excitations of a molecule. [186] The number of permutations required for such methods is susceptible to combinatoric explosion, rendering calculations infeasible for large molecules.

The Hohenberg–Kohn theorems prove the existence and uniqueness of a ground-state wavefunction and derived properties corresponding to any given electron density $\rho(\mathbf{r})$. [187] Hohenberg and Kohn introduced the concept of a *'universal functional'* $F[\rho(\mathbf{r})]$ which establishes this relation, [184] although it must be understood that no truly universal functional is guaranteed to exist beyond a notional infinitely large bijective mapping.[d] Whilst the concept of DFT was originally intended to address the structure of the *electron gas* (*e.g.* models of the sea electrons of a metal lattice in condensed matter physics), the principles of DFT are presently fully generalised and of great utility in the chemistry of molecules. [184]

Crucial to DFT is the accuracy and simplicity of functional approximations of the true electron exchange and correlation energies of a multielectron molecule, absent explicit modeling of dynamical interactions. Therefore, a chief interest in DFT research is the development of efficient eXchange–Correlation (XC) functionals. These functionals may be parameter-free or parameterised against known desirable properties of species of interest, which has led to the proliferation of a large range of XC functionals, with both general and specialist use-cases in mind. The availability of computationally cheap approximants of electronic structure in the form of density functionals must however be tempered with an understanding that any given density functional is only designed

[d] Frankcombe, T.J.; *Personal Communication*

to be *'good enough'* within its nominated field of applicability. This observation is a source of both self-aware reflection[e] and criticism [189] in the theoretical chemistry community.

DFT's approximate treatment of electron correlation—a crucial component of accurate electronic structure—distinguishes it from the older Hartree–Fock theory, which only considers the exchange interaction. Indeed, the correlation energy is defined as the difference between the Hartree–Fock and exact ground state energies of a molecule (Equation 3.1). [190]

$$E_C \equiv E_0 - E_{HF} \tag{3.1}$$

Nevertheless, the desirable accuracy of the Hartree–Fock exchange potential means that so-called *'hybrid functionals'* such as B3LYP and PBE0—which include a weighted combination of Hartree–Fock and DFT exchange—are quite popular. It should however be noted that not all DFT methods directly account for correlation; the $X\alpha$ method, for instance, does not explicitly model correlation. [191]

Density functionals can be divided on the particular property or properties of the density that the exchange and correlation energies are obtained from. Density functionals belonging to the Local Density

[e]Swart, *et al*. produce a yearly poll of density functional popularity, from which a so-called 'PACO' (Popularity Adapted Consensus Object) Functional is derived, itself being a weighted admixture of the most popular functionals. [188]

Approximation (LDA) and Generalised Gradient Approximation (GGA) families obtain exchange and correlation energies from the density and first derivative of the density at points in space, respectively. [192] Meta-GGA and MetaHybrid functionals are those which include a term which scales with the Laplacian of the electron density or kinetic energy density, which Tao, *et al*. characterise as *'the third rung of Jacob's Ladder of [density] approximations [...]'*. [193] In addition to an approximate hierarchy of sophistication with respect to approximations of electron exchange and correlation, functionals have also been developed with an eye to more accurate modeling of long-range behaviour (appropriate to modeling of long-range charge transfer, *e.g*. [194]) as well as dispersion correction (the attractive component of the van der Waals interaction), which is of particular interest in the modeling of intermediate-range, noncovalent interactions such as hydrogen bonding, noble gas dimerisation, and multipole interactions (*e.g*. π stacking). [195] Density functionals intended for use in obtaining specialised properties at high accuracy—without general applicability—also exist. As with other electronic structure methods, molecular orbitals are typically represented by a linear combination of atomic orbitals generated by an orthonormal basis, the features of which can be arranged into a rough qualitative hierarchy of sophistication. Additionally, density functionals relating the electron exchange and correlation to the electron density can be grouped into loose hierarchical

families (Figure 3.6).

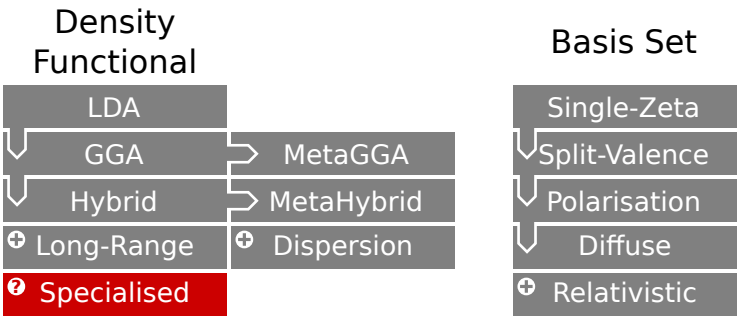


Figure 3.6: Illustration of approximate hierarchies in density functionals and basis sets.

3.5.4 Born–Oppenheimer Approximation

Within quantum-mechanical methods, the Born–Oppenheimer approximation is typically used to reduce the number of degrees of freedom (and hence complexity) of the simulation. The Born–Oppenheimer approximation involves the neglect of nuclear motion from an electronic structure simulation, noting that to a first order approximation the electronic wavefunction and nuclear wavefunction are separable. [196] This approximation can be compared to the neglect of plate tectonics when planning air routes. The approximation is made on the basis that electrons are much less massive than nuclei and that the relative timescales of nuclear motion with respect to electronic motion are much greater. An additional feature of this approximation is that it separates nuclear symmetry from dynamical desymmetrisation by the electronic structure.

Levine [197] cites the Born–Oppenheimer approximation as a potential source of error in electronic structure calculations but notes that it is typically negligible for ground-state calculations.[f]

3.5.5 Broken-Symmetry

Electronic structure calculations can be performed either in spin-restricted or spin-unrestricted capacities. Spin-restricted calculations do not permit unpaired electrons, whereas spin-unrestricted calculations permit breaking of the spin symmetry of a molecule by separating the molecular wavefunction into magnetic orbitals. Spin-restricted calculations are inappropriate to multiplet states, as well as open-shell singlet states, where α and β spin density may not be homogeneously distributed over a molecule. Electronic structures which break spin symmetry are denoted *broken-symmetry solutions*. These states are admixtures of non-interacting pure spin states. [199]

The prevalence of unpaired *d* electrons in transition metal complexes—including bioinorganic cofactors—motivates the use of broken-symmetry approaches in modeling these species, particularly to derive the Heisenberg magnetic coupling constants of such a system. In the most elementary case—that of a dinuclear, magnetically coupled complex—the

[f] Although consequential counterexamples exist, *e.g.* Pisana, *et al.* [198] and references therein.

Heisenberg coupling energy is simply the difference in energy between the antiferromagnetic broken-symmetry state and the maximum-spin pure ferromagnetic state. [199] More complex polynuclear complexes have a rich combinatoric structure of broken-symmetry states and multiple pairwise Heisenberg coupling constants that can be calculated by evaluation of these states.

3.5.6 Basis Sets

The electronic structure of a molecule can be represented as a linear combination of atomic orbitals (the so-called LCAO approximation). In practice, these atomic orbitals are themselves represented as linear combinations of primitive functions relating wavefunction probability magnitude to radial distance from an atomic nucleus. These *basis functions* constitute a *basis set*, and variational optimisation of the coefficients attached to these functions is the objective of an electronic structure calculation. Basis sets for molecular species are most commonly composed of Gaussian or Slater functions, however plane-wave (PW) functions are typically used for periodic modeling,[g] and numerical and Daubechies

[g]Plane wave basis sets have the useful property of being themselves periodic. PW basis sets are selected for energetic convergence with respect to the maximum spatial frequency of the basis functions. Whilst PW basis sets are typically used for calculations on structures that are periodic in at least one dimension, they can also be used to model aperiodic molecules, provided that they unit cell in which the molecule is situated is sufficiently large that self-interaction error is negligible. Unfortunately the requirement to minimise the self-interaction in molecular systems renders PW basis sets relatively inappropriate for aperiodic systems. [200]

wavelet [200] bases are also known.

Gaussian basis functions have the convenient properties of analytic integrability and that the product of two Gaussian functions is another Gaussian function. However, in comparison to Slater functions, Gaussians are poor approximants of actual atomic wavefunctions. As such, the sum of several Gaussian functions is required to adequately model both the nuclear cusp and the radial attenuation of atomic orbitals. Conversely, Slater functions are not analytically integrable due to the presence of a singularity at $r = 0$. Accordingly, numerical integration must be used in calculations using Slater basis sets. This integration comes at increased computational expense for comparable accuracy with respect to the analytic result, however implementations that use Slater basis sets such as ADF permit adjustment of the numerical integration accuracy according to the specifications of the end user. As a point of reference, the STO–3G basis set (Slater-Type Orbital, Three Gaussian) is a Gaussian basis set that minimally approximates Slater functions through a linear combination of three Gaussian functions, providing a rough indication of the relative qualities of Gaussian and Slater functions. Basis sets may have their outermost (valence) orbital representations subdivided so as to improve accuracy; such basis sets are known as 'split valence' basis sets, and are distinguished from so-called single-zeta basis sets, which have unsplit valence levels. On top of this, basis sets may be given polarisation

functions, which provide the additional flexibility to accurately model polarised electron densities in molecules, and diffuse functions, which are used to accurately model highly radially-diffuse electron densities, such as seen in anions and excited electronic states. [201], [202] Conversely, pseudopotentials or Effective Core Potentials (ECPs) may be used to eliminate costly basis functions associated with core electrons, where those electrons are considered irrelevant to the molecular properties. [203]

3.5.7 Relativistic Behaviour

A consequence of the increasing nuclear charge Z as one descends the periodic table is that bound states of electrons with these nuclei (*i.e.* atomic or molecular orbitals) must have higher velocities associated with them. The relation $v_{1_s} = c\alpha Z$, where $\alpha \approx 1/137$,[h] places apparent limits on the maximum atomic number of an element at around $Z = 137$, in the non-relativistic approximation. As electrons gain mass due to relativistic motion, their orbitals undergo radial contraction with consequences for molecular structure and energetics that become serious for 2$^{\text{nd}}$-row transition metal and heavier elements. [205] A computationally simple solution is to use relativistically-contracted basis sets in concert with the Zeroth Order Regular Approximation. [206]

[h] $\alpha^{-1} = 137.035\,999\,6(37)$ [204]

3.5.8 Solvent interactions and the Polarisable Continuum Model

Solvent/solute interactions are potentially important to the quantitative and qualitative accuracy of modelling solvated systems. Whilst molecular dynamics calculations are sufficiently computationally inexpensive that it is common practice to explicitly model solvent molecules surrounding a system of interest under hydrostatic equilibrium (a so-called *'water box'*), this practice is generally inappropriate for electronic structure methods (multilevel methods notwithstanding, *e.g.* §3.6).

The importance of accurate modelling of solvent/solute interactions at a high level of theory therefore motivates the development of mean-field theoretic methods that represent the effect of a solute interacting with a homogeneous dielectric medium representing the average effect of explicit solvation. The use of a homogeneous and isotropic dielectric is defensible on the basis that under physiological conditions, hydrogen bonding networks such as those seen in liquid water undergo total decorrelation on timescales on the order of 50 fs. [207] (although see van der Post, *et al.* for a more recent caveat indicating considerably longer characteristic timescales). [208]

Contrariwise to the use of explicit solvation methods that would be extremely computationally expensive in the context of *ab initio* electronic

structure methods, solvent interactions in these systems are typically treated with implicit continuum solvation models such as the Polarisable Continuum Model (PCM), and related cognates such as the Conductor-Like Solvent MOdel (COSMO). Briefly, the general principle of PCM is to model the electrostatic interactions with a molecule at a Solvent-Excluding Surface (SES) created by rolling a ball of the effective van der Waals radius of the solvent over the surface of the union of van der Waals radii of the solute atoms.[i] From this point, the screening interaction is modelled by finite element electrostatics with respect to the mesh of the solvent-excluding surface, at some notional bulk dielectric constant ϵ.[209]

The principal difference between explicit and implicit solvation models is that implicit models do not consider specific interactions between real solvent and solute molecules but rather represent a mean effect, whereas molecular dynamics calculations involving explicit solvation achieve mean solvent/solute interactions by way of ensemble and time averaging. Implicit solvent models achieve high computational efficiency by only considering interactions germane to the solute (which is to say, interactions occuring at the interface defined by the solvent excluding surface) whereas explicit models involve computation of inter-solvent interactions in the bulk solvent phase.

[i] See Tomasi, *et al*. p. 3005 for a visual representation. [209]

The use of solvent correction is not only relevant to molecular systems in a fluid dielectric medium, but also relevant to regions of interest of proteins divorced from their wider context, *i.e.* embedded in a polypeptide matrix. In these instances it is uncontroversial to use a low-dielectric solvent model to account for the average dielectric screening interactions present in the region's surrounding protein environment.

3.6 Reconciliation of QM and MM/MD calculations by multilevel methods

Advanced methods that combine the normally incommensurate natures of QM and MM calculations into a coherent whole exist in the form of so-called QM/MM methods such as ONIOM [210] and related generalisations such as QUILD. [211] ONIOM is named as a pun on 'onion', reflecting the analogy between the layered nature of an ONIOM calculation and the concentric layers of the bulb of an onion, whilst QUILD is a pun on 'quilt', representing the patchwork application of levels of theory to a chemical system that it facilitates. Methods such as ONIOM and QUILD are particularly well-suited to modeling active sites and cofactors of enzymes, as they permit the quantum-mechanical treatment of a cofactor of interest within the context of a protein matrix treated

at a molecular-mechanical level of theory, where a reasonable approximation of forces and steric influence at the interface between high- and low-level regions can be achieved without recourse to electronic structure modeling.

3.7 Vibrational Structure

The infrared response of a molecule is obtained by the eigendecomposition of the mass-weighted Hessian matrix **H**. The Hessian is the matrix of second partial derivatives of energy with respect to the degrees of freedom of a molecule. Ideally, such matrices are symmetric about their diagonal. Eigenvectors of the Hessian represent orthogonal vibrational modes, whilst the eigenvalues of the Hessian are the square roots of the frequencies of the corresponding mode. In instances where the eigenvalue of a mode is imaginary, this corresponds to a negative frequency and represents a point of positive curvature in at least one dimension on the molecular PES at the geometric configuration of the molecule. Therefore, evaluating the Hessian has important diagnostic properties in locating potential transition state geometries and positively characterising geometric extrema in general.[j]

[j] See [212] for deeper mathematical discussion.

3.7.1 Vibrational analysis of molecular regions

The evaluation of the Hessian is potentially very computationally expensive. In the numerical case, Hessian evaluation requires numerical differentiation of every degree of freedom in a molecule, which in the Cartesian frame may be up to $3N$, where N is the number of atoms in the molecule. A number of techniques have been developed to mitigate the computational expense of Hessian computation, where only part of the Hessian is of interest and some number (potentially large) of degrees of freedom may be suppressed without appreciably harming the accuracy of calculated observables of interest, such as the vibrational frequencies of highly localised modes. These techniques are discussed in Chapter 7 and the associated publication. These techniques—especially the Mobile Block Hessian (MBH) technique—bear a degree of similarity to coarse-grained methods in molecular dynamics (§3.4), in that they freeze out degrees of freedom in nominated regions of a molecule.

3.8 Conclusion

This chapter has briefly introduced Density Functional Theory and its position within the complexity/accuracy hierarchy of computational chemistry techniques. It is apparent that quite disparate computational

chemistry techniques may share both thematic and technical commonality, although they seek to answer different questions about chemical systems. Therefore, it is important to explore computational chemistry as a broad field in order to understand DFT in context.

The approximations used in DFT are highly successful in producing accurate results for a wide range of observables of practical chemical interest with excellent scaling characteristics, which accounts for the wide popularity DFT enjoys. Nevertheless, these approximations may also carry with them them varying degrees of domain-specific applicability of which researchers should be aware.

This chapter has also introduced the state of the art of computational chemistry as it relates to computer hardware, and the extent to which the solubility of chemical questions is informed by technological development. Whilst conventional microprocessor speeds have not meaningfully improved in the past decade, parallelism has taken up the slack, whilst heterogeneous processing and specialised hardware implementations hold great promise in the acceleration of certain computational chemistry tasks, insofar as developers are motivated to, and succeed in producing efficient algorithms for these architectures. Whilst the hardware acceleration of MM/MD calculations is a rapidly maturing field, it remains to be seen whether a similar efflorescence occurs for QM methods.

Introduction to Paper I: *What does the Sr-substituted 2.1 Å resolution crystal structure of Photosystem II reveal about the water oxidation mechanism?* 4

> Like storm winds [the archons] persisted in blowing, that they might try to capture that image which had appeared to them in the waters.
>
> *The Hypostasis of the Archons*
>
> *(author unknown)*

4.1 Citation

Terrett, R.; Petrie, S., Pace, R.J.; Stranger, R.; *What does the Sr-substituted 2.1 Å resolution crystal structure of photosystem II reveal about the water oxidation mechanism?*; Chemical Communications 50 (**2014**) pp. 3187–3190 doi: 10.1039/C3CC49324E

4.2 Introduction

Substitution of the Ca^{2+} site with other metal ions has been undertaken, eliciting varying degrees of suppression of photosynthetic activity. In most instances, substitution of the Ca^{2+} ion results in complete suppression of oxygen-evolving activity. Contrariwise, it is known that substitution of PSII with Sr^{2+} retains a viable degree of catalytic activity (aproximately 50%). The evidence that Sr^{2+} substitutes Ca^{2+} in the Mn_4CaO_5 cluster of the functional form comes from a 210 pm crystal structure published by Koua, *et al*. of substituted PSII from cyanobacterium *Thermosynechococcus vulcanus*. [213] The observed rate modulation is seen as evidence that Ca^{2+} binds a substrate water in the physiological photosystem, and therefore substitution of the Ca^{2+} ion provides important diagnostics as to the mechanistic aspects of water oxidation within the OEC.

4.3 General Ca^{2+} substitution

Ghanotakis, *et al*. have also reported [214] that La^{3+}-substitution maintains a small ($\approx 4\%$) oxygen-evolving activity in PSII, however it seems plausible that imperfect washout and exclusion of Ca^{2+} in the preparation of La^{3+}-substituted PSII may result in the retention of some essentially native catalytic activity. This suggestion is motivated by observation of the high degree of similarity of the Ca^{2+} and Sr^{2+} anions with respect to other metal substitutions surveyed (in terms of *aquo* complex pK_a, K_D, and ionic radius), [215] and the considerable decrease in oxygen-evolving activity in spite of this similarity (Figure 4.1). Ghanotakis, *et al*. report that introduction of La^{3+} into PSII results in dismemberment of the holoenzyme and evident destruction of the OEC signalled by the release of Mn ions, whilst introduction of Ca^{2+} after La^{3+} treatment does not restore oxygen-evolving activity.[a]

A principal determinant of the liability of a metal ion to participate

[a]More generally, lanthanum is considered a poison of photosynthetic metabolism and therefore the the study of La^{3+} interaction with the photosynthetic apparatus is motivated by a desire to understand the interactions between soils with high concentrations of the lanthanide elements and plants growing in those soils. This is of special interest in Chinese soil geochemistry due to the anomalous accumulation of lanthanides in southern Chinese clays by ion absorption processes. [216] The contention also exists, however, that lanthanum may constitute a soil micronutrient; as an example, Zeng, *et al*. found [217] that treating *Cucumis satives* (cucumber) with La^{3+} negatively impacted oxygen-evolving activity, however this activity was preserved by inclusion of Ca^{2+} in the incubation media suggesting competition between Ca^{2+} and La^{3+} for high-affinity binding sites. Moreover, the concentration of chlorophyll in chloroplasts treated with La^{3+} was found to be increased beyond the reference level of experimental controls, establishing a case for non-obvious (and possibly hormetic) interactions between La^{3+} and plant metabolism. [214]

in substitution of a bioinorganic cofactor ion appears to be compatibility of ionic radii. This is a relation borne out in metal substitutions in many other biological systems, and is responsible for the toxicity of a number of metals in biological systems. In particular, Ca^{2+} and Sr^{2+} are quite similar in ionic radius (99 pm and 112 pm, respectively), [215], [218] and therefore the ease of replacement of Ca^{2+} by Sr^{2+} in functional PSII is not altogether surprising. Moreover, Ca and Sr have very similar *aquo* complex pK_a values, which accounts for the maintenance of water oxidising activity in PSII under Sr substitution. [215]

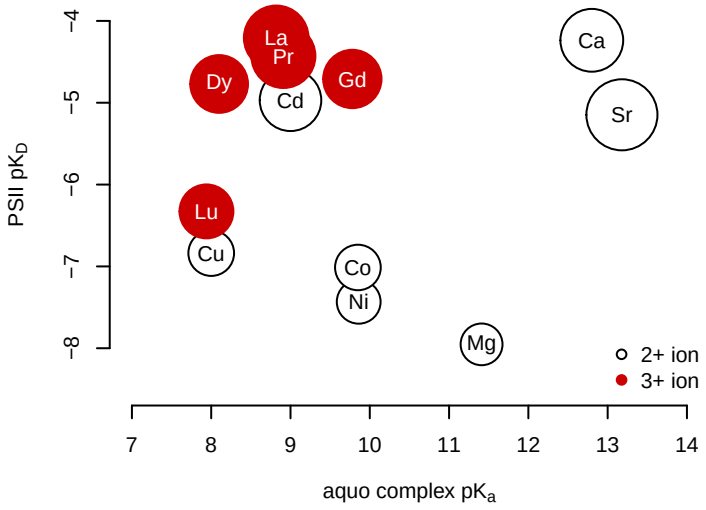


Figure 4.1: Plot of *aquo* complex pK_a and PSII pK_D of various metal ions, based on data tabulated by Vrettos, *et al.* and citations therein. [215] Relative ionic radii are expressed by the relative radii of circles in figure.

The Sr-substituted PSII structure of Koua, *et al.* is distinguished from wild-type PSII structures by its deactivated Sr–W_4 bond (227 pm to

232 pm) with respect to wild-type Ca–W_4 bonds (249 pm to 251 pm in the 195 pm structure of Suga, *et al*. [34]). This short bond is in spite of the high degree of positional similiarity between Sr-substituted and wild-type W_4 with respect to D1-Tyr161 and D1-Gln165, which are in presumptive hydrogen-bonding contact with the ligand. It was considered that accounting for this difference in bond length may provide insight into the difference in oxygen evolving activity between Ca- and Sr-PSII. Accordingly, we sought to computationally characterise the structural similarities and differences between the native form of the OEC, and the Sr-substituted form *via* Density Functional Theory, and this paper represents the culmination of those efforts.

4.4 Contributions to this work

My contribution to this work consists of development and iterative refinement of a molecular model that captures the essential structural traits of the Sr-substituted PSII structure of Koua and coworkers, as distinguished from wild-type PSII structures (*i.e*. those containing the conventional Mn_4CaO_5 cluster). This process consisted of:

4.4.1 Exploratory calculations

I performed exploratory calculations consisting of initial assessment of structural changes occuring in the PSII OEC under a variety of metal substitutions[b] by density functional theory calculation, followed by differential visual inspection and difference-of-distance-matrix analysis. These calculations provided a basis for understanding the behaviour of a Mn_4MO_5 cluster where M is a non-calcium heterometal ion, especially in terms of the geometric response of μ^x-*oxo* caps, *aqua* ligands, and *carboxylato* ligands coordinating M, and the secondary long-range implications of these structural distortions on the bulk geometry of crystallographically adapted PSII OEC models.

These exploratory calculations also provided an opportunity for the development of code to efficiently extract and visualise structural correlations and differences between differing metal substitution scenarios.

4.4.2 Model construction and evaluation

I constructed crystallographically-adapted Sr-substituted PSII models to explore a variety of ligand dispositions, constraints, and supporting hydrogen-bonding environments. In general, this process consisted of iterative elaboration of the model until a stable geometry modeling

[b] Namely, Ba, Be, Cd, Mg, and Sr.

the anomalous Sr–W_4 geometry of the crystallographic structure was achieved.

4.4.3 Model refinement and identification

I performed refinement and elaboration of the OEC model, permitting identification of a plausible aetiology of the short Sr–W_4 vector, namely W_4 existing as a hydroxide crystallographically.

4.4.4 Contributions to manuscript

Having established a plausible identity for W_4 in the crystallographic structure of Koua, *et al.*, I contributed structural data obtained from these calculations, the methodological section describing how the calculations were undertaken, and chemical visualisation of the results (specifically, Figure 3) to the manuscript.

ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Chem. Commun.*, 2014, **50**, 3187

Received 9th December 2013,
Accepted 31st January 2014

DOI: 10.1039/c3cc49324e

www.rsc.org/chemcomm

What does the Sr-substituted 2.1 Å resolution crystal structure of photosystem II reveal about the water oxidation mechanism?†

Richard Terrett, Simon Petrie, Ron J. Pace and Robert Stranger*

A density functional study of the Sr-substituted photosystem II water oxidising complex demonstrates that its recent X-ray crystal structure is consistent with a $(Mn^{III})_4$ oxidation state pattern, and with a Sr-bound hydroxide ion. The Sr-water-hydroxide interactions rationalize differences in the exchange rates of substrate water and kinetics of dioxygen bond formation relative to the Ca-containing structure.

Photosystem II in green plants and algae, is responsible for oxidising water to molecular oxygen. This occurs within a Mn_4Ca catalytic metal cluster known as the water oxidising complex (WOC). During the four-electron oxidation process, the WOC cycles through several intermediate states, $S_0 \cdots S_4$, of increasingly higher mean oxidation level.[1] However, both the mechanism of water oxidation and the metal oxidation levels throughout the S states, remain contentious.

Over the past decade, X-ray crystal structures of Photosystem II (PS II) have appeared at progressively improved resolution, from 3.5 to 2.9 Å.[2] These structures, presumed to represent the photosystem in the dark stable S_1 state, have revealed considerable detail concerning the Water Oxidising Complex in PS II, in particular, the positions of the metal atoms and their coordination. In 2011, Umena *et al.*[3] published an X-ray structure of PS II at 1.9 Å, the first at atomic resolution, which clearly resolved the positions of the Mn and Ca atoms in the WOC, the oxo-bridges connecting the metal atoms and a number of coordinated waters (Fig. 1).

Nevertheless, certain aspects of the 1.9 Å structure have been controversial.[4] The long Mn1–Mn2 and Mn2–Mn3 distances (2.8 and 2.9 Å, respectively) conflict with those derived from high-precision EXAFS measurements[5] and earlier lower-resolution X-ray structures[2c–e] of PS II. The Mn EXAFS data for the functional WOC in the S_1 and S_2 states indicate the presence of two Mn–Mn vectors of ~2.70 Å and these are consistent with values of 2.65 and 2.70 Å reported for the Mn1–Mn2 and Mn2–Mn3 distances in the earlier 3.0/2.9 Å resolution XRD structures.[2d,e] This discrepancy has prompted suggestions that the 1.9 Å structure is compromised by photo-reduction of the Mn centres during the data collection,[4a–d,6] possibly reduced to the S_{-3} state.[4d,e] However, the dosage levels used were at least 10 times less than those employed previously in XRD studies of PS II, resulting in no more than ~25% reduction to Mn^{II}.[7] Even more disconcerting is that the 2.9 Å structure, which was subject to much higher X-ray dosages, shows no significant perturbation of the Mn centres on the basis of the reported Mn–Mn distances. The other major concern in the 1.9 Å structure is the identity and position of the unusual O5 species (Fig. 1) which weakly binds to four metals of the WOC, Mn1, Mn3, Mn4 and Ca, at distances between 2.4 and 2.7 Å. If O5 is an oxo or a hydroxide group, it is unlikely to be at such long distances from the metal centres, particularly if one or more Mn are in the IV oxidation state.

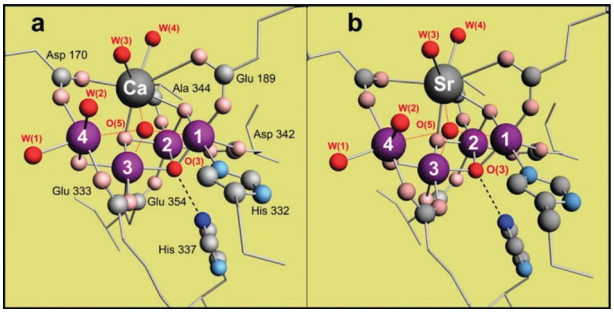


Fig. 1 Diagram detailing the close correspondence between the WOC geometries of (a) Ca-containing and (b) Sr-containing PSII, as determined from the respective 1.9 Å resolution[3] and 2.1 Å resolution[13] XRD studies on these structures.

Recent computational studies[4] have addressed these anomalous features of the WOC in the 1.9 Å crystal structure, but were unable to reproduce the long Mn–Mn distances nor the unusual position of O5. Generally, these calculations assumed a mean oxidation level of +3.5 for the Mn centres in the S_1 state of the WOC. However, spectroscopic studies on the WOC do not preclude a mean Mn oxidation level of +3.0.[8] In particular, on the basis of the 'multiline' signal observed in the EPR spectrum of the S2 state of PS II, either a

Research School of Chemistry, College of Physical Sciences and Mathematics, The Australian National University, ACT 0200, Australia.
E-mail: Rob.Stranger@anu.edu.au

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c3cc49324e

$(Mn^{IV})(Mn^{III})_3$ or $(Mn^{IV})_3(Mn^{III})$ oxidation state assignment is possible. We have labelled these alternatives as the 'Low Oxidation State' (LOS) and 'High Oxidation State' (HOS) paradigms, respectively. The HOS paradigm, with a $(Mn^{IV})_2(Mn^{III})_2$ oxidation pattern for S_1, is favoured by most groups, largely on the basis of XANES and EPR evidence.[8a] However, we recently showed that when the metal ligand environments are properly accounted for, the XANES data are more consistent with the LOS paradigm for S_1, namely $(Mn^{III})(Mn^{IV})(Mn^{III})(Mn^{II})$.[9] Further analysis of other spectroscopic data (X-ray, UV-Vis, NIR) also support this conclusion.[8b,c] Finally, recent experimental studies by Dismukes *et al.*,[10] counting electrons removed from Mn^{II} ions during functional photo-assembly of PS II, validates the LOS paradigm for the WOC.

In light of the above, we recently undertook a computational study of the WOC in the 1.9 Å crystal structure assuming that the Mn oxidation levels conformed to the LOS paradigm.[11] Importantly, in the 1.9 Å structure, His337 is directly orientated towards the μ_3-oxo bridge (O3) and, as our earlier calculations revealed,[12a] is close enough to engage in H-bonding (Fig. 1). Our modelling revealed a remarkable geometric change in the Mn_4Ca cluster, triggered by a proton transfer from a terminal water on Mn4 to the Nτ of His337. This allows His337 to engage in H-bonding with O3 which converts Mn2 from the IV to III oxidation state, thus lengthening the Mn1–Mn2 and Mn2–Mn3 bonds to around 2.9 Å, in good agreement with the crystal structure values. Deprotonation of a water on Mn4 favours the III oxidation state on this metal rather than II found in our earlier modelling of the 2.9 Å structure.[12] As a result, the sequence of Mn oxidation states change from (III)(IV)(III)(II) in the 2.9 Å structure to $(III)_4$ in the 1.9 Å structure. This change also resolved the identity of the O5 species which was found to be a water ligand, balanced at the intersection of the Jahn–Teller axes on Mn1, Mn3 and Mn4 and H-bonded to the now terminal hydroxide on Mn4 (Fig. 2a). It successfully reproduced the long bonds between O5 and Mn1, Mn3, Mn4 and Ca in the 1.9 Å crystal structure.

Attempts by others[4] to rationalize O5 as either an oxo or hydroxo group in the HOS paradigm resulted in its strong binding to a Mn centre in the IV oxidation state, either Mn1 or Mn4. To resolve this, it was suggested[4j,k] that the unusual position of O5 results from a resonance process, with O5 inter-converting between an oxo group bound to Mn1 or Mn4. The Mn centre binding O5 is always in the IV state and the other Mn in the III state. If this resonance occurs in the OEC, then it should reflect in the electron density associated with O5. However, the WOC electron density in the 1.9 Å Ca structure shows that O5 is similar to the other oxo bridges, O1 to O4, in the metal cluster, suggesting that the resonance mechanism described above is unlikely.

Given that the low oxidation paradigm successfully rationalises the WOC geometries inferred by EXAFS and seen in the 2.9 and 1.9 Å crystal structures, without invoking radiation induced changes, we now consider the recent Sr-substituted 2.1 Å resolution crystal structure of PS II.[13] This structure exhibits a number of parallels with its Ca analogue – long Mn1–Mn2 and Mn2–Mn3 distances of 2.8 and 2.9 Å, respectively, and an anomalous O5 group weakly bound (2.4 to 2.7 Å) to four metals in the cluster (Fig. 1). The His337 orientation and Nτ–O3 distance of approximately 3.0 Å suggests again that, analogous to the 1.9 Å Ca structure, His337 is undergoing H-bonding with the μ_3-oxo bridge. The fact that these features closely mimic those seen in the Ca structure would suggest that they are not an artefact of Mn photo-reduction.

To compare the two structures, we have undertaken DFT calculations,[14] assuming the LOS paradigm, on a model of the 2.1 Å Sr-substituted structure incorporating the Mn_4SrO_5 core, all coordinating amino acids (Asp170, Glu189, His332, Glu333, Asp342, Ala344, Glu354) constrained at their β-carbon positions using the X-ray structure coordinates, and bound water groups (W1···W4). Our model also included His337, Tyr161, His190, Gln165 and Arg357 as these residues H bond with W3, W4, O2 or O3. In addition, several non-coordinating crystallographic waters, in the vicinity of Sr, Tyr161 and His190, were incorporated due to their importance in the H-bonding network involving the WOC. Details of the computational model are elaborated upon in the ESI.†

The calculated metal–metal and selected metal–ligand distances for the antiferromagnetic (ABAB) coupling arrangement are listed in Table 1 along with the reported values from the 2.1 Å crystal structure (averaged over monomers A and B). For comparison, the corresponding calculated and crystallographic values for the 1.9 Å Ca

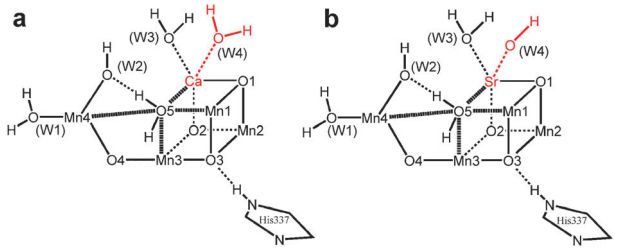


Fig. 2 A comparison of our best-fit models to (a) the 1.9 Å resolution PSII WOC[3] and (b) the 2.1 Å resolution Sr-substituted WOC.[13] For simplicity, only the metal atoms, oxo bridges, metal-coordinated water ligands, and the His337 residue are shown. The Ca- and Sr-containing models are very similar, except for the identification of ligand W4 (highlighted in red) in the two structures.

Table 1 Comparison of key crystallographic and computed metal–metal and metal–ligand distances for the Sr-substituted WOC and the Ca-containing WOC

	2.1 Å Sr structure		1.9 Å Ca structure	
Distance/Å	Expt[a]	Calc.	Expt[b]	Calc.
Mn1–Mn2	2.78	2.82	2.80	2.79
Mn2–Mn3	2.93	2.91	2.90	2.92
Mn3–Mn4	2.88	3.09	2.94	2.93
Mn1–Mn3	3.34	3.40	3.29	3.55
Mn1–Mn4	5.06	5.30	4.98	5.19
Sr/Ca–Mn1	3.56	3.77	3.48	3.56
Sr/Ca–Mn2	3.50	3.53	3.32	3.25
Sr/Ca–Mn3	3.64	3.86	3.43	3.62
Sr/Ca–Mn4	4.00	4.10	3.79	4.32
Sr/Ca–W3	2.64	2.62	2.40	2.54
Sr/Ca–W4	2.30	2.30	2.44	2.63
Mn1–O5	2.66	3.00	2.60	2.56
Mn3–O5	2.42	3.46	2.39	3.25
Mn4–O5	2.56	3.15	2.48	3.30
Sr/Ca–O5	2.59	2.90	2.64	2.50

[a] Ref. 13. [b] Ref. 3.

structure are also given. While the overall geometry of the WOC in the Sr-containing structure is very similar to the Ca structure, there are obvious differences. Firstly, the Sr atom has clearly moved outwards, away from the other three Mn ions in the Mn_3SrO_4 cubane, resulting in lengthening of all the Sr–Mn vectors relative to the Ca analogue by ~0.2 Å. This outward displacement of the Sr atom is also reproduced in the calculated structure with a similar increase of ~0.25 Å in the Sr–Mn vectors.

In contrast to the Ca/Sr–Mn distances, the differences between the reported Mn–Mn distances in the two crystal structures is much smaller, around 0.05 Å, and this is mostly consistent with the calculated values where an average difference of ~0.1 Å is seen. Similar to the Ca structure, the Sr structure exhibits long Mn1–Mn2 and Mn2–Mn3 distances of 2.78 and 2.93 Å. We were only able to reproduce these distances if H-bonding between His337 and O3 was invoked which ensured that the Mn oxidation pattern was $(III)_4$. Indeed, our calculations indicate that there is a propensity for His337 to transfer a proton to the O3 bridge. However, since the O3 protonated structures led to excessive expansion of the Mn1–Mn2 and Mn2–Mn3 distances to around 2.95 Å, subsequent calculations limited this interaction by constraining the N_{His337}–H distance to ~1.1 Å. As found in our modelling of the 1.9 Å Ca structure, in the absence of this H-bonding interaction, the Mn oxidation states adopt a (III)(IV)(III)(II) pattern and much shorter Mn1–Mn2 and Mn2–Mn3 distances are obtained.

Like the 1.9 Å Ca structure, the Sr structure possesses a weakly bound oxo group (O5) with long bonds, between 2.4 and 2.7 Å, to Mn1, Mn3, Mn4 and Sr. Our calculations show that this is only possible if two conditions are met. Firstly, O5 must be a water ligand and secondly, all the Mn centres must be in the III oxidation state which is only possible for the S_1 state of PS II in the LOS paradigm. For the $(III)_4$ oxidation pattern, O5 sits approximately at the intersection of the Jahn–Teller axes on Mn1, Mn3 and Mn4, accounting for the long bonds to these metals between 2.8 and 3.4 Å.

In addition to the longer Sr–Mn distances relative to the Ca structure, the other noticeable difference concerns the water groups (W3, W4) bound to Sr. Whereas the Ca–W3 and Ca–W4 bonds in the 1.9 Å structure are similar, averaging 2.42 Å over monomers A and B, the corresponding distances in the Sr structure are significantly different, 2.64 and 2.30 Å, respectively (Table 1). On the basis of the long Sr–W3 bond, it was suggested that this water is activated in the structure and thus a likely candidate for one of the substrate waters in PS II.[13] Conversely, since the Sr–W4 bond was only 0.1 Å shorter than the Ca–W4 bond, it was presumed to be in a non-activated form.

Initial calculations on our model system with two waters bound to Sr revealed that the calculated Sr–W3 bond was in good agreement with the observed crystallographic distance of 2.69 Å but the calculated Sr–W4 bond of 2.76 Å was almost 0.5 Å too long. Inclusion of hydrogen bonding partners Tyr161 and Gln165 reduces this distance to 2.49 Å. This prompted additional calculations where W4 was modelled as a hydroxide group rather than water (Fig. 2b), and returned Sr–W3 and Sr–W4 bond lengths of 2.62 and 2.30 Å, respectively, in excellent agreement with the crystal structure values (Fig. 3). These distances, particularly Sr–W4, are quite dependent on the H-bonding network seen in Fig. 3 involving W3, W4, Tyr161,

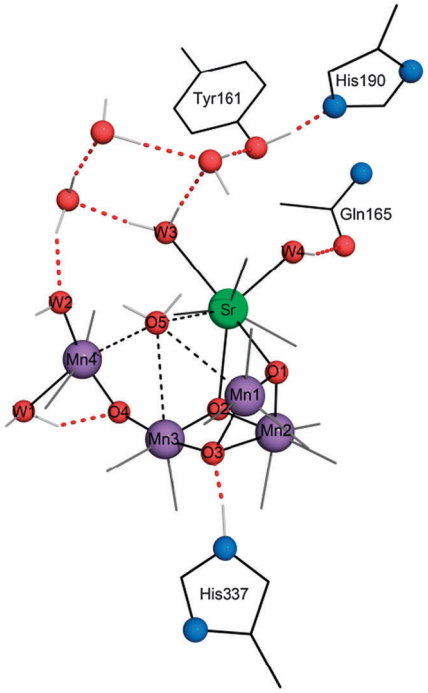


Fig. 3 Calculated WOC structure showing H-bonding network connecting Sr with Tyr161, His190 and Gln165. Inclusion of these residues, and the associated H-bonding network, yields good agreement with the crystallographic Sr–W3 and Sr–W4 distances (see Table 1). For clarity, only metal atoms, oxo bridges, and water–ligand O atoms are shown explicitly.[17]

His190, Gln165 and the nearby non-coordinating waters. When these additional residues are removed from the model, longer Sr–W3 and Sr–W4 distances are obtained. We find plausible stabilisation of the short Sr–W4 bond by coordination of W4 to either the phenolic oxygen of Tyr161 or the terminal amide oxygen of Gln165. Both orientations appear energetically competitive and elicit similarly good agreement of the Sr–W3 and –W4 distances with crystallographic data. The results strongly argue that W4 corresponds to a hydroxide and not a water ligand in the 2.1 Å Sr-substituted structure. Further, since our calculated Sr–W3 distance of 2.62 Å is very close to the reported crystal structure value, we conclude that this water is not in an activated form. However, on the basis of our earlier modelling,[11,12d,e] we consider that this ligand, and the O5 species, are likely to be the two substrate waters in PS II.

Isotope-exchange experiments using ^{18}O labelled water[15] show that in Sr-substituted PS II, the exchange rate for the slowly exchanging water increases by a factor of ~4 (in all S states). This increase is totally consistent with the Sr–W3 bond being approx. 0.2 Å longer, and thus more weakly bound, than in the Ca structure. Furthermore, while our modelling strongly suggests that W4 is deprotonated and bound to Sr in the form of a hydroxide, it also reveals that W4 undergoes H-bonding with W3 and Tyr161. This H-bonding interaction may provide insight to understanding the reduction, by a factor of 3–10, in the oxygen-evolving activity of Sr-substituted PS II compared to Ca-based PS II,[16] as it is likely to impact on the ability of W3 to form an O–O bond with the other substrate water, O5, in the higher S states, and consequently, the oxygen-evolving activity of the WOC.

In conclusion, the recent 2.1 Å resolution Sr-substituted crystal structure of PS II reveals the same unusual features found in the 1.9 Å Ca structure, namely long Mn1–Mn2 and Mn2–Mn3 distances and the anomalous O5 group weakly bound to four metals, Mn1, Mn3, Mn4 and Sr, in the WOC. Our calculations show that these features can only be rationalized in the LOS paradigm in which the Mn centres adopt a $(III)_4$ oxidation pattern as a result of His337 undergoing H-bonding with the μ_3-oxo bridge, O3. The key structural difference between the Ca and Sr systems involves the W3 and W4 groups bound to the alkaline earth metal. On the basis of our calculations, the short Sr–W4 bond of approximately 2.3 Å is consistent with a hydroxide group and not a water as previously proposed.[13] The reported Sr–W3 bond of 2.62 Å is $\sim$0.2 Å longer than the Ca–W3 bond. The weaker binding of W3 provides an obvious explanation for the observed faster exchange rate of the slow water in Sr-substituted PS II. Furthermore, the deprotonation of W4 in the Sr structure, and consequently its propensity to undergo H-bonding with W3 and Tyr161, may also be responsible for the reduced oxygen-evolving activity observed for Sr-substituted PS II. This suggests that the deprotonation of W4 to a hydroxide is not just an artefact of the 2.1 Å Sr-substituted crystal structure but the preferred form of W4 in Sr-substituted PS II.

Considering the above points, an interesting picture emerges which highlights structural factors largely ignored in mechanistic considerations to date. These have focussed closely on the Mn cluster itself and oxygen groups directly ligating it, which are presumed to include the substrate moieties themselves.[8a] The fact that the principal geometric differences between the Ca and Sr crystal structures occur in the region containing the alkaline earth metal and its associated H bonding network to Tyr161 (rather than the Mn cluster *per se*), suggests the intimate involvement of this region with substrate water binding and O–O bond formation. In particular, it is strongly implied, consistent with our earlier proposals,[12] that at least one substrate water (slow exchanging) is directly bound to the alkaline earth metal throughout the S state cycle. Further, the rate limiting events in the O–O bond forming step are likely to be directly influenced by the precise interaction of this water within a H-bonded chain proceeding to Tyr 161. This focus is quite different from other current mechanistic proposals.[4]

RS and RJP gratefully acknowledge financial assistance from the Australian Research Council, and the provision of supercomputing time on the platforms of the NCI (National Computational Infrastructure) Facility in Canberra, Australia.

Notes and references

1 K. Satoh and T. Wydrzynski, *Photosystem II: The Light Drive Water: Plastoquinone Oxidoreductase*, Springer, Dordrecht, The Netherlands, 2005, pp. 11–22.
2 (*a*) A. Zouni, H.-T. Witt, J. Kern, P. Fromme, N. Krauss, W. Saenger and P. Orth, *Nature*, 2001, **409**, 739–743; (*b*) N. Kamiya and J.-R. Shen, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 98–103; (*c*) K. N. Ferreira, T. M. Iverson, K. Maghlaoui, J. Barber and S. Iwata, *Science*, 2004, **303**, 1831–1838; (*d*) B. Loll, J. Kern, W. Saenger, A. Zouni and J. Biesiadka, *Nature*, 2005, **438**, 1040–1044; (*e*) A. Guskov, J. Kern, A. Gabdulkhalov, M. Broser, A. Zouni and W. Saenger, *Nat. Struct. Mol. Biol.*, 2009, **16**, 334–342.
3 Y. Umena, K. Kawakami, J.-R. Shen and N. Kamiya, *Nature*, 2011, **473**, 55–60.
4 (*a*) S. Luber, I. Rivalta, Y. Umena, K. Kawakami, J.-R. Shen, N. Kamiya, G. W. Brudvig and V. Batista, *Biochemistry*, 2011, **50**, 6308–6311; (*b*) P. E. M. Siegbahn, *ChemPhysChem*, 2011, **12**, 3274–3280; (*c*) W. Ames, D. A. Pantazis, V. Krewald, N. Cox, J. Messinger, W. Lubitz and F. Neese, *J. Am. Chem. Soc.*, 2011, **133**, 19743–19757; (*d*) A. Galstyan, A. Robertazzi and E. W. Knapp, *J. Am. Chem. Soc.*, 2012, **134**, 7442–7449; (*e*) A. Grundmeier and H. Dau, *Biochim. Biophys. Acta, Bioenerg.*, 2012, **1817**, 88–105; (*f*) P. E. M. Siegbahn, *Phys. Chem. Chem. Phys.*, 2012, **14**, 4849–4856; (*g*) M. Kusunoki, *J. Photochem. Photobiol., B*, 2011, **104**, 100–110; (*h*) T. Saito, S. Yamanaka, K. Kanda, H. Isobe, Y. Takano, Y. Shigeta, Y. Umena, K. Kawakami, J.-R. Shen, N. Kamiya, M. Okumura, M. Shoji, Y. Yoshioka and K. Yamaguchi, *Int. J. Quantum Chem.*, 2012, **112**, 253–276; (*i*) S. Yamanaka, T. Saito, K. Kanda, H. Isobe, Y. Umena, K. Kawakami, J.-R. Shen, N. Kamiya, M. Okumura, H. Nakamura and K. Yamaguchi, *Int. J. Quantum Chem.*, 2012, **112**, 321–343; (*j*) D. A. Pantazis, W. Ames, N. Cox, W. Lubitz and F. Neese, *Angew. Chem., Int. Ed.*, 2012, **51**, 9935–9940; (*k*) H. Isobe, M. Shoji, S. Yamanaka, Y. Umena, K. Kawakami, N. Kamiya, J.-R. Shen and K. Yamaguchi, *Dalton Trans.*, 2012, **41**, 13727–13740.
5 (*a*) M. Haumann, C. Müller, P. Liebisch, L. Iuzzolino, J. Dittmer, M. Grabolle, T. Neisius, W. Meyer-Klaucke and H. Dau, *Biochemistry*, 2005, **44**, 1894–1908; (*b*) J. Yano, Y. Pushkar, P. Glatzel, A. Lewis, K. Sauer, J. Messinger, U. Bergmann and V. Yachandra, *J. Am. Chem. Soc.*, 2005, **127**, 14974–14975.
6 J. Yano, J. Kern, K.-D. Irrgang, M. J. Latimer, U. Bergmann, P. Glatzel, Y. Pushkar, J. Biesiadka, B. Loll, K. Sauer, J. Messinger, A. Zouni and V. K. Yachandra, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 12047–12052.
7 J. Vinyard, G. M. Ananyev and G. C. Dismukes, *Annu. Rev. Biochem.*, 2013, **82**, 577–606.
8 (*a*) P. Gatt, R. Stranger and R. J. Pace, *J. Photochem. Photobiol., B*, 2011, **104**, 80–93; (*b*) R. J. Pace, L. Jin and R. Stranger, *Dalton Trans.*, 2012, **41**, 11145–11160; (*c*) R. J. Pace, R. Stranger and S. Petrie, *Dalton Trans.*, 2012, **41**, 7179–7189.
9 (*a*) A. R. Jaszewski, R. Stranger and R. J. Pace, *Phys. Chem. Chem. Phys.*, 2009, **11**, 5634–5642; (*b*) A. R. Jaszewski, S. Petrie, R. J. Pace and R. Stranger, *Chem.–Eur. J.*, 2011, **17**, 5699–5713.
10 D. R. J. Kolling, N. Cox, G. M. Ananyev, R. J. Pace and G. C. Dismukes, *Biophys. J.*, 2012, **103**, 313–322.
11 P. Gatt, S. Petrie, R. Stranger and R. J. Pace, *Angew. Chem., Int. Ed.*, 2012, **51**, 12025–12028.
12 (*a*) S. Petrie, P. Gatt, R. Stranger and R. J. Pace, *Phys. Chem. Chem. Phys.*, 2012, **14**, 4651–4657; (*b*) S. Petrie, P. Gatt, R. Stranger and R. J. Pace, *Phys. Chem. Chem. Phys.*, 2012, **14**, 11333–11343; (*c*) S. Petrie, R. Stranger, P. Gatt and R. J. Pace, *Chem.–Eur. J.*, 2007, **13**, 5082–5089; (*d*) S. Petrie, R. Stranger and R. J. Pace, *Angew. Chem., Int. Ed.*, 2010, **49**, 4233–4236; (*e*) S. Petrie, R. Stranger and R. J. Pace, *Chem.–Eur. J.*, 2010, **16**, 14026–14042.
13 F. H. M. Koua, Y. Umena, K. Kawakami and J.-R. Shen, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 3889–3894.
14 (*a*) *Amsterdam Density Functional v.2013.01*, S.C.M., Theoretical Chemistry, Vrije Universiteit, Amsterdam, the Netherlands, http://www.scm.com; (*b*) Calculations were performed in an unrestricted fashion using the GGA Becke–Perdew density functional and the TZP all-electron basis set. Relativistic corrections using the Zeroth-Order Regular Approximation (ZORA) were applied to the strontium basis set.
15 G. Hendry and T. Wydrzynski, *Biochemistry*, 2003, **42**, 6209–6217.
16 (*a*) A. Boussac and A. W. Rutherford, *Biochemistry*, 1988, **27**, 3476–3483; (*b*) K. L. Westphal, N. Lydakis-Simantiris, R. I. Cukier and G. T. Babcock, *Biochemistry*, 2000, **39**, 16220–16229.
17 *The PyMOL Molecular Graphics System, Version 1.6.x*, Schrödinger, LLC.

Introduction to Paper II:

Effect of concomitant oxidation and deprotonation of hydrated Mn centres in rationalising the FTIR difference silence of D1-Asp170 in Photosystem II 5

Altissima quæque flumina

minimo sono labuntur

Quintus Curtius Rufus

5.1 Citation

Terrett, R.; Frankcombe, T., Pace, R.; Stranger, R.; *Effect of concomitant oxidation and deprotonation of hydrated Mn centres in rationalising the FTIR difference silence of D1-Asp170 in Photosystem II*; Journal of Inorganic Biochemistry 155 (**2016**) pp. 101–104 doi: 10.1016/j.jinorgbio.2015.11.023

5.2 Introduction

Fourier transform infrared (FTIR) difference spectroscopy is a powerful tool for probing small changes in the ligand environment of protein cofactors. In the context of Photosystem II, difference spectroscopy can offer clues as to the oxidation state, ligation, and Jahn–Teller distortion of ions within the OEC, and has therefore been investigated in some detail by Debus and coworkers in combination with site-directed mutagenesis of the peptide matrix coordinating the Mn_4CaO_5 cluster. This paper explores the vibrational structure of a small molecule model of the PSII OEC, under repeated oxidation and deprotonation.

5.3 D1-Asp170 silence in PSII

A surprising result of Fourier transform infrared (FTIR) difference spectroscopy of the PSII OEC by Debus, *et al*. is the absence of evidence

that D1-Asp170 ligates a Mn ion that undergoes oxidation in S-state transitions between S_0 and S_3. [219] This result was obtained through differential analysis with respect to a mutant form of PSII that replaces D1-Asp170 with histidine.[a] Crystallographically, D1-Asp170 is consistently interpreted to ligate Mn_4 and Ca^{2+} in a $\mu^1 : \mu^1$ capacity, and D1-Asp170 was earlier identified by difference spectroscopy as either a directly coordinating ligand of the Mn_4CaO_5 cluster or as possessing a significant noncovalent interaction with the cluster. [220]

In isolation, it is not inconceivable that the OEC might advance through the S-state cycle without oxidation of Mn_4 ocurring between S_0 and S_3. However, what is more surprising is that Strickler, *et al*. also find no evidence of changes in D1-Glu189 (crystallographically ligating Mn_1) and D1-Asp342 (crystallographically ligating Mn_1 and Mn_2) between S_0 and S_3 in analogous studies. [221], [222] Contrariwise, Chu, *et al*. do find implicit evidence that Mn_2 undergoes oxidation in the $S_1 \rightarrow S_2$ transition, by FTIR difference spectroscopy probing D1-Ala344, which is the C-terminus of the D1 subunit and ligates Mn_2 and Ca^{2+} of the OEC crystallographically. [223] D1-Ala344 is distinct from other carboxylate-bearing amino acids ligating the OEC as it ligates by the α-carboxylate rather than a sidechain carboxylate, and as the C-terminus has a higher Debye–Waller factor (*B factor*) than precedent amino acids

[a] Denoted as a *D1-D170H mutant*.

in D1, however the significance of this fact to the detection of difference intensity reflecting oxidation of Mn_2 is uncertain.

5.4 Rationale for the work

This work originated as an informal suite of calculations validating the results of Chuah, *et al*. on small molecule models of the PSII OEC. [224] Chuah, *et al*. demonstrated a compensatory effect on carboxylate symmetric and antisymmetric stretching frequencies wherein simultaneous oxidation of a carboxylate-ligated Mn ion and deprotonation of an Mn-ligating *aqua* ligand occurred, lending credence to the idea promulgated by Debus, *et al*. that the D1-Asp170 difference signal may be suppressed by a similar compensatory effect involving deprotonation of W_1 or W_2 ligating Mn_4 in the physiological OEC.

Whilst the study of Chuah, *et al*. provided an elegant proof-of-concept of the aforementioned compensatory effect, considerable scope for fruitful expansion and refinement of the principle was quickly identified. We utilised the dinuclear small molecule model illustrated in Figure 5.1 as a model of the ligand environment around Mn_4 in the PSII OEC, modeling D1-Asp170 as a formate ligand bridging Mn and Ca. Di-bridging of the Mn and Ca ions (modeling D1-Glu333) was not undertaken as no

such species was investigated by Chuah, *et al*. It was also known that di-bridging of this form was unstable and liable to geometric isomerisation,[b] as observed in earlier project work performed by Kassetra von Nessi. Four avenues of additional inquiry were explored with respect to the work of Chuah, *et al*.:

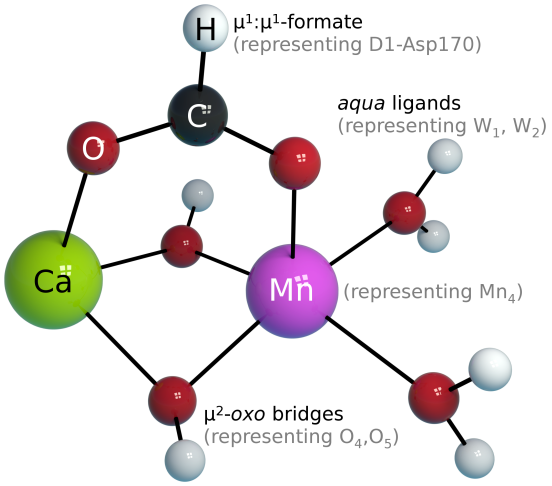


FIGURE 5.1: Representation of generic small molecule model used in FTIR difference spectroscopy study. This molecule approximates the chemical environment around D1-Asp170 and Mn_4 of the PSII OEC.

Conformational paramaterisation Chuah, *et al*. only considered single conformers of reduced and oxidised forms of their small molecule models, however the small molecule models offer a considerable degree of conformational flexibility. This flexibility was in turn anticipated to have a modulatory effect on the ligand vibrational structure that remained unexamined. This information in turn can

[b] To wit, deligation of one of the formates from Ca, producing a formate coordinated to Mn in a μ^1 fashion, with the pendant oxygen of this formate involved in hydrogen bonding with the *aqua* ligands coordinating Mn.

help parameterise the sorts of changes in ligand geometry that are consistent with the difference silence of D1-Asp170 and potentially with difference silence of carboxylates generically.

Spectral simulation The originating study only considered the migration of carboxylate symmetric and antisymmetric stretch normal mode frequencies in response to oxidation and deprotonation, without assessing changes in vibrational intensity. Nevertheless, the absorption intensity as a function of frequency is an indispensible component of difference spectra. Experimental difference spectra conceal information regarding the precise nature of their contributory direct spectra, and multiple combinations of direct spectra can give rise to the same difference spectrum. Notably, it is easy to demonstrate that the difference spectrum associated with pure migration of a single vibrational peak is indistinguishable with the symmetric interchange of absorption intensity between two proximate absorption peaks, as illustrated in Figure 5.2.

Multiple oxidation/deprotonation states A wide variety of oxidation and deprotonation patterns can be described for the model we explored, comprising a directed acyclic graph (DAG) that may be traversed in a variety of ways. Each transition may or may not permit FTIR difference silence.

Combinatoric optimisation of transitions The evaluation of the vibrational structure of multiple conformations for each oxidation/deprotonation states also permits the simulation of multiple difference spectra for each permissible transition between oxidation/deprotonation states, as a Cartesian product of initial and final vibrational structures.

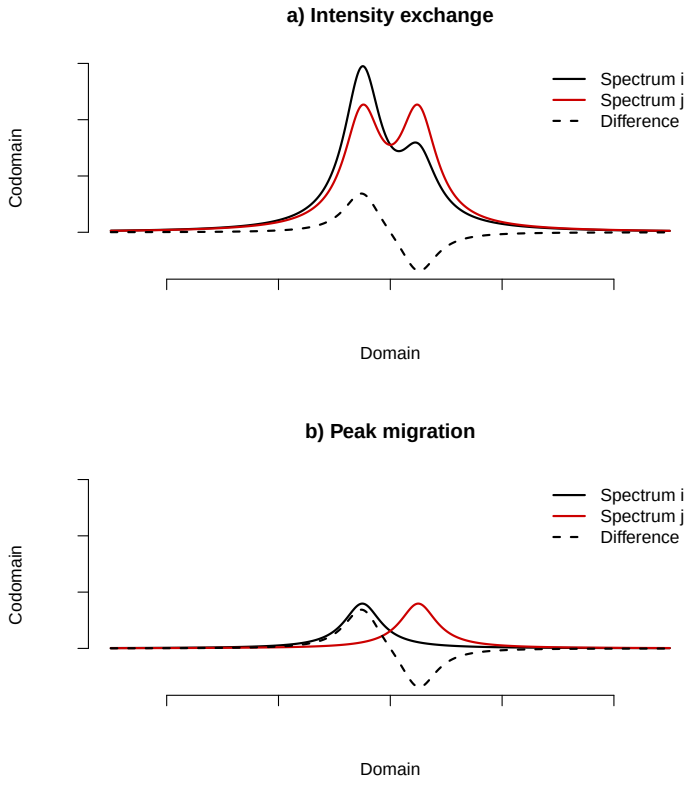


FIGURE 5.2: Idealised illustration of how multiple scenarios of direct infrared spectra i and j can account for the same difference spectra, $i - j$. **a**) Interchange of peak intensity between the left and right peaks yields a difference signal. **b**) Migration of a single peak yields an identical signal to that seen in **a**).

5.5 My contribution to this work

My contribution to this work consists of development of the underlying methodologies used in the work, their application to the difference spectra of interest, and the description, analysis, and discussion of the results obtained therein.

5.5.1 Preparation of computational models

I prepared computational models exploring the conformational space of the small molecule models under varying scenarios of oxidation and deprotonation by a combination of directed conformational modification and random perturbation, followed by structural optimisation using density functional theory.

I performed conformational exploration in both guided and unguided fashions. Unguided conformational exploration was performed by the random Gaussian perturbation of atoms within each molecule, followed by geometry optimisation. This method was chosen as it carries no parametric preconceptions about the flexibility of a molecule, in comparison to more sophisticated techniques geared towards organic species.

5.5.2 Normal mode analysis

I performed normal mode analysis of equilibrium geometries of computational models to obtain spectra for each conformer of each small molecule model.

5.5.3 Generation of combinatoric difference spectra

I implemented code in R [225] to achieve combinatoric production of difference spectra associated with compensated and uncompensated oxidations under varying patterns of deprotonation and conformational reconfiguration.

5.5.4 Spectral integration and optimisation

I implemented code in R to integrate generated difference spectra and produce optimal difference spectra with respect to combinations of oxidised and reduced small molecule models.

5.5.5 Code parallelism

I augmented the aforementioned spectral generation and optimisation code to take advantage of parallelism. This was undertaken to increase the time efficiency of combinatoric searches for optimal difference spectra

in light of the addition of oxidation and deprotonation pathways, in response to reviewer comments.

5.5.6 Molecular and spectral visualisation

I provided all molecular and spectral visualisation to the manuscript. I implemented spectral visualisation code in R for analysis and publication.

5.5.7 Compilation of manuscript

I produced the drafts of the papers, providing theoretical background, methodological information, analysis, and commentary with respect to computational results to the manuscript. I also produced additions and amendments to the manuscript as requested by reviewer commentary.

5.6 Opportunities for future research

A number of avenues are apparent for future research along the lines previously presented. One particular area of interest is the combinatoric simulation of deeper traversals of the deprotonation–oxidation graph. This is motivated by the observation that D1-Asp170 difference silence (as well as the difference silence of other residues coordinating the PSII OEC, §5.3) is typically seen over successive advances of the S-state cycle. Concretely, a potential elaboration of benefit would be to find pathways

through a graph of deprotonations and oxidations such that the sum of integrated difference spectra for some path $S_1, S_2, \ldots, S_n$ is minimised. This is contrasted with isolated pairwise evaluation of reduced and oxidised forms of the small molecule model, where optimised pairs of reduced and oxidised conformers for an initial oxidation may be inconsistent with those for a successive oxidation. This principle is illustrated in Figure 5.3. This concept is anticipated to be considerably more computationally expensive than pairwise optimisation.

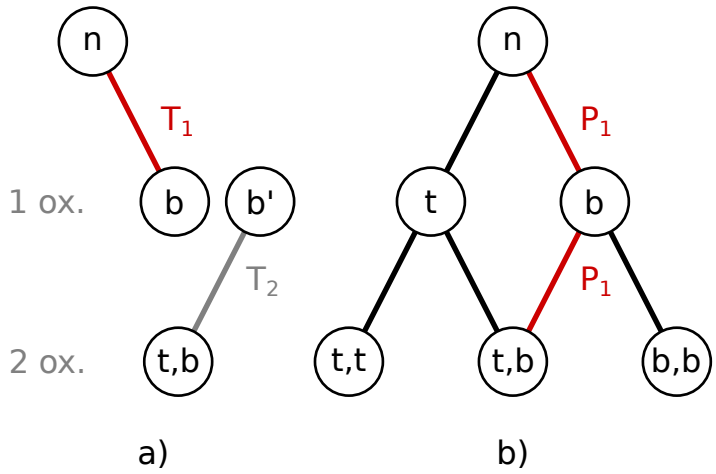


FIGURE 5.3: Abstract representation of **a**) locally and **b**) globally optimal difference paths. In **a**), only pairs of reduced and oxidised scenarios are considered, with differing patterns of terminal (t) or bridge (b) deprotonation. Whilst this locates pairs of conformers that optimally minimise difference intensity, this may mean that for a given path on the graph, two transitions T_1 and T_2 may be incommensurate (represented by *b* and *b'*). **b**) represents a globally optimal path traversal P_1, which connects three nodes—each with a nominal conformation—which when taken together minimise the sum of difference spectra for each edge in the path. However, finding such paths represents a potential *combinatoric explosion*.

Journal of Inorganic Biochemistry 155 (2016) 101–104

Contents lists available at ScienceDirect

Journal of Inorganic Biochemistry

journal homepage: www.elsevier.com/locate/jinorgbio

Effect of concomitant oxidation and deprotonation of hydrated Mn centres in rationalising the FTIR difference silence of D1-Asp170 in Photosystem II

Richard Terrett [a], Terry Frankcombe [a,b], Ron Pace [a], Robert Stranger [a,*]

[a] Research School of Chemistry, Australian National University, Acton, ACT 2601, Australia
[b] School of Physical, Environmental and Mathematical Sciences, University of New South Wales, Canberra, ACT 2610, Australia

ARTICLE INFO

Article history:
Received 23 June 2015
Received in revised form 29 September 2015
Accepted 24 November 2015
Available online 25 November 2015

Keywords:
Photosystem II
Density functional theory
Water oxidation

ABSTRACT

The observation of negligible FTIR differences in carboxylate vibrational modes for the D1-Asp170 residue of Photosystem II (PSII) on successive one-electron oxidations of the Mn_4CaO_5 oxygen-evolving complex (OEC) is counterintuitive in light of the apparent ligation of D1-Asp170 to an oxidisable Mn ion in the X-ray crystallographic structures of PSII. Here, we show computational support for the hypothesis that suppression of the FTIR difference spectrum in the 1100 cm^{-1} to 1700 cm^{-1} region of D1-Asp170 occurs by concomitant Mn oxidation and deprotonation of water ligands bound to the ligated metal centre. Density functional theory calculations on the model species $[Mn^{II}Ca(COOH)(OH)_2(H_2O)_2]^+$ over two successive oxidations of the Mn ion are performed, where those oxidations are accompanied by deprotonation of water and μ-*hydroxo* ligands coordinated to the Mn ion. In contrast, dramatically increased FTIR difference activity is observed where these oxidations are unaccompanied by proton loss.

1. Introduction

Photosystem II (PSII) is responsible for the light-driven four-electron oxidation of water effecting dioxygen evolution. This reaction is arguably the most significant biological process on Earth, with paleogeochemical evidence supporting the significant reconstitution of Earth's atmosphere by this process starting approximately 2.4 billion years ago [1]. The change is associated with the flourishing of photosynthetic autotrophs at or before this time [2].

The active site of water oxidation in PSII has been identified as a $CaMn_4$ cluster coordinated by the D1 and CP43 subunits of PSII. This site is known as the oxygen evolving complex (OEC). X-ray crystallographic evidence [3,4] at resolutions as high as 1.9 Å indicates (Fig. 1) that the cluster is comprised of an oxo-metal heterocubane consisting of three ions (Mn(1), Mn(2), Mn(3)) and a vertex with an additional pendant μ-*oxo* bridged ion (Mn(4)). This cluster accumulates four oxidative equivalents by reduction of the P680 complex followed by dioxygen evolution in a five state sequence called the Kok or *S*-state cycle (Fig. 2). Crystallographic, EXAFS/XANES, EPR and vibrational characterisation have provided rich data on the structural and electronic changes associated with the *S*-state cycle but the Mn oxidation levels in each *S*-state, precise OEC morphology and the mechanism of water oxidation remains highly contentious. [5–10].

Fourier Transform Infrared (FTIR) difference spectroscopy [11] of the carboxylate stretch signature region of Photosystem II indicates that no change in carboxylate vibrational mode frequencies, detected by FTIR difference spectroscopy, is observable in the Asp170 residue on advancing from the S_0 to S_3 oxidation states of the PSII OEC. This silence has been variously interpreted as indicating one of the following:

1. Asp170 does not coordinate a manganese ion during the S-state cycle, a hypothesis that is inconsistent with crystallographic structures indicating ligation of Asp170 to Mn(4). [3,4]. Site directed mutagenesis effecting D1-Asp170His is similarly FTIR difference silent which motivates this interpretation, [12] alongside the suggestion that crystallographic ligation of Asp170 may not represent the physiological situation. A preponderance of crystallographic evidence now exists that renders this explanation largely untenable.
2. Asp170 coordinates a manganese ion that does not undergo oxidation between the S_0 and S_3 states. The similarly observed FTIR difference silence [13] of D1-Asp342 ligating Mn_1 and Mn_2 indicates that FTIR difference silence of a residue can still be consistent with oxidation at an ostensibly ligated metal. This hypothesis is difficult to reconcile with the additional observed FTIR difference silence [12] of D1-Glu189 (which crystallographically ligates Mn(1)) and derivative mutants. If FTIR difference silence is interpreted consistently in this fashion it would require Mn(1), Mn(2) and Mn(4) to not undergo oxidation in states S_0 to S_3 or else not coordinate a manganese centre in these states.

* Corresponding author.
E-mail address: rob.stranger@anu.edu.au (R. Stranger).

http://dx.doi.org/10.1016/j.jinorgbio.2015.11.023
0162-0134/

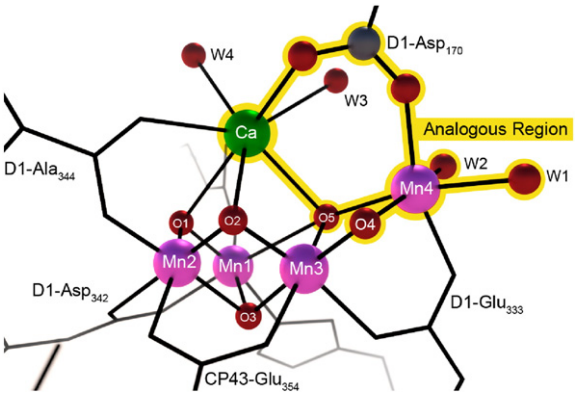


Fig. 1. Crystallographic environment of PSII OEC at 1.9 Å by femtosecond X-ray crystallography of Suga and co-workers (Protein Data Bank accession code 4UB6) [4]. The region of the OEC analogous to the small molecule model investigated in this work is highlighted in yellow.

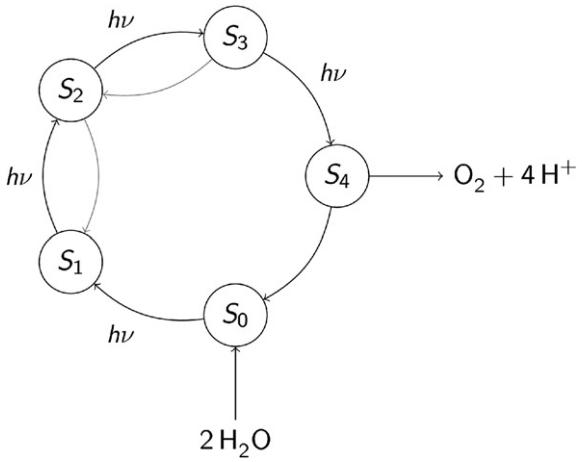


Fig. 2. Schematic representation of PSII OEC S-state cycle. Four one-electron photo-oxidations advance the OEC through S-states S_0 to S_4, with spontaneous water oxidation and regeneration of the S_0 state.

3. Asp170 is rendered FTIR difference silent by changes in ligation concomitant with oxidation. This hypothesis was initially suggested by Debus and co-workers [11].

This study addresses the third hypothesis: that the observed suppression of carboxylate symmetric and antisymmetric stretch level shifts upon oxidation of the manganese cluster in PSII occurs through concomitant deprotonation of water and/or *hydroxo* ligands bound to the manganese. It extends the recent work by Chuah and co-workers [14] involving Mn and Mn/Ca carboxylate species, by evaluating the variability in difference spectra with respect to a number of deprotonation and oxidation patterns as well as variation in ligand conformation.

2. Methodology

Calculations were executed in the Amsterdam Density Functional (ADF) suite version 2013.01 [15–17]. SCF convergence was achieved to the default criteria, whilst the gradient convergence threshold was set to one tenth of the ADF default (1×10^{-4} E_h/Å) for structures used to generate difference spectra. Frequencies[1] and intensities were calculated by analytic composition of the Hessian within the ADF software [24–26]. Calculations were performed using the Becke–Perdew exchange-correlation functional [18,19] and an all-electron triple-zeta basis set with one polarisation function (TZP), without relativistic correction to the electronic energy [20].

Calculations were undertaken on the following species which serve as convenient simplified models for the Mn4/Ca environment in the OEC, retaining all key features relevant here. The Mn(4) centre in the protein possesses two ligated water/hydroxyl groups, an *oxo* bridge, a bridging carboxylate (Glu333) and the Asp170 group which bridges to Ca.

1. $[Mn^{II}Ca(COOH)(OH)_2(H_2O)_2]^+$, the species under study prior to any oxidation or deprotonation.
2. $[Mn^{III}Ca(COOH)(OH)_3(H_2O)]^+$, representing oxidation at Mn with concomitant deprotonation of a water ligand coordinating Mn.
3. $[Mn^{III}Ca(COOH)(OH)O(H_2O)_2]^+$, representing oxidation at Mn with concomitant deprotonation of a bridging hydroxide ligand generating a μ-*oxo* ligand.
4. $[Mn^{III}Ca(COOH)(OH)_2(H_2O)_2]^{2+}$, representing oxidation at Mn without concomitant deprotonation of any ligand.
5. $[Mn^{IV}Ca(COOH)(OH)_2O(H_2O)]^+$, representing double oxidation at Mn with concomitant deprotonation of a water and the corresponding *cis*-hydroxide ligand coordinating Mn.
6. $[Mn^{IV}Ca(COOH)(OH)_2O(H_2O)]^+$, representing double oxidation at Mn with concomitant deprotonation of a water and the corresponding *trans*-hydroxide ligand coordinating Mn.
7. $[Mn^{IV}Ca(COOH)O_2(H_2O)_2]^+$, representing double oxidation at Mn with concomitant deprotonation of both hydroxide bridges to μ-*oxo* bridges.
8. $[Mn^{IV}Ca(COOH)(OH)_4]^+$, representing double oxidation at Mn with concomitant deprotonation of both water ligands coordinating Mn.
9. $[Mn^{III}Ca(COOH)(OH)_2(H_2O)_2]^{3+}$, representing double oxidation at Mn without concomitant deprotonation of any ligand.

Calculations were performed on C_s and C_1-symmetric conformers of the above species. Notably, species 1, 4, and 5 (see Fig. 3) permit structures in either symmetry. However, deprotonation of either *aqua* or *hydroxo* ligands necessarily breaks symmetry, reducing it to C_1.

These species model two principal scenarios for oxidation of Mn with differential ligand deprotonation, namely oxidation of Mn with charge compensation by deprotonation, and uncompensated oxidation with induction of a +2 or +3 charge on the species. The initial oxidation converts a Mn^{II} ion into Mn^{III}, whilst the subsequent oxidation generates Mn^{IV}. Optimised geometries of selected conformers and the reaction paths connecting these species are depicted in Fig. 3. Simulated difference spectra were produced using bespoke code developed in R, and molecular visualisations were produced in PyMOL [21,22].

3. Results

The effect of oxidation with and without concomitant deprotonation on the described species was ascertained by identification of pairs of conformations for each oxidative transition giving rise to the smallest integrated difference spectra in the signature region 1100 cm^{-1} to 1700 cm^{-1}. These minimal difference spectra (spectra with the smallest absolute integrated magnitude) and the direct IR spectra that give rise to these difference spectra are illustrated in Figs. 4 and 5.[2] In total, 2409 difference spectra were evaluated, representing the sum of Cartesian products of all redox pairs. The signature region is dominated by the characteristic presence of a

[1] In some cases, equilibrium geometries gave rise to low-energy negative frequencies. However, we find that the presence or absence of these modes does not appear to meaningfully perturb the spectrum in the carboxylate stretch region between 1100 cm^{-1} and 1700 cm^{-1}.

[2] These difference spectra were obtained by combinatoric subtraction of computed spectra convolved with a Cauchy–Lorentz kernel of scale factor $\gamma = 25$ cm^{-1} (corresponding to a full width at half maximum (FWHM) of 50 cm^{-1}). The numerical integration of difference spectra was achieved through the method of trapezoids with a discretisation interval of 1 cm^{-1}.

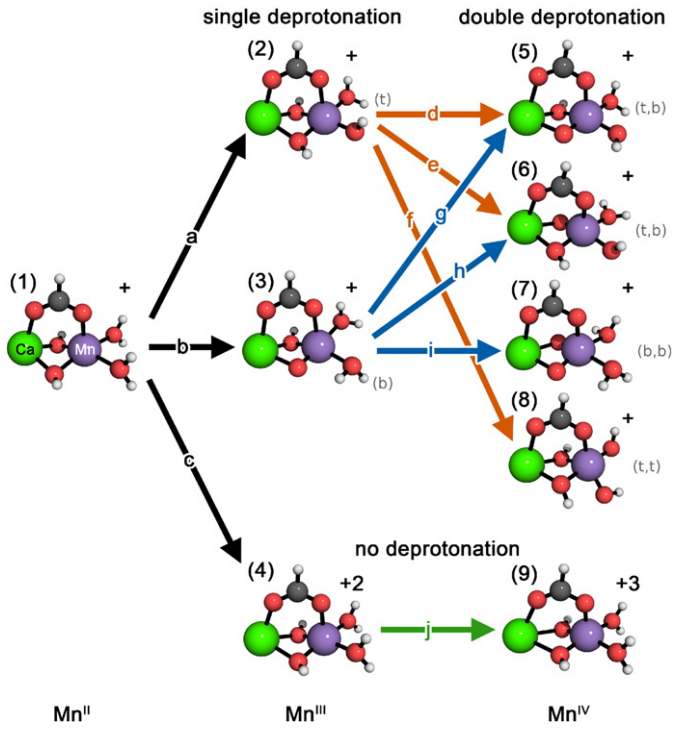


Fig. 3. Scheme representing connectivity of species under investigation and the oxidation/deprotonation pathways connecting these species. Terms in parentheses indicate the deprotonation pattern of each species: t = *terminal* (i.e. deprotonation of an *aqua* ligand), b = *bridging*.

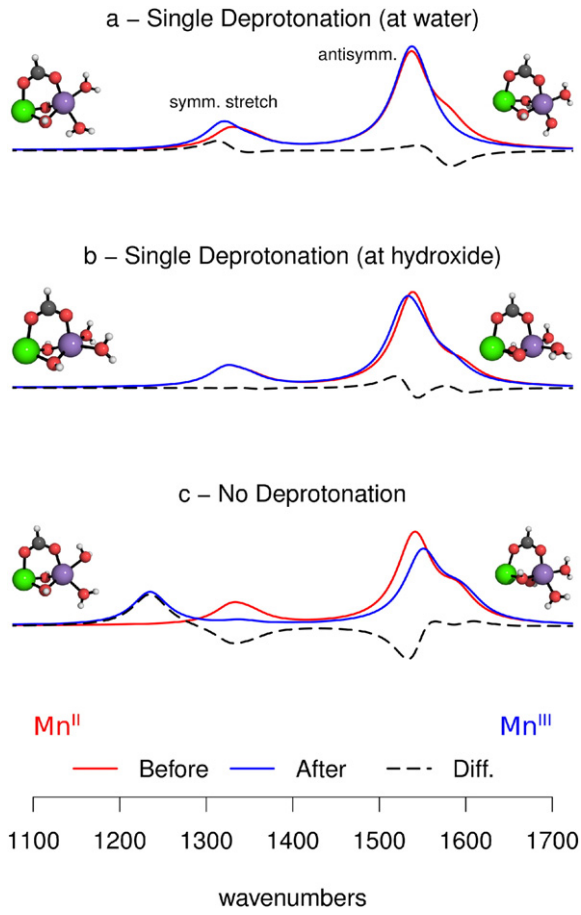


Fig. 4. (a,b) Difference spectra identified as possessing minimal difference integrals in the region 1100 cm^{-1} to 1700 cm^{-1} for transitions a and b, corresponding to a single oxidation of the manganese ion with deprotonation of either a terminal water or bridging hydroxide ligand. (c) Difference spectra … without deprotonation. See Fig. 2 for species involved in transitions a to c. The specific frequency range illustrated has been selected to correspond to that of the FTIR difference spectra of Debus and co-workers. The corresponding conformations for *before* and *after* states are depicted inset on the left and right hand sides of the figure respectively.

carboxylate symmetric stretch around 1200–1350 cm^{-1} and an antisymmetric stretch around 1500–1600 cm^{-1}. These positions are consistent with expectations for carboxylate-terminated ligands.

4. Discussion and conclusion

The IR spectra we have generated reveal an insensitivity to the specific order of deprotonation of water or hydroxide ligands, and the suppression of difference intensity is of the same character in both singly and doubly deprotonated complexes, although patterns f (double deprotonation of terminal waters) and h (*trans* double deprotonation) possess superior suppressions to other double deprotonation patterns. Contrarily, oxidation without deprotonation consistently generates negative shifts in the symmetric stretch peak position, as well as a positive shift in the antisymmetric stretch peak position, that is moderate in the case of single oxidation (Mn^{II} to Mn^{III}) but large in the case of a second oxidation (Mn^{III} to Mn^{IV}). Whilst deprotonation with concomitant oxidation generally yielded suppression of difference spectra with respect to uncompensated oxidation, it was found that deprotonation of both bridging hydroxide ligands (i) uniquely produced a large difference spectrum.

The calculations we have presented indicate a conditional suppression of FTIR difference intensity in the carboxylate stretch region of the modelled Asp170 group with concomitant oxidation and deprotonation of the complex. This is apparently not possible in oxidation-only scenarios. These data clearly support the cancellation effect hypothesis, indicating that there exist patterns of deprotonation and oxidation that yield no appreciable change in both stretch peak position and intensity.

Conformational exploration reveals a nontrivial dependence of the carboxylate symmetric and antisymmetric stretches on the Mn bound water and hydroxide ligands. Moreover, the modulation of modes associated with motions of water and hydroxide protons nearly coincident with the carboxylate antisymmetric stretch represents a source of difference intensity not directly related to the carboxylate ligand, but nevertheless expressed in the carboxylate antisymmetric stretch region.

Despite the simplicity of our model, it retains a close connection with the local environment of Asp170 in the OEC of PSII. This residue is bound to the 'dangler' Mn4 centre and is largely decoupled from the remaining carboxylate residues bound to metal centres forming the $CaMn_3$ cubane structure. Unfortunately, it is not possible to directly compare our calculated peak positions with those for the physiological Asp170 residue, due to the FTIR difference silence of the latter. However, we do calculate symmetric stretches that correspond well with the Ala344 symmetric stretch between 1320 cm^{-1} and 1356 cm^{-1} identified in S_1 and S_2-state PSII studies reported by Chu and co-workers [23]. Since the Ala344 residue also coordinates the Ca ion, it serves as a useful guide for the expected location of the Asp170 symmetric stretch. Chu and co-workers also nominate an antisymmetric stretch region of 1500–1640 cm^{-1} which is consistent with the range of antisymmetric stretch peaks we calculate.

Whilst these results have implications for understanding the absence of observable FTIR differences associated with Asp170 during *S* state turnover, some caution is required. There is significantly greater conformational flexibility afforded by the model system with respect to the environment surrounding the Mn(4) ion of the PSII OEC, as a Glu333 simulant coordinatively saturating the model Mn ion is absent. In practice, inclusion of a second carboxylate bridge in the model

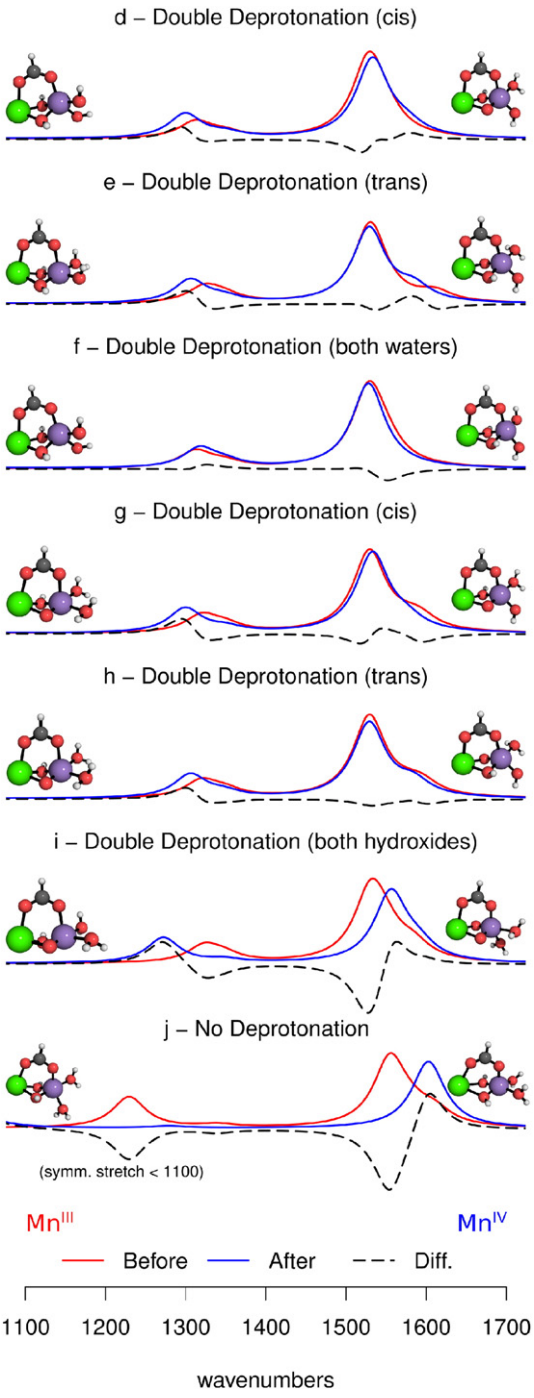


Fig. 5. (d–h) Difference spectra identified as possessing minimal difference integrals in the region 1100 cm^{-1} to 1700 cm^{-1} for oxidative transitions d to h, corresponding to a second oxidation either with or without deprotonation of a terminal water or bridging hydroxide ligand. (i) Difference spectra for transition i involving deprotonation of both bridging hydroxides. (j) Difference spectra without deprotonation. See Fig. 2 for species involved in transitions d to j. Note that the symmetric stretch for transition j has migrated below 1100 cm^{-1} (to 1035 cm^{-1}), and is therefore outside of the area of consideration of the experimental characterisation of Debus and co-workers.

complex does not result in consistent molecular connectivity. In future work, we intend to apply the above approach to larger, more realistic models of the $CaMn_4$ cluster in PSII based on the most recent XRD structures [3,4].

This research was undertaken with the assistance of resources from the National Computational Infrastructure (NCI), which is supported by the Australian Government.

References

[1] J.F. Kasting, Science 293 (2001) 819–820.
[2] H.D. Holland, Philos. Trans. R. Soc. B 361 (2006) 903–915.
[3] Y. Umena, K. Kawakami, J.-R. Shen, N. Kamiya, Nature 473 (2011) 55–60.
[4] M. Suga, F. Akita, K. Hirata, G. Ueno, H. Murakami, Y. Nakajima, T. Shimizu, K. Yamashita, M. Yamamoto, H. Ago, J.-R. Shen, Nature 517 (2015) 99–103.
[5] N. Cox, D.A. Pantazis, F. Neese, W. Lubitz, Acc. Chem. Res. 46 (2012) 1588–1596.
[6] D.R.J. Kolling, N. Cox, G.M. Ananyev, R.J. Pace, G.C. Dismukes, Biophys. J. 103 (2012) 313–322.
[7] M. Kusonoki, J. Photochem. Photobiol. B 104 (2011) 100–110.
[8] S. Petrie, R. Stranger, R.J. Pace, Chem. Eur. J. 14 (2008) 5482–5494.
[9] S. Petrie, R. Stranger, R.J. Pace, Chem. Eur. J. 21 (2015) 6780–6792.
[10] X. Li, P.E.M. Siegbahn, Phys. Chem. Chem. Phys. 17 (2015) 12168–12174.
[11] R.J. Debus, M.A. Strickler, L.M. Walker, W. Hillier, Biochemistry 44 (5) (2005) 1367–1374.
[12] R.J. Debus, Coord. Chem. Rev. 252 (3–4) (2008) 244–258.
[13] M.A. Strickler, L.M. Walker, W. Hillier, R.D. Britt, R.J. Debus, Biochemistry 46 (11) (2007) 3151–3160.
[14] W.Y. Chuah, R. Stranger, R.J. Pace, E. Krausz, T.J. Frankcombe, J. Phys. Chem. 118 (2014) 3553–3558.
[15] . ADF2013, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, 2013 (http://www.scm.com).
[16] C. Fonseca Guerra, J.G. Snijders, G. te Velde, E.J. Baerends, Theor. Chem. Accounts 99 (1998) 391–403.
[17] G. te Velde, F.M. Bickelhaupt, S.J.A. van Gisbergen, C. Fonseca Guerra, E.J. Baerends, J.G. Snijders, T. Ziegler, J. Comput. Chem. 22 (2001) 931–967.
[18] J.P. Perdew, Phys. Rev. B 33 (1986) 8822–8824.
[19] A.D. Becke, Phys. Rev. A 38 (1988) 3098–3100.
[20] E. van Lenthe, E.J. Baerends, J. Comput. Chem. 24 (2003) 1142–1156.
[21] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2015 (http://www.R-project.org).
[22] The PyMOL Molecular Graphics System Version 1.6, Schrödinger, LLC, 2015 (http://www.pymol.org).
[23] H.-A. Chu, W. Hillier, R.J. Debus, Biochemistry 43 (2004) 3152–3166.
[24] A. Bérces, R.M. Dickson, L. Fan, H. Jacobsen, D. Swerhone, T. Ziegler, Comput. Phys. Commun. 100 (1997) 247–262.
[25] H. Jacobsen, A. Bérces, D.P. Swerhone, T. Ziegler, Comput. Phys. Commun. 100 (1997) 263–276.
[26] S.K. Wolff, Int. J. Quant. Chem. 104 (2005) 645–659.

Introduction to Paper III: *What computational chemistry and magnetic resonance reveal concerning the oxygen evolving centre in Photosystem II* 6

Infinite dark rotation,

Magnetic isolation

Frank Michael Spinath

6.1 Citation

Terrett, R.; Petrie, S.; Stranger, R.; Pace, R.J.; *What computational chemistry and magnetic resonance reveal concerning the oxygen evolving centre in Photosystem II*; Journal of Inorganic Biochemistry 162 (**2016**) pp. 178–189; doi: 10.1016/j.jinorgbio.2016.04.009

6.2 Introduction

Due to the presence of four manganese ions in close proximity (both in space and connectivity on the molecular graph), PSII has the potential for rich EPR structure, with considerable diagnostic value in terms of mapping topological and magnetic changes in the PSII OEC over the course of the S-state cycle. Prior to the availability of high-resolution X-ray structures unambiguously furnishing the OEC cluster geometry, rudimentary characterisation of the relative geometry of Mn ions within the cluster was achieved through EPR. [36]

In this paper, we present calculations of the J_{ij} magnetic coupling parameters of the Mn ions of the OEC within the Heisenberg–Dirac–van-Vleck (HDvV) Hamiltonian (Equation 6.1) formalism.[a] These data are then coupled to EPR nuclear hyperfine calculations.

[a] *n.b.* that the HDvV Hamiltonian has a variety of definitions distinguished by a multiplicative constant. The convention adopted by our publication eschews the factor of −2 seen in other works (*e.g.* Retegan, *et al.* [226])

$$\hat{H} = -J_{i \neq j} \sum_{ij} \hat{S}_i \cdot \hat{S}_j \tag{6.1}$$

This paper presents simulated anisotropic nuclear hyperfine tensors for PSII OEC models (**A**-tensors). The nuclear hyperfine interaction arises from the interaction of unpaired electron spin density with the nuclear magnetic dipole, where nuclei possess a nonzero nuclear magnetic moment (^{55}Mn possesses a nuclear magnetic moment tensor of $\mathbf{I} = \frac{5}{2}$ [227]). The OEC has the convenient property that ^{55}Mn has a 100% natural abundance. [228] The total nuclear hyperfine interaction consists of two components: an *isotropic* component arising from the Fermi contact interaction[b] and an *anisotropic* component, **A** which arises from coupling between the orbitally distributed electron magnetic moment and the nuclear magnetic dipole moment. [229] The individual components of the nuclear hyperfine anisotropy may be resolved in ordered systems, whereas disordered phases present a *'powder pattern'* response.

The orientation and aspect ratio of the nuclear hyperfine tensors can provide diagnostic information regarding the local geometry and ligation of transition metal ions, their oxidation state, and the nature of superexchange interactions coupling these ions. This information is salient to

[b]To wit, the magnetic interaction of an unpaired electron where the electron is located inside the atomic nucleus. This phenomenon can only arise in *s* orbitals, due to orbitals with $l \geq 1$ possessing a node of zero expectation density at the nucleus, however in practice the partial *s* character of transition metal valence electrons gives rise to an isotropic component.

structural investigation of the PSII OEC. In contrast to computational modeling of molecular geometry and bonding energy, simulation of nuclear hyperfine interactions requires particular attention to modeling of core electron density, which make effective core potential-like approaches such as the frozen core basis sets offered by the ADF software inappropriate. Moreover, whereas relativistic effects are of limited relevance to the valence orbital structure of light transition metals such as Mn, core orbital contractions were considered to be important in the context of nuclear hyperfine structure calculations, accounting for the application of the Zeroth Order Regular Approximation (ZORA) scalar relativistic correction to nuclear hyperfine calculations. [230] Whilst computational chemistry suites typically model the atomic nucleus as a point-like entity, the finite radius of real atomic nuclei is cited as a factor that becomes significant to the nuclear hyperfine structure of heavier atoms. [231] Nevertheless, benchmarking comparing point-like and Gaussian models of nuclear structure did not reveal a measurable effect on our ^{63}Cu model **A**-tensor components.

6.3 My contribution to this work

6.3.1 Preliminary benchmarking

I performed exploratory EPR benchmarking in ADF at a variety of levels of theory on a collection of Cu^{II}DETA imidazole complexes, to gain intuition as to the behaviour and stability of EPR calculations in ADF. This step included comparison of LDA, GGA and hybrid functionals, basis sets composition, finite or point-like nuclear models, and spin-orbit or perturbative nuclear hyperfine interaction models.

6.3.2 Spin-ladder preparation

I prepared diabatic and adiabatic spin-ladders based on optimisation of electronic structures and geometries of pre-existing molecular models.[c] Not all spin states could be obtained for each S-state, however the complete spin ladder over-determines the J_{ij} couplings for the OEC and therefore a subset of these spin states was sufficient. Ultimately, only diabatic spin ladders were reported, however adiabatic spin ladders were produced as a side-effect of identifying the attendant geometry of the ground state spin configuration.

Briefly, if one considers a collection of n interacting transition metal

[c]Crystallographically-adapted models of PSII in putative S_1 through S_3 states were furnished by Dr Simon Petrie.

centres with locally ferromagnetic spin density (*i.e.* one in which all unpaired *d* electrons possess either α or β polarisation, a *single-determinant* configuration), there exist 2^{n-1} unique spin configurations, taking into account spin-symmetry in the absence of an external field (an example is given in Table 6.1, for four identical Mn ions, generating 8 configurations). The precise expansion of the HDvV Hamiltonian depends on the particular topology of the cluster, and in highly symmetric cases degenerate J_{ij} couplings may result (*e.g.* Schmitt, *et al.* [232]). Magnetic coupling between ions such as seen in the OEC or in related models such as the $Mn_4ac_3Cl_7$ complex of Schmitt, *et al.* occurs through the mechanism of Kramers–Anderson superexchange, where magnetic communication between metals is effected by the spin-polarisation of intervenient ligands (typically μ^n-coordinated anions). [39]

Table 6.1: Generic spin ladder of four identical manganese ions. Note that the individual ferromagnetic and anti-ferromagnetic relations between ions are maintained under the inversion of spin, producing $(n-1)^2$ distinct configurations.

Mn_1	Mn_2	Mn_3	Mn_4	Comment
↑	↑	↑	↑	Ferromagnetic
↑	↑	↑	↓	
↑	↑	↓	↑	
↑	↑	↓	↓	Anti-Ferromagnetic
↑	↓	↑	↑	
↑	↓	↑	↓	Anti-Ferromagnetic
↑	↓	↓	↑	Anti-Ferromagnetic
↑	↓	↓	↓	

6.3.3 EPR calculation

I performed EPR calculations on S-state models at a variety of levels of theory. These calculations were used to obtain anisotropic nuclear hyperfine tensors for analysis.

6.3.4 Data postprocessing and tensor visualisation

I developed a methodology to visualise ADF EPR tensor data in ParaView [233], [234] and implemented code to produce the requisite Visualisation Toolkit (VTK)-compatible tensor visualisation data from ADF output, as well as writing associated parsers to extract tensor information from ADF output. Visualisation involved fusing molecular geometry output with tensor data corresponding to selected atoms, and was implemented in Python. I chose ParaView for visualisation as it can display generic rank-3 tensor glyphs situated in $\mathbb{R}^3$ using either cuboid, ellipsoid or a superquadric surfaces.[d] I considered none of these visualisation techniques to be appropriate as we required a visualisation that clearly displayed tensor components without occluding atoms or bonds. Therefore our visualisation consisted of representing the principle axes of the tensors as rods, generated by triplicating a cuboidal representation of the tensors and scaling the aspect ratios of the three representations to $\{1.0, 0.1, 0.1\}$,

[d] The use of superquadrics is a fairly recent innovation in tensor field representation, principally intended to reduce visual ambiguity of diffusion tensor fields in biomedical imaging with respect to ellipsoidal representations. [235]

{0.1, 1.0, 0.1} and {0.1, 0.1, 1.0}. The methodology is illustrated in Figure 6.1 with respect to conventional tensor representations. In some instances highly anisotropic tensors resulted in components that were submerged within manganese ions; in these cases approximate orientations were indicated by hand.

6.3.5 Graphic preparation

I prepared all molecular visualisations, EPR tensor visualisations, and J_{ij} coupling pattern illustrations in the manuscript, not including EPR spectra. Molecular visualisations were produced in PyMOL and Blender, whilst EPR tensor visualisation was achieved through fusion of ParaView tensor geometry and Blender data.

6.3.6 Contributions to manuscript

I contributed the methodology section and data germane to the aforementioned EPR and spin-ladder calculations to the manuscript and supporting information. Additionally, I edited the manuscript for clarity and consistency and assisted in implementing reviewer recommendations.

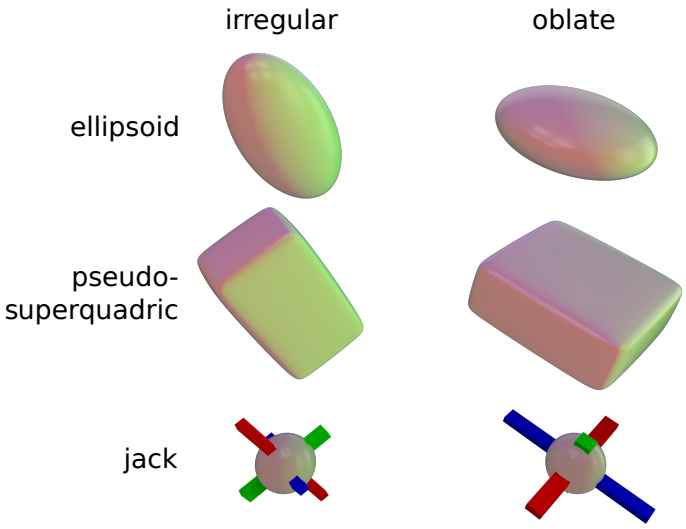


FIGURE 6.1: Representation of three tensor glyph schemes. Ellipsoidal and superquadric glyphs obscure underlying molecular geometry, whereas the *'jack'* representation maintains legibility.

Journal of Inorganic Biochemistry 162 (2016) 178–189

Contents lists available at ScienceDirect

Journal of Inorganic Biochemistry

journal homepage: www.elsevier.com/locate/jinorgbio

What computational chemistry and magnetic resonance reveal concerning the oxygen evolving centre in Photosystem II☆

Richard Terrett, Simon Petrie, Rob Stranger, Ron J. Pace *

Research School of Chemistry, College of Physical and Mathematical Sciences, Australian National University, Canberra, ACT 0200, Australia

ARTICLE INFO

Article history:
Received 10 November 2015
Received in revised form 1 March 2016
Accepted 4 April 2016
Available online 8 April 2016

Keywords:
Oxygen Evolving Complex
Photosystem II
Water oxidation
Electron Paramagnetic Resonance
Hyperfine coupling
Computational chemistry

ABSTRACT

Density Functional Theory (DFT) computational studies of the Mn_4/Ca Oxygen Evolving Complex (OEC) region of Photosystem II in the paramagnetic S_2 and S_3 states of the water oxdizing catalytic cycle are described. These build upon recent advances in computationally understanding the detailed S_1 state OEC geometries, revealed by the recent high resolution Photosystem II crystal structures of Shen et al., at 1.90 Å and 1.95 Å (Petrie et al., 2015, *Angew. Chem. Int. Ed.*, 54, 7120). The models feature a 'Low Oxidation Paradigm' assumption for the mean Mn oxidation states in the functional enzyme, with the mean oxidation levels being 3.0, 3.25 and 3.5 in S_1, S_2 and S_3, respectively. These calculations are used to infer magnetic exchange interactions within the coupled OEC cluster, particularly in the Electron Paramagnetic Resonance (EPR)-visible S_2 and S_3 states. Detailed computational estimates of the intrinsic magnitudes and molecular orientations of the ^{55}Mn hyperfine tensors in the S_2 state are presented. These parameters, together with the resultant spin projected hyperfine values are compared with recent appropriate experimental EPR data (Continuous Wave (CW), Electron-Nuclear Double Resonance (ENDOR) and ELDOR (Electron-Electron Double Resonance)-Detected Nuclear Magnetic Resonance (EDNMR)) from the OEC. It is found that an effective Coupled Dimer magnetic organization of the four Mn in the OEC cluster in the S_2 and S_3 states is able to quantitatively rationalize the observed ^{55}Mn hyperfine data. This is consistent with structures we propose to represent the likely state of the OEC in the catalytically active form of the enzyme.

1. Introduction

Water is the reducing substrate for production of molecular oxygen in Photosystem II (PSII), in a process whose mechanistic detail remains as yet unresolved. It is known that during this cyclic four-electron reaction, the Mn_4Ca cluster in the PSII water-oxidising complex (WOC, also known as the Oxygen Evolving Centre, OEC) moves through a series of four metastable, spectroscopically accessible intermediates and one intrinsically short-lived unstable state, collectively known as the S states. Electron transfer from the cluster to the photo-oxidized P680 chlorophyll reaction centre is by means of a redox active tyrosine residue, Y_Z. The metastable states are labelled S_0, S_1, S_2 and S_3, and the short-lived 'final' state is S_4, where the subscript refers to the number of stored oxidising equivalents in the catalytic centre [1]. However, the precise, catalytically relevant geometry of the Mn_4Ca site, as well as the overall metal centre oxidation level at which water oxidation occurs, are still the subjects of debate [2]. Atomic level detail of the nominally 'dark stable' S_1 state of the Mn_4Ca catalytic site has now been revealed by PSII structures at 1.9 Å [3a] and 2.1 Å resolution for the (still functional) Sr/Ca-substituted enzyme [3b] and very recently at 1.95 Å for the native enzyme, with very low X-ray exposure [3c]. These advance earlier studies at lower resolution (2.9 Å, 3.0 Å and 3.5 Å) [3d,e,f] and supply detail of oxo bridging and possibly substrate water binding sites within the cluster. At present, these structures should be conservatively regarded as nominally S_1, as the possibility of some contribution from more reduced forms (likely S_0) cannot be definitely excluded.

The 1.9–2.1 Å resolution PSII structures reveal the disposition of Mn, Ca/Sr and terminal/bridging O species, along with protein supplied ligating side chains. In each case there is a distinguishable difference between the precise OEC metal positions within the two monomers (labelled A and B) of the homo-dimeric enzyme as crystallized. In the 1.9 and 2.1 Å structures, the closest Mn–Mn separations are generally ≈2.8–2.9 Å (or more), rather than the 2.7–2.8 Å, seen consistently for the two closest Mn–Mn separations resolved in EXAFS (Extended X-ray Absorption Fine Structure) [4]. However, in both the Ca and Sr structures, monomer B has a shortest Mn–Mn distance of 2.76 Å, similar to that seen in EXAFS. Nonetheless, this has led to speculation that the cluster has been perturbed by photo-reduction during the XRD data collection [5]. However, we have now shown [6] by large scale computational modelling that the structural forms seen by Shen et al. are

☆ Dedication: The authors dedicate this publication to the memory of a dear friend and colleague, Graeme Hanson at University of Queensland, who passed away recently, after tragic illness. Graeme was an individual of very high standing in the international EPR community, who has made lasting contributions to the field, particularly in biological EPR. It has been our great privilege and pleasure to have known and worked with him, particularly on Photosystem II.

* Corresponding author.
E-mail address: Ron.Pace@anu.edu.au (R.J. Pace).

http://dx.doi.org/10.1016/j.jinorgbio.2016.04.009
0162-0134/

simple proton displaced tautomers of alternative S_1 configurations. One is consistent with the 1.95 Å structure and lower resolution structure of Guskov et al. [3e] with 'short' EXAFS, Mn–Mn distances, while the other shows 'long' Mn–Mn distances, as seen in the 1.9 Å and 2.1 Å structures. This is possible only because we assign a particular oxidation state level (low oxidation paradigm, see below) to the Mn ions within the OEC and the proton displacement triggers a re-disposition of the individual Mn oxidation states, while maintaining the same mean Mn oxidation level of 3.0. The same redox assignment quantitatively rationalizes the monomer A and B differences in the 1.9 and 2.1 Å structures, as a consequence of a subtle symmetry breaking that alters the Jahn–Teller axis alignment on one particular Mn^{III} ion in the cluster [6b].

Two possibilities for the redox levels of the Mn ions within the functional OEC cluster have been proposed, the so called 'low' and 'high' paradigms [2]. These differ consistently by two oxidation equivalents throughout the S cycle, with the mean oxidation levels in S_1 being 3.0 and 3.5 in the low and high series, respectively. The high paradigm has been supported by particular interpretations of Mn XANES measurements ([4], but see also [7]) and computational modelling of the OEC (extensive, but see [5,8] for representative examples), while the low paradigm has been supported by alternative analyses of XANES and related spectroscopic data [9], as well as a recent direct determination by electron counting during photo assembly of the functional OEC, from *apo*-PSII and exogenous Mn^{II} [10].

The only high resolution XRD structures of PSII currently available are for the S_1 state, although two groups [11] have very recently presented data from flash advanced PSII microcrystals (extensively signal averaged), in the nominal S_2 and S_3 states. These are at present only to 5–6 Å resolution. However, Mn EXAFS results from several groups [4] indicate that there is less structural change on the $S_1 \rightarrow S_2$ transition than for any other transition in the S cycle. Both states have a very similar pair of 'short' (≈2.7 Å) Mn–Mn vectors, with any differences being in the region of Mn–metal (Mn or Ca) distances of ≈3.0 Å or larger. Both S_0 and S_2 have an odd number of electrons and exhibit well known EPR signals at low temperature [12]. The most studied of these is the S_2 Mn hyperfine structured spin ½ ground state, known as the multiline signal (ML), although other signals of spin ≥5/2 may also form as ground state (or near ground state) species or as cryo-trapped forms generated by low temperature NIR (Near Infrared) illumination of the ML state [13]. From ^{55}Mn ENDOR (Electron-Nuclear Double Resonance) studies on the ML and modelling by several groups [14], it had been proposed that all four Mn in the exchange coupled cluster contribute near equivalently to the total hyperfine structured signal. However, we have recently shown by ^{55}Mn ENDOR studies on the S_2 ML signal performed at temperatures significantly lower (2.5 K) than any employed previously [15], that the relative net contributions of individual Mn to the total magnetic state of the ML form are quite different from those earlier suggested. Two Mn ions are virtually magnetically 'silent' in the cluster with small effective Mn hyperfine couplings (one essentially zero). This requires that the overall cluster coupling be dominantly antiferromagnetic, but the Mn ions are in effect magnetically organized as two dimers, with one dimer consisting of a Mn^{III}–Mn^{IV} pair (net spin ½) and the other as a Mn^{III}–Mn^{III} pair (low paradigm) or Mn^{IV}–Mn^{IV} pair (high paradigm, both cases net spin near 0). This is despite the *topological* arrangement of the OEC Mn being close to a 'trimer plus one' ('dangler' [6], or 'funnel' arrangement [12a]). These 'dimer' pairs are relatively strongly anti-ferromagnetically coupled within themselves, but with only weak effective (net) coupling between them (ferro- or antiferromagnetic), at least in the S = ½ ML state. Interestingly, such a magnetic organization is actually predicted for the OEC by very recent high level 'quantum entanglement' calculations [16], with computed effective hyperfine contributions from two Mn being small (<20% of nominal, isolated ion values).

EXAFS studies reveal [4] that there is a significant change in geometry of the OEC metal positions in the S_3 state compared with the lower S states. The detailed interpretation of this change differs somewhat between experimental groups. Yachandra et al. found expansion of one 'close' ≈ 2.7 Å Mn–Mn vector (as seen prominently in the lower S states), to ≈2.9–3.0 Å [4a], while Dau et al. saw an additional close vector of ≈2.8 Å form [4c] (both studies from higher plant derived PSII). A recent study on cyanobacterial PSII by the Berkeley group [4f], found results similar to that of Dau et al. The 'expansion' of the close distance is superficially surprising as the cluster undergoes single electron oxidation and proton loss on the $S_2 \rightarrow S_3$ transition. Yachandra et al. [4a] have interpreted this to mean that oxidation of a bridging oxo had occurred rather than a metal centre. This is compatible with the small Mn XANES edge shift these authors see on the $S_2 \rightarrow S_3$ transition. However, this interpretation is not favoured by Dau et al. experimentally [4d], nor by any group who has studied the matter with computational chemistry, all of whom see Mn oxidation on passage to S_3 (*e.g.* see discussion in [2]).

The S_1 and S_3 states are integer spin. A broad, featureless parallel polarized derivative signal at g ≈ 4.8, with ≈60 mT peak–trough width, is sometimes seen in S_1 (see [12b]). As yet no ^{55}Mn hyperfine has been resolved from this state. However, Cox et al. have recently reported high-field (Q, W band) pulsed EPR on the S_3 state in thermophilic PSII core complexes [17]. The S_3 state is heterogeneous in these systems, one population (≈60%) exhibits a high-spin, S = 3 magnetic state while the other population is EPR-silent. W Band EDNMR (ELDOR-Detected NMR) on the EPR-visible state yields several well-resolved ^{55}Mn nuclear couplings.

Here we examine the detailed structures of the S_2 and S_3 states in PSII, building on our recent modelling of the S_1 state of the photosystem [6] within the low oxidation state paradigm. The consequences of this new modelling for magnetic interactions within these higher S states are examined, in particular, for exchange interactions between the Mn ions in the OEC cluster and the individual ^{55}Mn ion hyperfine interaction parameters seen in S_2 and S_3 experimentally.

2. Methods

All density functional calculations were performed in the Amsterdam Density Functional suite (ADF), version 2014.01 on 64-bit Linux [18a,b,c].

2.1. Geometry optimization

Geometry optimizations were performed on crystallographically adapted models of the Photosystem II OEC active site in the S_2 and S_3 oxidation states possessing expected average Mn oxidation states of +3.25 and +3.5, respectively. Two initial geometries were investigated in these states, distinguished by the relative arrangement of the Asp170 group and associated water ligands in the structure and designated '*open*' and '*closed*'. The crystal structure used to derive optimized geometries is the 2.9 Å X-ray structure from *Thermosynechococcus elongatus* (Guskov et al., '*Berlin*' structure [3e]). This structure has the open form, which we believe (see below) more closely resembles the functional state of the OEC. The structure used was that of OEC monomer 'B' from the RCSB Protein Databank. Derived structures consist of the Mn_4CaO_5 decorated hemi-cubane cofactor with coordinating propanoate and 5-ethylimidazole ligands representing amino acids D1–Asp170, D1–Glu189, D1–His332, D1–Glu333, D1–Asp342, D1–Ala344 and CP43–Glu354.

Optimized geometries were calculated using the default convergence criteria. These comprise a ΔE threshold of 1×10^{-3} E_h, with maximum and root-mean-square gradient criteria of 1×10^{-3} E_h/Å and 6.67×10^{-4} E_h/Å, respectively, and a maximum and root-mean-square Cartesian step size of 1×10^{-2} and 6.67×10^{-3} Å, respectively. Geometry optimization was performed in a Cartesian coordinate system. In all geometry optimizations, the terminal methyl carbons of 5-ethylimidazole and propanoate ligands representing coordinating amino acids were pinned at the XRD structure relative positions. Geometry optimizations were performed in the generalized gradient

approximation (GGA) using the Becke–Perdew '86 (BP86) combined exchange–correlation functional [19a,b]. The calculations were performed in the spin-unrestricted Kohn–Sham formalism, in which the α and β spin symmetry of the molecule is broken. Each species was spin-polarized to obtain the 8 symmetry-unique single-determinant coupling configurations between Mn ions 1–4. The TZP (Triple Zeta, one Polarization function) all-electron Slater basis set was used without relativistic correction, owing to the absence of heavy atoms [19c]. Calculations employed the default numerical integration accuracy exponent of 4, and SCF convergence was achieved to the satisfaction of the default convergence criterion.

2.2. Single-point spin ladders

Non-adiabatic spin ladders were generated by single-point calculations on the optimized geometry of the minimum energy broken-symmetry state. Where complete suites of optimized geometries for all spin states were unavailable, the optimized geometry of the ABAB spin coupling pattern (A = 'spin up', α; B = 'spin down', β) was instead used. These spin ladders are produced by single-point calculations of the electronic energy at a fixed geometry under varying constraints of total molecular spin and Mn ion spin-spin coupling.

2.3. A-tensor determination

Hyperfine interactions were modelled using the ESR (Electron Spin Resonance) directive within the ADF2014.01 suite [20] using the BP86/TZP-ZORA and PBE0/TZP-ZORA functional/basis set combinations. This produces *A*-tensors in the molecular Cartesian and atomic principal axis frames, however eigenvalues and eigenvectors of the *A*-tensors were produced manually by eigen-decomposition using the *linalg* module of the Numerical Python (NumPy) package [21]. ESR calculations also employed the Zeroth Order Relativistic Approximation (ZORA) [22] in computing *A*-tensor components. Isotopic defaults of ^{55}Mn were retained in ESR calculations.

3. Results

3.1. Structural models

The structural models used for the S_2 and S_3 states derive from our recent computational studies of the S_1 state, as revealed by the XRD data at 2.9 Å, 1.9 Å and 1.95 Å [6]. The computational structures, shown in Figs. 1, 2 are hydrated (6 waters) and fully model all protein supplied metal ligating amino acid side chains, with the β carbons 'pinned' at relative spatial positions consistent with the 2.9 Å XRD structure and modelled as methyl groups (as in [6]). Our S_1 state studies revealed two basic structural types, an 'open' form, generally consistent with the 2.9 Å structure and a 'closed' form, seen in both the 1.9 and 1.95 Å structural types. These differ principally in that the Asp170 carboxylate side chain is ligated only to Mn4 (so called 'dangler' Mn, see [6]) in the open form (Fig. 1a), while Asp170 bridges between Mn4 and Ca in the closed form (Fig. 1b). The imidazole side chain of His337 is definitely not H-bonded to the μ_3 oxo bridging Mn1, 2, 3 in the open form, but is so bonded in the closed variants in the 1.9 Å and 1.95 Å XRD structures [6]. Also, in these forms the O5 moiety (water or hydroxide) is relatively distant from Mn1 (>2.6 Å, XRD and computational), while computationally it is ≈2.3 Å in the open form (along the Mn1 Jahn–Teller axis) in S_1 and S_2. We believe this closer interaction to be more consistent with the relatively strong ^{17}O hyperfine interaction seen in S_2 by Rapatskiy et al. [8l] from $H_2^{17}O$-labelled solvent introduced by fast exchange in S_1.

Fig. 1a, b show the S_2 state computational models employed here, for the open and closed forms. Both have the low paradigm Mn mean oxidation state level of 3.25, consistently giving a redox pattern of III,IV,III,III for the S_2 state (for Mn1···4, [9]). The μ_3 oxo group is assumed to be un-protonated and His337 is not included. This is the form (with solvent water access to Mn1) that we regard most likely to actually exist in the functional OEC [6,9]. Both structures are consistent with the EXAFS data, with two close ≈2.7–2.8 Å Mn–Mn distances [4]. The W2 water on Mn4 is deprotonated in our S_2 models, which are charge neutral, like S_1. However, this proton is not seen to be released externally on the $S_1 \rightarrow S_2$ transition and is likely trapped internally (as seen by Fourier transform IR spectroscopy [23]). The Mn3···4 distance is ≈3.3 Å in both cases.

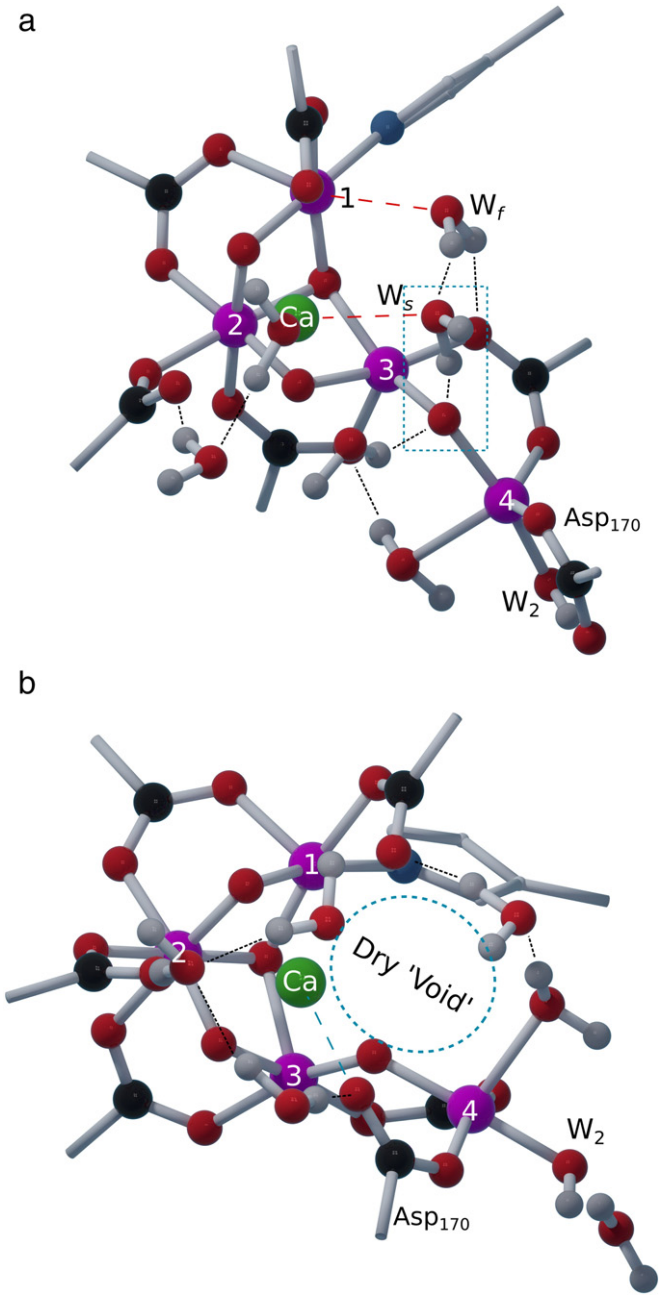


Fig. 1. a) Computational model of the OEC S_2 state in 'open' form with carboxylate of Asp170 pointing away from Ca (as in some XRD structures, see text). All structures in Figs. 1 and 2 have protein supplied ligand β carbons modelled as methyl groups, fixed at their relative positions in the 2.9 Å XRD structure (see Methods). Mn oxidation state pattern is III, IV, III, III. Proposed substrate species starred. One water binds to Mn1 along its Jahn–Teller axis ('fast' water, *Wf*). 'Slow' substrate water (*Ws*) is bound near Ca and Mn3 (see [6]). The W2 water on Mn4 is mono-deprotonated. Carboxylate and ethylimidazole ligands visually truncated for clarity in Figs. 1, 2, with selected hydrogen bonding patterns indicated by black dashed lines. b) Computational model of the OEC S_2 state in 'closed' form with the carboxylate of Asp170 now bridging between Mn4 and Ca (as in some XRD structures, see text). Mn oxidation state pattern is III, IV, III, III. In this form no water/hydroxide groups approach within ≈4 Å of Mn1.

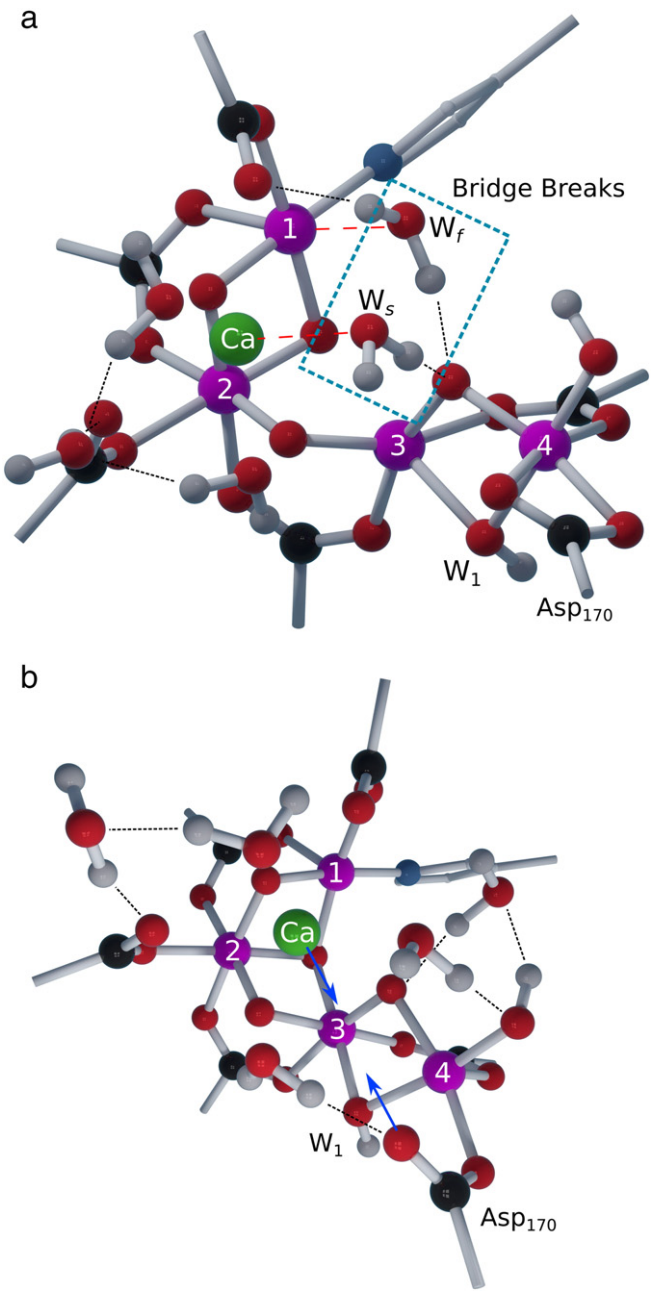


Fig. 2. a) Computational model of the OEC S_3 state in 'open' form. Carboxylate of Asp170 is bi-dentate ligand to Mn4. Mn oxidation state pattern is III, IV, III, IV. The substrate water molecules (*Wf,s*, Fig. 1a) are in close proximity and H bonded to the same oxo bridge between Mn3, 4. The W1 water has lost one proton and now bridges Mn3, 4. The O3 μ_3-oxo bridge between Mn1, 2, 3 is now a μ_2 bridge between Mn1, 2, allowing the Mn2, 3 vector to elongate (≈3.0 Å). This and the W1 bridge between Mn3, 4 ensures 'Coupled Dimer'-like magnetic organization for the cluster (see text). b) Computational model of the OEC S_3 state in 'closed' form. Carboxylate of Asp170 now a mono-dentate to Mn4, with the pendant carboxylate oxygen pointing towards Ca. Mn oxidation state pattern is III, IV, IV, III. W1 bridges Mn3, 4, as in Fig. 2a, but O3 is a μ_3-oxo bridging Mn1, 2, 3 and the magnetic ground state of the cluster is diamagnetic.

No high resolution XRD data are available for the S_3 state. However, low–medium (≈4.5–5.5 Å) resolution structures of the S_2, S_3 and S_0 states in flash advanced, femtosecond XRD studies on PSII microcrystals have become recently available from two groups [11]. The data are currently at the limits of interpretability due to the resolution, however a consistently discernible difference in the OEC cluster geometry is seen in S_3, compared to the lower S states. Further, as noted above, there is a significant, Mn EXAFS detectable change in the OEC cluster geometry on going from the earlier S states to S_3. This affects mostly the two 'short' (≈2.7 Å) Mn–Mn vectors seen in S_1 and S_2, either lengthening one of them, or adding a new short distance (≈2.8 Å), or both [4]. Further, detailed analysis of the S_2 and S_3 Mn XANES edge shapes [4d] suggests that a $Mn^{III}L_5 \rightarrow Mn^{IV}L_6$ change (L = ligand) occurs on this S state transition, which would be totally consistent with formation of a new oxo or hydroxo bridge between adjacent Mn (presumably leading to a shortening of a Mn–Mn vector).

The possibility of open and closed forms of the OEC, as seen above for S_2, mediated by the flexibility of the Asp170 ligation, provides a ready, quantitative explanation of these observations for S_3. This is illustrated in Fig. 2b. In the open form (Fig. 2), the low paradigm redox pattern is III, IV, III, IV and now the O3 oxo bridges only Mn1, 2. A proton is lost from the W1 water and the W2 hydroxide now bridges to Mn3, *i.e.* the Mn3, 4 pair is now oxo/hydroxo bridged. This shortens the Mn3⋯4 distance to ≈2.9 Å, while the Mn1⋯2 distance remains at ≈2.7 Å and the Mn2⋯3 distance increases to ≈2.95 Å, due to the loss of O3 bridging between Mn1, 3. The O5 water again has direct access to the Jahn–Teller site on Mn1 and the Asp170 carboxylate is now bidentate to Mn4. In the closed form (Fig. 2b), the Asp170 bridges Mn4 and Ca, ensuring the O3 group remains as a μ_3 oxo bridging Mn1, 2, 3, retaining two 'short' ≈ 2.7 Å Mn–Mn vectors. Mn3, 4 is again oxo/hydroxo bridged, with the Mn–Mn distance shortening to ≈2.9 Å. The O5 water now has only distant access (>3 Å) to the Jahn–Teller site on Mn1. In total these results are in excellent, quantitative agreement with the range of possibilities seen in EXAFS for the S_3 state.

3.2. Magnetic interactions

It is convenient to express the magnetic interactions between the Mn ions in the OEC in terms of an effective, phenomenological exchange Hamiltonian. From the range of low paradigm structures in both S_2 and S_3, Mn4 interacts only with Mn3 to excellent approximation. Then the exchange interaction pattern is as indicated in Fig. 3, with Hamiltonian $\mathcal{H}$, given by:

$$\mathcal{H} = -[J_{12}\,\mathbf{S}_1 \cdot \mathbf{S}_2 + J_{13}\,\mathbf{S}_1 \cdot \mathbf{S}_3 + J_{23}\,\mathbf{S}_2 \cdot \mathbf{S}_3 + J_{34}\,\mathbf{S}_3 \cdot \mathbf{S}_4]. \quad (1)$$

In Eq. 1, positive *J* indicates ferromagnetic coupling, negative indicates *anti*-ferromagnetic coupling. The single determinant DFT calculations performed here assume any centre with unpaired spins (*e.g.* a Mn ion), has all spins fully polarized locally, as α (A) or β (B). So a state designated AAAA has the spins on Mn1⋯4 all polarized α (*i.e.* 'up') *etc.* Formally, a single determinant of *e.g.* AAAA configuration is equivalent, in spin and relative energy terms, to the product electronic spin function, $|M_{S1}(\alpha), M_{S2}(\alpha), M_{S3}(\alpha), M_{S4}(\alpha)>$, where for a Mn^{III} centre say, with 4 unpaired d electrons, $M_S(\alpha) = +2$ and $M_S(\beta) = -2$ *etc.* Then the *relative total energies* of the different spin polarized DFT determinants (ABAB, *etc.*) may be related to the *expectation values* of Eq. 1, with the corresponding product spin functions [24]. If DFT energies for

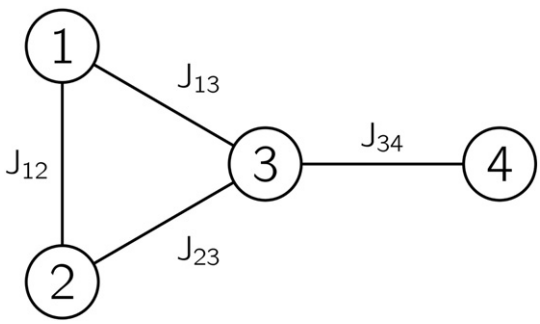


Fig. 3. Schematic depiction of magnetic coupling interactions between Mn ions of OEC active site.

sufficient, distinct spin polarization states can be obtained, then the *J* values in Eq. 1 are determined, or over-determined. In practice, one obtains first the configuration of lowest energy (*e.g.* ABAB) and then determines the energies of other spin configuration states, whose total geometries are fixed at that of the lowest energy (reference) state. This does not always lead to converged energies of particular 'excited' state configurations, but in the present instances, sufficient results were obtained to at least minimally determine the *J* values, listed in Table 1 (see Supplementary material § S1, 2). For the S_2 structures, the minimum energy states for both the open and closed forms were both dominantly *anti*-ferromagnetic, with net cluster spin $S_T = ½$ and configurations ABAB and ABBA, respectively. For the S_3 structures, the closed form was generally anti-ferromagnetically coupled, with $S_T = 0$. For the open form, configurations with $S_T = 0$ or 3 were admitted, within the likely uncertainties of the DFT energy analyses. Zero spin states are of course EPR inactive and will not be considered in detail here.

3.3. S_2 state

The S_2 intermediate, with odd total electron spin, is known to exhibit a range of Kramers state EPR signals [12,13]. These appear centred at apparent g values of ≈2.0, 4.4, 8–9 and correspond to total net spin states of $S_T = ½$, 3/2 or 5/2 and possibly higher (≈7/2). By far the most studied and information rich is the $S_T = ½$, so called Multi-Line state (ML) [12]. It is this species which we concentrate on here, as it is the only S_2 form from which Mn hyperfine interactions have been clearly resolved. The intrinsic anisotropy of the ^{55}Mn hyperfine tensor for a Mn^{III} ion (d^4) differs from those of the closed d sub-shell species, Mn^{II} (d^5) and Mn^{IV} (d^3). Illumination of PSII in the dark stable S_1 state at ≈200 K (*e.g.* [12]) produces the hyperfine structured S_2 ML signal. Typical examples are given in Fig. 4a, b (see [15] for details). The spectrum contains 18–20 resolved hyperfine lines, centred close to g = 2.00 and is at least 180 mT (*i.e.* ≈5000 MHz) wide at X-band (one resonance obscured by the Tyr_D radical signal near g = 2.00 [12]). The signal resembles the 'classic' 16 line EPR signals from anti-ferromagnetically coupled $Mn^{III\text{-}IV}$ oxo bridged dimers below ≈20 K [25a–c], although these signals (width < 130 mT) are significantly narrower than the S_2 ML, with fewer resonances. The S_2 ML is narrower than the only known example from a linear, $Mn^{III}Mn^{IV}{}_3$ net spin ½ tetramer (≈26 resonances, ≈195 mT [25d]), but broader than the g ≈ 2 hyperfine signal from a very recently presented Mn_4/Ca synthetic OEC cluster model, in the $Mn^{III}Mn^{IV}{}_3$ oxidation state (≈18 main resonances, ≈150 mT width [25e]). We discuss this further below.

Several experimental factors subtly influence the detailed, hyperfine structure of the S_2 ML spectrum in functional PSII [12b,26d]. A possible basis for this structural 'flexibility' is seen computationally above. So called 'broad' and 'narrow' (or sharper) ML forms [12b] exist (depending on PSII species, buffer additions *etc.*). Fig. 4a, b gives typical examples of each.

In magnetic systems like the OEC, with strongly exchange coupled centres, the magnetic state (*e.g.* ML) is defined by the total spin operator S_T. Under these circumstances (*e.g.* see [15]), the observed individual ^{55}Mn nuclear hyperfine couplings, $\boldsymbol{A_i}$, are given by,

$$\boldsymbol{A}_i = \rho_i \boldsymbol{A}_{\mathbf{i},ion}. \quad (2)$$

Table 1
Computational exchange coupling constants (cm^{-1}) for OEC magnetic model in Fig. 3.

	J_{12}	J_{13}	J_{23}	J_{34}
S_2 (open)	−130	−50	−105	−45
S_2 (closed)	−145	−50	+30	−100
S_3 (open)	−160	≈0	≈−15 − (+30)	≈20 ± 5
S_3 (closed)	–	–	–	–

Here ρ_i is the spin projection coefficient (dimensionless) of the ion, *i*, within the $S_T = ½$ manifold. It is determined by the precise details of the exchange coupling. $\boldsymbol{A}_{i,ion}$ is a hypothetical 'isolated ion' value for the metal in a chemically 'equivalent' environment (ligands *etc.*). This is in principle obtainable computationally. The ρ_i formally sum to unity. In the anti-ferromagnetically coupled $Mn^{III\text{-}IV}$ dimers above, whose spectra significantly resemble the S_2 ML, $\rho_{III} = 2$ and $\rho_{IV} = -1$, so that the nuclear hyperfine coupling seen from the Mn^{III} ion is approximately twice that from the Mn^{IV}, giving the characteristic, approximately sixteen line pattern. The $\boldsymbol{A}_{III}$ term is anisotropic (Jahn–Teller ion) and typically near axial, with $\boldsymbol{A}_{\mathbf{II}} < \boldsymbol{A}_{\perp}$ in magnitude. Statistically, $\boldsymbol{A}_{\perp}$ dominates the shape in a powder pattern spectrum. Normally $\boldsymbol{A}_{IV}$ is near isotropic, giving the ≈9 mT main line spacing seen in these systems [25a–c] and the ML at moderate resolution (Fig. 4a). At higher resolution (Fig. 4b), a substructure is reproducibly seen in the latter. Both the broad and narrow spectra have virtually identical field positions for the main hyperfine resonances, but differ in the substructure splitting seen, particularly above g = 2. This must arise from an *anisotropic* hyperfine interaction, as it is *not* seen fully on all main resonances (roughly every second resonance in the 300–400 mT region, Fig. 4b). This coupling is of order 80 MHz (*i.e.* ≈3 mT) in the broad form and ≈50 MHz (≈2 mT) in the narrow form. Thus, at least one centre has a small (≈0.2 magnitude) ρ value, while at least one has |ρ| ≈ 1, from the 9.0 mT main pattern spacing (as in the model dimers, likely a near isotropic Mn^{IV} coupling). A centre with |ρ| ≈ 0.2 is inconsistent with most previous phenomenological modelling of the ML signal ([14], see below).

Very recently, Kurashige et al. presented the most advanced quantum chemical calculations yet undertaken on OEC models [16]. These involved Density Matrix Renormalization Group (DMRG) theory, generating 'near exact' many-electron wave function solutions. A robust conclusion from these, even in an assumed high oxidation state paradigm, was that $\rho_1 = 2.0$, $\rho_2 = -1.0$ and, and ρ_3, ρ_4 were each of small, equal magnitude and opposite sign (approximately 0.2 to 0.3 magnitude). Because the *total* ML pattern width is around 5.5 GHz (≈190 mT) and the dominant line spacing is ≈250 MHz (*i.e.* 9.0 mT), these calculations essentially require that one coupling be large, *i.e.* of order 600 MHz, since two Mn couplings are small. This is because, in simple terms [15],

$$\text{ML width}(5.5\ \text{GHz}) \approx 5(A_{1\ \text{eff}} + A_{2\ \text{eff}} + A_{3\ \text{eff}} + A_{4\ \text{eff}}) \quad (3)$$

where the $A_{i,\text{eff}}$ are essentially angularly averaged (effective) ^{55}Mn hyperfine magnitudes. One, (*e.g.* $A_{2\ \text{eff}}$) must be ≈250 MHz (as observed, Fig. 4a), two (*e.g.* $A_{3,4\ \text{eff}}$) are 'small' (<70 MHz) (computational and seen here, Fig. 4b). So $A_{1\ \text{eff}}$ is at least 600 MHz, also consistent with the quantum computational value of $\rho_1 = 2.0$. Kurashige et al., assumed a high paradigm S_2 oxidation state pattern of III, IV, IV, IV. Mn1 and Mn2 have near '$Mn^{III\text{–}IV}$ dimer-like' oxidation states and contributions, while Mn3 and Mn4 are near coupled silent. Since at least one of the latter two has a substantially anisotropic hyperfine coupling, *i.e.* is likely a Mn^{III} ion, then both ions would have to be in the low paradigm S_2 state.

In principle, ^{55}Mn ENDOR spectroscopy on the ML state should resolve this. ENDOR, particularly pulsed (Davies) ENDOR, directly reveals the hyperfine couplings within an electron spin manifold by inducing nuclear transitions (typically MHz frequencies) within the rotating frame of the electron magnetization (see [27]). Britt et al. [14a,b] made the first such studies, followed by Bittl et al. [14c,d] and Lubitz et al. [14e,f] These measurements were taken at ≈4.2 K, but also recently by Pantazis et al. at 2.5 K. [14g] To first order, a hyperfine coupling of *A* MHz gives ENDOR frequencies at ≈|*A*/2| MHz, for this spin ½ system, with strong nuclear hyperfine coupling.

Essentially similar results were reported, at X or Q band frequencies, including the 2.5 K data, from plant or cyanobacterial PSII. Hyperfine couplings from ≈170–340 MHz (*i.e.* ≈85–170 MHz in ENDOR) were

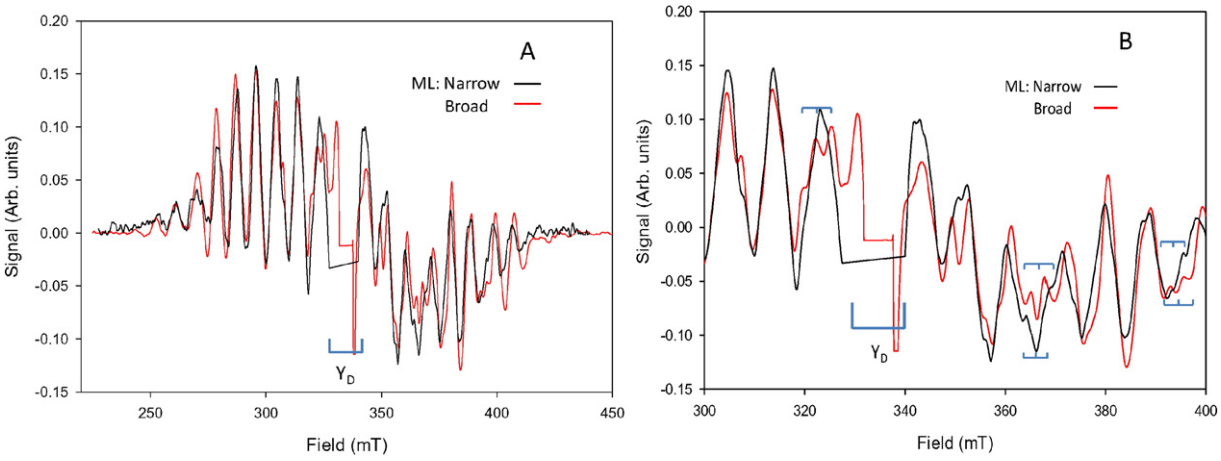


Fig. 4. a) X Band CW EPR difference ($S_2 - S_1$) spectra of typical Multiline forms of the 'broad' (red, PSII Cores) and 'narrow' (black, PSII Membranes) types. Central region (indicated) about $g = 2$ is obscured by Tyr_D subtraction artifact. b) Expansion of (A) between 300 and 400 mT. Stick spectra (blue) reveal several examples of similarly patterned peak splittings. Data from Jin et al. [15], q.v. for experimental details.

seen in a broad, variably shaped envelope, with a coupling maximum at ≈240–280 MHz. No couplings above 400 MHz or below ≈160 MHz were reported. Many sample and instrumental factors influence ENDOR transition intensities (see discussion in [15]), generally not well-defined or controlled in a particular experiment. Hence, the *total number* of distinct nuclear transitions contributing to a particular ENDOR resonance region is typically unclear, unless the transitions are sufficiently narrow or separated from others (then presumed to be single transitions). So these results were interpreted in terms of three nearly equivalent isotropic Mn couplings (≈200–230 MHz magnitude, approximately overlapping in the spectra) and one axial anisotropic coupling, with perpendicular component at ≈350 MHz [14]. The latter was assigned to a single Mn^{III} and the former to three Mn^{IV} ions in the high oxidation state models assumed. The spin projections were about 2, −1, 1, −1, with the first assigned to the Mn^{III}, while sign ordering varied for the others, subject to them summing to −1. No interpretations of line substructure (if seen), as in Fig. 4b, were made.

One factor which may have influenced all ENDOR studies on the OEC ML is Mn nuclear relaxation. If too high, this causes dephasing and loss of ENDOR intensity (see [15,27]). These nuclear rates depend strongly on the electron spin relaxation (T_1 rate), often declining sharply near 0 K. As all previous work to that time had been performed at ≈4.2 K or above, Hanson, Pace et al. recently undertook Davies ENDOR studies on plant PSII core complexes, at temperatures down to 2.5 K [15]. These PSII complexes have excellent optical properties, high O_2 activity [15] and ML electron T_1 rates (at 4 K) ≈ 30 times lower than for any PSII material elsewhere reported (see [15]). At 2.6 K, this gave ML T_1 rates ≈150 times slower than any previously obtained, to that point. The majority sample rate (≈50 s^{-1}) is still an order of magnitude *slower* than that recently reported by Pantazis et al. [14g] for cyanobacterial PSII complexes at 2.5 K (≈500 s^{-1}). The plant core complexes display the broad ML form in S_2, with ^{55}Mn ENDOR transitions far sharper than any previously reported (Fig. 5A, B). As a consequence of this, precise field position sampling effects can be pronounced, particularly for the Mn1 contribution, which is the most anisotropic centre, with large quadrupole component, Table 2 (see [15] for further details). The observed transitions correspond to ^{55}Mn hyperfine couplings in the range ≈20–670 MHz, as expected from the simple arguments above and the high level calculations of Kurashige et al. [16]. It was concluded in [15] that all previous ENDOR studies on the ML signal had essentially seen *only* resonances from the (single) Mn^{IV} centre in the S_2 OEC. This was rationalizable, in magnetic relaxation terms, as Mn^{III} centres typically experience faster T_1 relaxation than Mn^{IV} or Mn^{II} centres at equivalent temperatures (see Discussion). This will also be significant below, when we consider EPR results on the S_3 state.

The continuous wave (CW) data in Fig. 4a, b and the ENDOR data in Fig. 5a, b were earlier simulated (see [15] for full details). All simulation spectra (CW and ENDOR) assumed powder pattern averaging of the sample, although some orientation selection effects likely occur experimentally, as noted above. These appeared to be too subtly dependent on precise observing field position, sample variation *etc.* to reliably address experimentally or in simulation, given the high anisotropy of some of the hyperfine interactions. The CW simulations for the broad and narrow forms are given in [15]. In Fig. 5, the red simulation spectra for the broad form (from [15]) are shown, with representative experimental data as indicated in the figure captions (and [15]). In all cases the fittings were used to obtain the experimental hyperfine parameters. No spectral simulations using computationally derived hyperfine parameters were undertaken. Table 2 lists the best fit CW and ENDOR derived parameters. The hyperfine and quadrupole couplings are second rank, symmetric Cartesian tensors (the quadrupole tensor is traceless) and the tensor orientations, relative to a convenient reference frame, were originally defined in [15] by Euler rotation angles. The reference frame is that of Mn1, (**g** tensor frame, see Fig. 6 and [15] for details). The CW simulations were at 2nd order perturbation level (excluding quadrupole interactions). The ENDOR frequency simulation calculations (Fig. 5a, b) are exact, but the powder pattern simulation was approximate, due to substantial orientation selection effects arising in this system (see [15] for details). ENDOR data were only available for the broad ML form (PSII core complexes), the narrow form data (more dilute PSII membrane samples) are from CW simulations. The Mn hyperfine parameters from the CW and ENDOR simulations agreed to within 1.5% for the broad form (Table 2) and averaged values were used in subsequent analyses. The estimated orientations of the Mn2, 3, 4 hyperfine tensors, relative to the Mn1 frame, were determined from the CW fittings. For these smaller couplings, the 2nd order fitting is largely insensitive to the actual sign of the hyperfine tensor. Thus the fitting effectively gives only the unsigned orientation of a particular principal axis (*i.e.* X or −X direction, say), whereas the DFT calculations (below) give the signed principal axis directions. Hence in the following (see Tables 3a, 3b, 3c, 3d, 3e) we quote simply the magnitudes of scaler products (acute angle cosines) of particular fitted unit axes directions with the reference frame directions, *i.e.* $X_1 \cdot X_i$ *etc.*, where $i = 2, 3, 4$. It is known that the intrinsic sign of the ^{55}Mn hyperfine tensor is negative and this is assumed in Table 2. Only the hyperfine interactions were considered computationally in the present study.

The pattern of magnitudes seen in the experimentally determined hyperfine parameters in Table 2, are totally consistent with the above arguments, both empirical and theoretical, on the likely ρ_i values for the OEC Mn ions in S_2. These are indicated (from [15]), based mainly

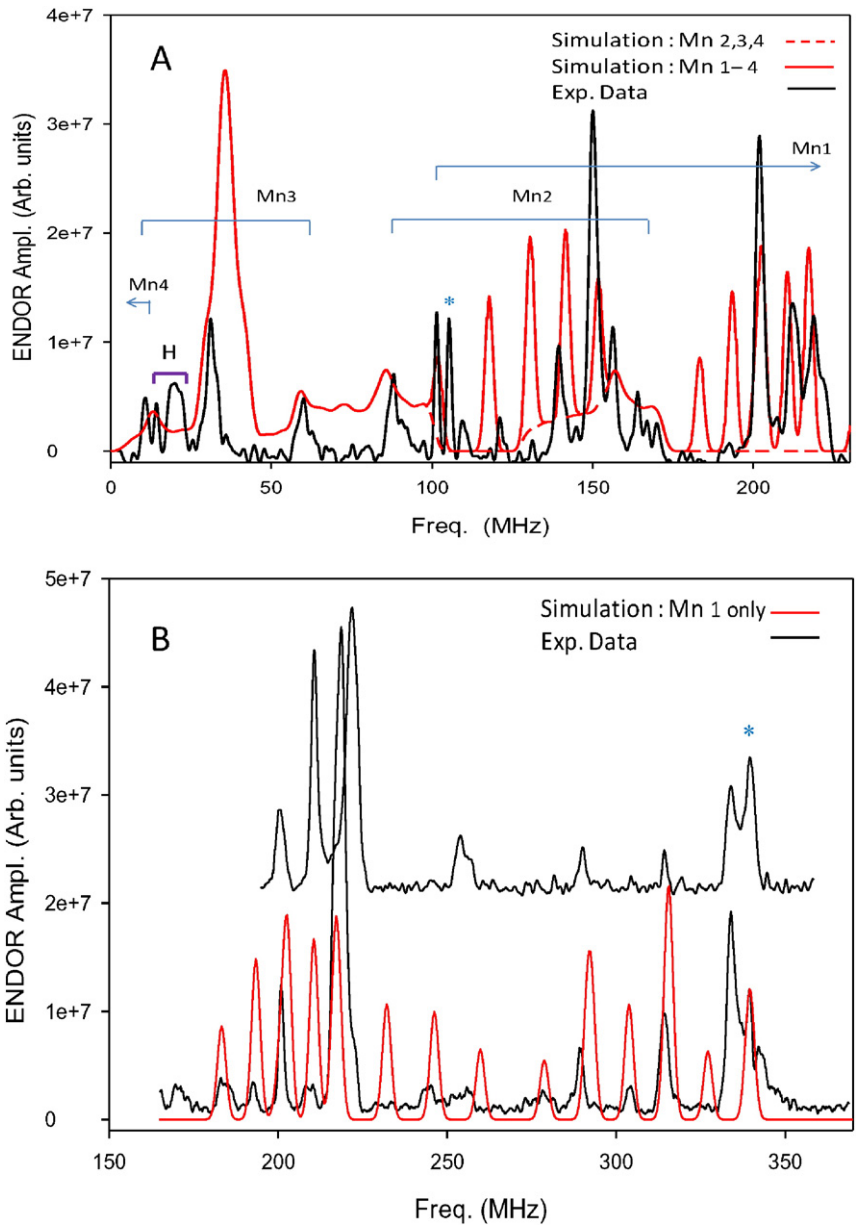


Fig. 5. X Band PSII Core complex ^{55}Mn ENDOR spectra of ML S_2 state for Low (A) and High (B) frequency regions at 2.5 K. A is average of several independent data sets with nominal ENDOR sampling field position at g = 2.05 ± 0.02 (~325 mT, see [15]). Data (black) and total simulation (solid red) using the spin Hamiltonian parameters listed in Table 2 (ENDOR). Individual Mn regions indicated in figures. Red dashed line in (A) is powder pattern simulation from Mn2, 3, 4, only which constitutes the total simulation below 100 MHz, (where it is shown full). Mn1 contribution is subject to orientational selection emphasis (see [15] and main text). In the High frequency region, particularly around 200 MHz, significant orientational selection at the ENDOR measurement point occurs (two representative examples shown, separate samples, nominal sampling field, g = 2.05), emphasising or suppressing individual transitions from Mn1, whose hyperfine interactions are highly anisotropic (see text and [15] for details). The resonances with ≈10 MHz spacing are from the quadrupole splitting on Mn1 (see [15]).

on the observed, fitted tensor magnitudes and known values for typical Mn systems (below). Mn1 and Mn2 are near Mn^{III}–Mn^{IV} 'dimer like', ($\rho_1 = +2$, $\rho_2 = -1$, using standard expressions, [12,25c]) while Mn3 and Mn4 are near coupled 'silent' ($\rho_3 \approx \rho_4 \approx 0$). Although we have not undertaken (computationally expensive) entangled state calculations on our model structures, the results in Table 1 allow us to readily rationalize how this coupling pattern could arise within our structural models of the OEC. The degenerate ground state total spin functions, $|M_S> = \pm$ ½, must contain, to good approximation, spin basis functions like:

$|\alpha_1\beta_2>[|\alpha_3\beta_4>-|\beta_3\alpha_4>]$ and $|\beta_1\alpha_2>[|\alpha_3\beta_4>-|\beta_3\alpha_4>]$.

These give 'dimer like' spin projections for Mn1 and Mn2, while the projections (from spin expectation values) for Mn3 and Mn4 are small (spin of $S_3 + S_4 \approx 0$). For these two M_S states to be degenerate, it is necessary, in simple terms, for the ABAB, ABBA *etc.* single determinant energies to be approximately equal. This is equivalent to $J_{23} = (4/3)J_{13}$ in Fig. 3. From Table 1, it seen that the open S_2 form is not far from this condition. Both J_{23} and J_{13} are antiferromagnetic and $J_{23} > J_{13}$, in magnitude, but $-J_{23}$ should be closer to ≈70 cm^{-1}, rather than ≈100 cm^{-1}, for the condition to closely hold. However, from Table 1, the J_{23} term is by far the most geometry sensitive, in fact changing sign between the open and closed forms, while J_{12}, J_{13} and J_{34} remain consistently, strongly antiferromagnetic. It seems quite reasonable that in the actual protein environment of the open form, the $J_{23} = (4/3)J_{13}$ condition could nearly apply. This would appear less likely for the closed form. Variants for the

Table 2
CW EPR and ENDOR simulation parameters for the S_2 multiline signal.

Spin Hamiltonian interaction[a]	X	Y	Z	Isotropic value	ρ
A: Broad form					
g[b] (plant)	1.989	2.008	1.964	1.987	–
A_{Mn1}[c,d]	−668 (−*658*)	−536 (−*534*)	−337 (−*329*)	−514 (−*507*)	≥2.0
Q_{Mn1}[c]	(−*8*)	(*4.5*)	(*3.5*)	–	–
A_{Mn2}[c]	185(*185*)	242 (*242*)	273 (*273*)	233 (*233*)	−1
A_{Mn3}	−78 (−*78*)	−12 (−*12*)	−46 (−*46*)	−45 (−*45*)	<0.2
A_{Mn4}	1 (*1*)	20 (*15*)	1(*1*)	7 (*6*)	< |0.05|
Euler angles[e]	α	β	γ		
R_{Mn2}	0.422	0.75	−0.737	–	–
R_{Mn3}	−0.135	0.21	0.284	–	–
R_{Mn4}	0	0	0	–	–
B: Narrow form[f]					
g (plant)	1.989	2.008	1.964	1.987	–
A_{Mn1}[c,d]	−668	−536	−337	−514	≥2.0
Q_{Mn1}[c]	–	–		–	–
A_{Mn2}[c]	185	242	273	233	−1
A_{Mn3}[c]	−45	−8	−32	−28	<0.1
A_{Mn4}[c]	–	–	–	–	–
Euler angles[e]	α	β	γ		
R_{Mn2}	0.173	0.35	−0.338	–	–
R_{Mn3}	−0.14	−0.06	−0.08	–	–
R_{Mn4}	–	–	–	–	–

[a] ENDOR parameters in brackets. See [15] for additional simulation parameters, line widths *etc.*
[b] Data for plant PSII. Cyanobacterial values very similar, see Rapatskiy et al. [14h].
[c] Units for *A* and *Q*, MHz.
[d] Total effective values (ρA_{ion}, see text).
[e] Units for Euler angles, radians. Referenced to Hyperfine frame of Mn1.
[f] CW simulation only. See [15] for details.

latter may be related to the higher spin states ($S_T = 5/2$ *etc.*) seen in S_2, but this is still under examination by us.

There are several additional notable features concerning the experimental hyperfine values in Table 2. The fitted principal axis values for Mn1 and Mn2 are virtually identical in the broad and narrow forms, although the Mn2 tensor is more closely aligned to the Mn1 reference in the narrow form, compared to the broad. The possible significance of this is uncertain, as the anisotropy of the Mn2 tensor is not large (it arises from a Mn^{IV} ion) and it is this anisotropy which drives the fitting for the tensor orientation. The fitted Mn1 tensor values are large and almost purely rhombic, not near axial as is typical for Mn^{III} in model systems with more ligand symmetry [12,14]. This is certainly consistent, qualitatively, with the low symmetry of the Mn1 site in the OEC. Further, the principal distinction between the broad and narrow ML forms is in the net hyperfine coupling to the Mn3 and Mn4 centres, which is stronger (but still small) in the broad form (making it broader) than the narrow. No coupling to Mn4 is discernable in the CW fitting of the narrow form. From the above arguments, this suggests a slightly closer effective alignment to the $J_{23} = (4/3)J_{13}$ condition in the narrow, compared to the broad form. Both, however, would be 'open' with water access to the Mn1 axial site. From the fact that for either Mn^{III} or Mn^{IV}, the isolated ion A_{iso} value is ≈−200 to −250 MHz [12,25c], estimates of the spin projection values for each ion in the OEC ML state are indicated.

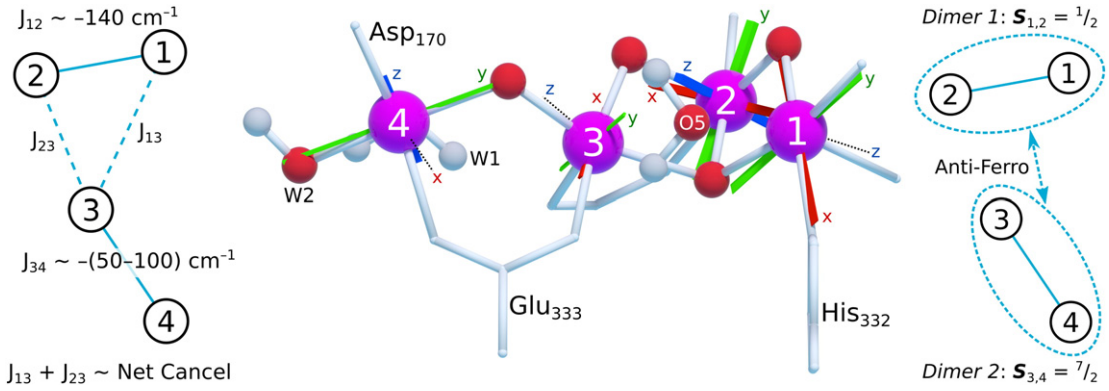


Fig. 6. Geometric and electronic structures of the OEC, from present work and [15]. Left: Magnetic coupling scheme for the S_2 OEC, based on DFT calculations of model OEC clusters in Figs. 1, 2. Near net cancellation of the J_{23} and J_{13} interactions would need to occur for consistency with S_2 EPR data, supported by current studies and recent 'entangled state' quantum calculations [16]. Center: Structural model of the OEC (S_2 state), based on the XRD structures and present computational modelling (see text). The O5 species is identified as a water molecule, as is the W1 species, but W2 is likely a hydroxide in S_2 (see Fig. 1). The molecular axis system, on Mn1, consistent with data in [15] and present calculations shown. This has Z axis along the Mn1–O5 direction. The X axis points along the Mn1–N_τ (His332) direction. Individual Mn hyperfine tensor axes and relative magnitudes indicated. Right: Coupled dimers model appropriate for high spin ($S = 3$) state of S_3. *Dimer 2* is ferromagnetically coupled internally (see text). In low spin states of S_3, S_2, this coupling is antiferromagnetic. *Dimer 1* is always antiferromagnetic and coupling between dimers is weakly antiferromagnetic.

Table 3a
S_2 Wet ABAB A-tensor parameters at PBE0/TZP-ZORA level of theory (Experimental values in brackets, here and in subsequent tables).

Spin Hamiltonian parameter	Mn	A_{aniso}			Isotropic value A_{iso}	Assumed ρ
		X	Y	Z		
Isolated ion hyperfine $A_{aniso,XYZ}$, A_{iso} (MHz)	1	−143 (−77)	−62 (−12)	+205 (+89)	−370 (−255)	+2.0
	2	+21 (+24)	−14 (−4)	−8 (−20)	−366 (−233)	−1.0
	3	−70 (−165)	+204 (+165)	−135 (≈0)	−261 (−225)	≈0.2
	4	+33 (+120)	−200 (−220)	+167 (+100)	−382 (−140)	≈−0.05

Isolated ion ^{55}Mn hyperfine tensors were computed, from the ABAB and ABBA configurations, using two different functionals (BP86 and PBE0) with relativistic corrections (see Methods). The results are summarized in Tables 3a-d (for open form structures). A typical feature of such calculations is that they reasonably reproduce the anisotropy of the relevant Mn ion, which arises from the spatial distribution of the unpaired d electrons, but give poorer performance in calculating the isotropic coupling. This arises from polarization of the inner core electrons, generally computable with accuracy only by very high level methods [28]. This was observed here, with the computed (pure anisotropic, traceless) $A_{aniso\ X,Y,Z}$ terms, for all ions, consistently agreeing, within <5 MHz, between the two functionals, for the averaged ABAB and ABBA configuration values. This averaging is performed on the grounds that the relative energies of these spin configurations are within the chemical accuracy of density functional methods and interchange between the 'open' and 'closed' forms of the OEC, rendering the minimal configuration ambiguous. With the A_{iso} values however, estimates from the two functionals could vary by up to a factor of ≈2, with BP86 typically underestimating (compared to typical ^{55}Mn values, *i.e.* ≈200–250 MHz) and the PBE0 functional overestimating. The values of all computed hyperfine parameters quoted in Table 3e are averages from the two functional results. We note that the functional-averaged isotropic values are generally quite reasonable with respect to the experimental results, suggesting that the BP86 and PBE0 functionals 'bracket' the experimental behaviour. This however constitutes an empirical observation only and at present we have no direct theoretical support for the averaging. The matter would obviously need a more comprehensive analysis of the specific dependence of *A*-tensor characteristics on particular functional type. The corresponding experimental parameters (Table 2) are for the broad form only, as these data include ENDOR results and a resolved contribution from Mn4. They are quoted now in the same format as the computational results (separate A_{aniso} and A_{iso} components) and have been scaled to the corresponding isolated ion estimates, using the earlier estimated spin projection factors (see [15]). This procedure is regarded as quite robust for Mn1 and Mn2, but for Mn3 and particularly Mn4 it is more uncertain, as the actual observed couplings for these ions are small, with larger attendant uncertainties from the fitting.

From Table 3e, the computations reproduce, qualitatively and even semi-quantitatively, the rhombic character of the hyperfine tensors, in sign and general magnitude. The DFT methodology calculates the anisotropic hyperfine component from the orbitally integrated, total nuclear-unpaired electron spin magnetic dipolar interaction only [28]. Computationally, the total anisotropy is very similar (within 10%) to that estimated experimentally for Mn3 and Mn4, within 30% for the (small) value for Mn2, but exceeds that for Mn1 by a factor of ≈2. Also, the axes of largest signed magnitude anisotropy for the Mn^{III} centres are consistently, correctly identified. For Mn2, 3 the hyperfine tensor orientations relative to the Mn1 axes are generally close to the experimental estimates. For Mn4 the latter are quite uncertain, but the computational values are reasonable by comparison. However, an interesting trend is evident in Table 3e. While the computational total anisotropies generally exceed the experimental values for the Mn^{III} ions (Mn1, 3, 4), the observed total anisotropy for the inferred Mn^{IV} centre (Mn2) is likely greater than the computational value. It is known [29] that in ground (spin ½) coupled Mn^{III-IV} systems, the relatively large (several cm^{-1}) Zero Field Splitting (ZFS) term (*D*) on the d^4 Jahn–Teller active Mn^{III} ion allows a contribution, from higher total spin states, to the effective hyperfine interactions on all the Mn ions present. This scales as *D*/*J*, where for the dimers *J* is the single exchange coupling present. Typically, this effect is small (fractionally) for the Mn^{III} ion, whose large anisotropy is normally dominated by the electron/nuclear dipole interactions. However if $J \approx 100\ cm^{-1}$ or less, the ZFS effect can be significant, relative to the normally small dipolar hyperfine anisotropy on the Mn^{IV} centre (d^3 ion) [29]. This may be occurring here, with the Mn1, 2 'effective dimer' having a *J* coupling of ≈130 cm^{-1} (Table 1).

S_3 State: The S_3 intermediate is even spin. From Table 1, in the open form, the coupling is now clearly 'coupled dimers' (Fig. 6), as $J_{13} \approx 0\ cm^{-1}$. *Dimer 1* has strong *anti*-ferromagnetic internal coupling, while *Dimer 2* exhibits moderate ferromagnetic internal coupling, on balance, computationally. However some DFT determinant energy combinations are consistent with a moderate anti-ferromagnetic interaction (Supplementary material § S2). This is likely experimentally significant (below). The coupling between these dimers (J_{23}) is weak, either ferro- or anti-ferromagnetic. In the closed form, the ground state appears to be clearly diamagnetic (ABAB or ABBA), spin zero, and so will not be EPR active.

Cox et al. have recently reported high field (Q, W band) pulsed EPR on the S_3 state in thermophilic PSII core complexes [30]. The S_3 state is heterogeneous in these systems, one population (≈60%) exhibits a high spin, $S_T = 3$ magnetic state (with small fine structure, near pure rhombic, $D \approx 0.2\ cm^{-1}$, $E/D \approx 0.3$). The other population is EPR silent, but responds to NIR turnover, forming the $S_2/Y_Z^{\cdot}$ split signal state (see [9]). This is readily rationalized within our modelling of the S_3 state. In the $S_T = 3$ form, the open cluster structure exists and *Dimer 2* is ferromagnetically coupled internally (net spin 7/2) and then (weakly) anti-ferromagnetically coupled to *Dimer 1*. In the NIR sensitive, diamagnetic population, the structure is likely of the closed form, but the weak, variable coupling predicted for the Mn2⋯Mn3, and Mn3⋯Mn4 interactions (Table 1) means that some variant on the open form could also be consistent with this. Thus, if the *Dimer 2* internal coupling is anti-ferromagnetic (net spin ½) and the Mn2⋯Mn3 interaction also anti-

Table 3b
S_2 Wet ABBA A-tensor parameters at PBE0/TZP-ZORA level of theory.

Spin Hamiltonian parameter	Mn	A_{aniso}			Isotropic value A_{iso}	Assumed ρ
		X	Y	Z		
Isolated ion hyperfine $A_{aniso,XYZ}$, A_{iso} (MHz)	1	−134 (−77)	−73 (−12)	+208 (+89)	−363 (−255)	+2.0
	2	+5 (+24)	−18 (−4)	−13 (−20)	−325 (−233)	−1.0
	3	−66 (−165)	+201 (+165)	−135 (≈0)	−171 (−225)	≈0.2
	4	+31 (+120)	−198 (−220)	+167 (+100)	−376 (−140)	≈−0.05

Table 3c
S_2 Wet ABAB A-tensor parameters at BP86/TZP-ZORA level of theory.

Spin Hamiltonian parameter	Mn	A_{aniso}			Isotropic value A_{iso}	Assumed ρ
		X	Y	Z		
Isolated ion hyperfine $A_{aniso,XYZ}$, A_{iso} (MHz)	1	−135 (−77)	−59 (−12)	+195 (+89)	−236 (−255)	+2.0
	2	+22 (+24)	−19 (−4)	−3 (−20)	−304 (−233)	−1.0
	3	−68 (−165)	+197 (+165)	−129 (≈0)	−116 (−225)	≈0.2
	4	+33 (+120)	−201 (−220)	+168 (+100)	−179 (−140)	≈−0.05
Hyperfine tensor orientation cosines relative to Mn1 axes	1	1.0	1.0	1.0		
	2	0.71(0.88)	0.91 (0.77)	0.63 (0.74)		
	3	0.95(0.99)	0.99 (0.97)	0.93 (0.98)		
	4	31) 0.84 (≈1)	0.79 (≈1)	0.83 (≈1)		

ferromagnetic, the total cluster spin would be 0. Ferromagnetic coupling in *Dimer 2* could reflect H-bonding to the Mn3···4 oxo bridge, by likely substrate water species (Fig. 2a).

The W Band ^{55}Mn EDNMR is most readily interpreted, as the high field separates the electron m_S sub-manifolds in the EPR absorption spectrum (see [30]). The data (*e.g.* Fig. S9 in [30]) reveal two clear nuclear couplings (called Class 1 and 2 in [30]), of ≈100 MHz and ≈30 MHz magnitude, respectively. The Class 1 resonances are narrow (<10 Hz width, generally un-split peaks) and have been assigned as two virtually identical, overlapping near isotropic couplings ($A_{iso} \approx -99$ and −96 MHz). Only one coupling, $A_{iso} \approx -26$ MHz, is assigned to Class 2 (a fourth coupling is ≈0). Within the signal to noise of the spectra presented, it is completely unclear that each class contains more than one spectral contribution, *i.e.* only *two* distinct couplings are seen. The measurements were made at ≈5 K and assuming the nuclear relaxation effects discussed above on the Mn^{III} centres apply, only resonances from the Mn^{IV} centres are expected to be visible. In the low paradigm, these are Mn2 and Mn4 in S_3. Again, within the simple dimer magnetic interaction scheme of Fig. 6, the spin projected values of the ^{55}Mn hyperfine interactions are readily determined. Each Mn centre first projects onto its *Dimer* (*1* or *2*, spins ½ or 7/2), then the dimer value projects onto the total cluster spin, $S_T = 3$. Using standard expressions (*e.g.* [12a, 25c]), this gives:

$$\rho(\mathrm{Mn4}) = \rho(Dim2)\cdot\rho_4 = (9/8)\cdot(3/7) = 0.482,\ \ \rho(\mathrm{Mn2}) = \rho(Dim1)\cdot\rho_2 = (-1/8)\cdot(-1) = 0.125. \quad (4)$$

The $A_{iso,\ ion}$ value for Mn2 (in S_2) is estimated from Table 2 to be −233 MHz, giving a projected A_{iso} value in S_3 of ≈ −29 MHz. There is no independent empirical estimate for Mn4 in the IV state, but Mn^{IV} centres in catalase mimic dimers [25c] have $A_{iso,\ ion}$ values of ≈ −200 to −240 MHz (only moderately anisotropic). This gives a range for the projected Mn4 A_{iso} value in S_3 of ≈ −96 to −115 MHz. These two assignments are in excellent agreement, both magnitude and sign, with the data of Cox et al. and in fact better agreement than their own estimates from DFT modelling of the S_3 OEC cluster, or empirically derived $A_{iso,\ ion}$ values from S_2 ENDOR fitting. Mn4 gives the Class 1 coupling and Mn2 the Class 2. The larger apparent line-width assigned to the Class 2 coupling (14 MHz, *vs* ≈6 MHz for Class 1) is a consequence of the relatively large (for a Mn^{IV} ion), near-rhombic hyperfine anisotropy of Mn2 (see Tables 3a, 3b, 3c, 3d, 3e).

4. Discussion, conclusions

The Mn spin projection pattern earlier deduced for the S_2 ML state, from ENDOR [15] and recently from DMRG computations [16], namely $\rho_1 \approx 2.0$, $\rho_2 \approx -1$, $|\rho_{3,4}|$ small ($\rho_4 \approx 0$ from ENDOR), is non-intuitive, given the OEC geometry. It is a consequence of the cluster internal interactions being magnetically close to a 'coupled dimers' configuration. (Fig. 6) As the calculations here show, this is quite plausible in the low oxidation state paradigm for an 'open' form of the cluster geometry. Then a water molecule (labelled W_f in Fig. 1a) has some access to Mn1 along its Jahn–Teller axis direction. This corresponds to the O5 group in the high resolution XRD structures [3a–c] and has been proposed by us to be a substrate species (so-called 'fast water', [6]). In S_2, one dimer (Mn1···Mn2, *Dimer 1*) is anti-ferromagnetically coupled internally, with net spin ½. The other (Mn3···Mn4, *Dimer 2*) is also anti-ferromagnetically coupled internally, with net spin ≈0. The coupling between the dimers is not small, by individual Mn–Mn interactions, but the Mn pairs nearly 'uncouple' overall, because of 'balancing' Mn1···Mn3 and Mn2···Mn3 interactions and the fact that Mn4 couples only to Mn3. In S_3 however, the open form of the OEC cluster now resembles, magnetically, a linear chain, with negligible Mn1···Mn3 interaction. As such it readily accommodates the effective 'coupled dimers' organization, if the Mn2, 3 interaction is sufficiently small, which our DFT calculations support.

As noted earlier by Jin et al. [15], the hyperfine parameters for Mn1 are unique by comparison with Mn^{III} centres in model $Mn^{III,IV}$ systems and superoxide catalase (the closest protein analogue, [25b,c]). The hyperfine tensor is almost purely rhombic, with $A_X \approx 660$ MHz and fractional anisotropy ≈65% of A_{iso}. For various model systems [25], the symmetry of the Mn^{III} hyperfine tensor ranges from near axial, to moderately rhombic, with a maximum fractional anisotropy of ≈45% of A_{iso} and maximal (A_X) component of ≈500 MHz (not both for the same compound). Compared to the model systems in [25b] (including catalase), the OEC (Fig. 1) has the lowest symmetry (C_1) of any. Further, *all* Mn centres assigned here as Mn^{III} in S_2 (Mn1, 3, 4) are near five

Table 3d
S_2 Wet ABBA A-tensor parameters at BP86/TZP-ZORA level of theory.

Spin Hamiltonian parameter	Mn	A_{aniso}			Isotropic value A_{iso}	Assumed ρ
		X	Y	Z		
Isolated ion hyperfine $A_{aniso,XYZ}$, A_{iso} (MHz)	1	−126 (−77)	−73 (−12)	+199 (+89)	−231 (−255)	+2.0
	2	+11 (+24)	−24 (−4)	−13 (−20)	−264 (−233)	−1.0
	3	−67 (−165)	+199 (+165)	−133 (≈0)	−6[a] (−225)	≈0.2
	4	+31 (+120)	−199 (−220)	+168 (+100)	−174 (−140)	≈−0.05
Hyperfine tensor orientation cosines relative to Mn1 axes	1	1.0	1.0	1.0		
	2	0.93 (0.88)	0.96 (0.77)	0.92 (0.74)		
	3	0.90 (0.99)	0.99 (0.97)	0.90 (0.98)		
	4	0.59 (≈1)	0.80 (≈1)	0.19 (≈1)		

[a] Abnormally small, not used for averaging.

Table 3e
Comparison of average computational and Broad Form Experimental ^{55}Mn Hyperfine Parameters for the OEC S_2 ML Signal.

Spin Hamiltonian parameter[a]	Mn	A_{aniso}			Isotropic value A_{iso}	Assumed ρ[b]
		X	Y	Z		
Isolated ion[c] hyperfine $A_{aniso,XYZ}$[d], A_{iso} (MHz)	1	−136 (−77)	−67 (−12)	+201 (+89)	−299 (−255)	+2.0
	2	+16 (+24)	−18 (−4)	+2 (−20)	−314 (−233)	−1.0
	3	−67 (−165)	+200 (+165)	−133 (≈0)	−139 (−225)	≈0.2
	4	+31 (+120)	−199 (−220)	+168 (+100)	−278 (−140)	≈−0.05
Hyperfine tensor orientation cosines[c] relative to Mn1 axes	1	1.0	1.0	1.0		
	2	0.83 (0.88)	0.94 (0.77)	0.78 (0.74)		
	3	0.92 (0.99)	0.99 (0.97)	0.92 (0.98)		
	4	≈0.7 (≈1)	≈0.8 (≈1)	≈0.6 (≈1)		

[a] Experimental values in brackets.
[b] Assumed value to convert experimental parameter to isolated ion values, see text.
[c] Averages of ABAB and ABBA for computational values.
[d] $A_{X,Y,Z}$, as in Table 1, $= A_{aniso,XYZ} + A_{iso}$.

coordinate (*vs* six coordinate in model complexes), as O5 is at least 2.5 Å from any of these Mn^{III} ions, along their Jahn–Teller axes [6].

As noted, the DFT estimates of the anisotropic hyperfine interactions, particularly for the Mn^{III} centres, are semi-quantitatively accurate, with correct sign and relative orientation in almost all cases. Consistent with these centres being near five co-ordinate, the tensors are all rhombic prolate ellipsoids, with the minimum (magnitude) axis pointing along the singly co-ordinated direction. The inferred computational hyperfine EPR axis frame, for Mn1, sits naturally within the structural model (Fig. 6b). The (largest magnitude) X axis points closely along the Mn1–N_τ His332 vector, the Y axis is normal to this in the di-μ-oxo plane of Mn1 (close to the O3–Mn1 direction) and the Z axis is normal to the plane, approximately along the O5–Mn1 direction. This is in fact the reverse of the Mn (d orbital) unpaired spin density distribution (most extensive along Z direction), due the effective opposite signs of the (traceless) electron-nuclear dipole-dipole interaction and the (negative) sign of the contact (*i.e.* isotropic) hyperfine coupling for ^{55}Mn ions [27]. Thus, the metal-ligand covalency is largest along the Mn-imidazole direction, which is expected.

The open S_3 computational OEC model (Fig. 2a) provides a totally 'natural' framework within which to quantitatively interpret the recent W Band ^{55}Mn EDNMR results of Cox et al. This assumes again that relaxation effects suppress the resonances from the two Mn^{III} ions expected to be present in the S_3 cluster. The unusual location of the O5 (W_f) water may underlay this 'relaxation suppression' of ENDOR /EDNMR *etc.* from the Mn^{III} centres. The O5 (*i.e. Wf*) water sits in a shallow potential well, approximately at the intersection of the Jahn–Teller axes on Mn1, 3, 4 in S_2 and Mn1, 3 in S_3 (Figs. 1a, 2a). As such it will execute large amplitude, low frequency thermal libration, even at low temperatures. This can act as an efficient relaxation mechanism for electron (and nuclear) spins, as Mn^{III} has a singly occupied e_g 3d orbital, with large Zero Field Splitting (ZFS) term. Thermal lattice modulation of this interaction (by Jahn–Teller splitting modulation *etc.*) can strongly enhance magnetic relaxation locally [27]. It could operate in both the S_2 *and* S_3 states within the low paradigm. Thus, Mn^{III} ENDOR suppression may occur more readily in open (wet) ML forms (typically employed in most reported studies) as the closed (dry) form has O5 excluded from significant interaction with the OEC. This matter is currently under computational examination by us.

Lastly, a very recent OEC cluster model of Zhang et al. [25e] illustrates (perhaps in an unanticipated way) key points here. This Mn_4/Ca complex has mostly oxygen ligation, a (Mn_3Ca)–Mn organization and unquestionably resembles the OEC, structurally. It is a 'high paradigm' system, as all Mn are 6 coordinate and O5 is a μ_3-oxo bridge, as typical in high paradigm computational modelling. [5,8] In its 'S_2' state, the oxidation pattern is likely (III, IV, IV, IV), closely resembling that assumed by Kurashige et al. [16] in their DMRG calculations of S_2. The Mn1 centre is near symmetrically coordinated (O ligands only), with Mn4 having the sole N ligand. The 'ML' pattern has approximately 18 main lines, with ≈8.5 mT (≈240 MHz) spacing downfield, showing internal splittings of ≈100 MHz upfield. The overall width is ≈150 mT (≈4.2 GHz). This is close, in simple terms (from Eq. 3) to that expected when A_{Mn1}, A_{Mn2} have their Mn Catalase values (≈ − 420(⊥), 250 MHz, respectively [25c]) and|$A_{Mn3,4}$| are 'small' (≈100 MHz), *i.e. precisely* the pattern predicted by Kurashige et al. for a high paradigm OEC state! So, the Zhang et al. complex shows us what would be expected in S_2 (or close to it) if the high paradigm actually obtained. However, the observed OEC S_2 ML is >25% wider than that of the Zhang model, with a different internal line splitting pattern. The results presented here, both experimental and computational, in total give yet further support for the low oxidation state paradigm in the OEC.

Abbreviations

OEC	Oxygen Evolving Complex
PSII	Photosystem II
EPR	Electron Paramagnetic Resonance
ESR	Electron Spin Resonance
ENDOR	Electron-Nuclear Double Resonance
NMR	Nuclear Magnetic Resonance
FTIR	Fourier Transform Infrared
ELDOR	Electron-Electron Double Resonance
EDNMR	ELDOR-Detected NMR
ML	Multiline

Acknowledgements

This research was undertaken on the NCI National Facility in Canberra, Australia, which is supported by the Australian Commonwealth Government.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jinorgbio.2016.04.009.

References

[1] K. Satoh, T. Wydrzynski, Govindjee, Photosystem II: The Light Driven Water: Plastoquinone Oxidoreductase, Springer Dordrecht, The Netherlands, 2005.
[2] P. Gatt, R. Stranger, R.J. Pace, J. Photochem. Photobiol. B 10 (2011) 80–93.
[3] a) Y. Umena, K. Kawakami, J.-R. Shen, N. Kamiya, Nature 473 (2011) 55–60;
b) H.M.K. Koua, Y. Umena, K. Kawakami, J.-R. Shen, PNAS 110 (2013) 3889–38941;
c) M. Suga, F. Akita, K. Hirata, G. Ueno, H. Murakami, Y. Nakajima, T. Shimizu, K. Yamashita, M. Yamamoto, H. Ago, J.-R. Shen, Nature 517 (2014) 99–103, http://dx.doi.org/10.1083/nature13991;
d) B. Loll, J. Kern, W. Saenger, A. Zouni, J. Biesiadka, Nature 438 (2005) 1040–1044;
e) A. Guskov, J. Kern, A. Gabdulkhakov, M. Broser, A. Zouni, W. Saenger, Nat. Struct. Mol. Biol. 16 (2009) 334–342;
f) K.N. Ferreira, T.M. Iverson, K. Maghlaoui, J. Barber, S. Iwata, Science 303 (2004) 1831–1838.

[4] a) W. Liang, T.A. Roelofs, R.M. Cinco, A. Rompel, M.J. Latimer, W.O. Yu, K. Sauer, M.P. Klein, V.K. Yachandra, J. Am. Chem. Soc. 122 (2000) 3399–3412;
b) J.H. Robblee, J. Messinger, R.M. Cinco, K.L. McFarlane, C. Fernandez, S.A. Pizarro, K. Sauer, V.K. Yachandra, J. Am. Chem. Soc. 124 (2002) 7459–S7471;
c) M. Haumann, C. Müller, P. Liebisch, L. Iuzzolino, J. Dittmer, M. Grabolle, T. Neisius, W. Meyer–Klaucke, H. Dau, Biochemistry 44 (2005) 1894–1908;
d) H. Dau, M. Haumann, Coord. Chem. Rev. 252 (2008) 273–295;
e) J. Yano, J. Kern, K. Sauer, M.J. Latimer, Y. Pushkar, J. Biesiadka, B. Loll, W. Saenger, J. Messinger, A. Zouni, V.K. Yachandra, Science 314 (2006) 821–825;
f) C. Glöckner, J. Kern, M. Broser, A. Zouni, V. Yachandra, J. Yano, J. Biol. Chem. 288 (2013) 22607–22620;
g) M. Grabolle, M. Haumann, C. Müller, P. Liebisch, H. Dau, J. Biol. Chem. 281 (2006) 4580–4588;
h) Y. Pushkar, J. Yano, P. Glatzel, J. Messinger, A. Lewis, K. Sauer, U. Bergmann, V. Yachandra, J. Biol. Chem. 282 (2007) 7198–7208;
i) A. Grundmeier, H. Dau, Biochim. Biophys. Acta Bioenerg. 1817 (2012) 88–105;
j) Y. Pushkar, J. Yano, K. Sauer, A. Boussac, V. Yachandra, PNAS 105 (2008) 1879–1884.

[5] S. Luber, I. Rivalta, Y. Umena, K. Kawakami, J.-R. Shen, N. Kamiya, G.W. Brudvig, V. Batista, Biochemistry 50 (2011) 6308–6311.

[6] a) P. Gatt, S. Petrie, R. Stranger, R.J. Pace, Angew. Chem. Int. Ed. 51 (2012) 12025–12028;
b) S. Petrie, R. Stranger, R.J. Pace, Chem. Eur. J. 21 (2015) 6780–6792;
c) S. Petrie, R.J. Pace, R. Stranger, Angew. Chem. Int. Ed. 54 (2015) 7120–7124;
d) R. Terrett, S. Petrie, R.J. Pace, R. Stranger, Chem. Commun. 50 (2014) 3187–3190.

[7] A.R. Jaszewski, R. Stranger, R.J. Pace, Chem. Eur. J. 17 (2011) 5699–5713;
b) A.R. Jaszewski, R. Stranger, R.J. Pace, Phys. Chem. Chem. Phys. 11 (2009) 5634–5642.

[8] a) E.M. Sproviero, J.A. Gascón, J.P. McEvoy, G.W. Brudvig, V.S. Batista, J. Am. Chem. Soc. 130 (2008) 3428–3442;
b) E.M. Sproviero, J.A. Gascón, J.P. McEvoy, G.W. Brudvig, V.S. Batista, J. Chem. Theory Comput. 2 (2006) 1119–1134;
c) P.E.M. Siegbahn, J. Am. Chem. Soc. 131 (2009) 18238–18239;
d) M. Kusunoki, Biochim. Biophys. Acta Bioenerg. 1767 (2007) 484–492;
e) P.E.M. Siegbahn, Biochim. Biophys. Acta Bioenerg. 1827 (2013) 1003–1019;
f) P.E.M. Siegbahn, J. Am. Chem. Soc. 135 (2013) 9442–9449;
g) W. Ames, D.A. Pantazis, V. Krewald, N. Cox, J. Messinger, W. Lubitz, F. Neese, J. Am. Chem. Soc. 133 (2011) 19743–19757;
h) A. Grundmeier, H. Dau, Biochim. Biophys. Acta 1817 (2012) 88–105;
i) K. Yamaguchi, H. Isobe, S. Yamanaka, T. Saito, K. Kanda, M. Shoji, Y. Umena, K. Kawakami, J.-R. Shen, N. Kamiya, M. Okumura, Int. J. Quantum Chem. 113 (2013) 525–541;
j) A. Galstyan, A. Robertazzi, Knapp EW, J. Am. Chem. Soc. 134 (2012) 7442–7449;
k) J. Schraut, M. Kaupp, Chem. Eur. J. 20 (2014) 7300–7308;
l) L. Rapatskiy, N. Cox, A. Savitsky, W.M. Ames, J. Sander, M.M. Nowaczyk, M. Rögner, A. Boussac, F. Neese, J. Messinger, W. Lubitz, J. Am. Chem. Soc. 134 (2012) 16619–16634.

[9] R.J. Pace, L. Jin, R. Stranger, Dalton Trans. 41 (2012) 11145–11160.

[10] D.R.J. Kolling, N. Cox, M. Gennady, G.M. Ananyev, R.J. Pace, G.C. Dismukes, Biophys. J. 103 (2012) 313–322.

[11] a) C. Kupitz, S. Basu, I. Grotjohann, R. Fromme, N.A. Zatsepin, K.N. Rendek, M.S. Hunter, R.L. Shoeman, T.A. White, D. Wang, D. James, J.-H. Yang, D.E. Cobb, B. Reeder, R.G. Sierra, H. Liu, A. Barty, A.L. Aquila, D. Deponte, R.A. Kirian, S. Bari, J.J. Bergkamp, K.R. Beyerlein, M.J. Bogan, C. Caleman, T.-Z. Chao, C.E. Conrad, K.M. Davis, H. Fleckenstein, L. Galli, S.P. Hau–Riege, S. Kassemeyer, H. Laksmono, M. Liang, L. Lomb, S. Marchesini, A.V. Martin, M. Messerschmidt, D. Milathianaki, K. Nass, A. Ros, S. Roy–Chowdhury, K. Schmidt, M. Seibert, J. Steinbrener, F. Stellato, L. Yan, C. Yoon, T.A. Moore, A.L. Moore, Y. Pushkar, G.J. Williams, S. Boutet, R.B. Doak, U. Weierstall, M. Frank, Chapman HN, Spence JCH, P. Fromme, Nature 513 (2014) 261–265;
b) J. Kern, R. Tran, R. Alonso–Mori, S. Koroidov, N. Echols, J. Hattne, M. Ibrahim, S. Gul, H. Laksmono, R.H. Sierra, Gildea RJ, G. Han, J. Hellmich, B. Lassalle–Kaiser, R. Chatterjee, A.S. Brewster, C.A. Stan, C. Glöckner, A. Lampe, D. DiFiore, D. Milathianaki, A.R. Fry, M. Seibert, J.E. Koglin, E. Gallo, J. Uhlig, D. Sokaras, T.-C. Weng, P.H. Zwart, D.E. Skinner, M.J. Bogan, M. Messerschmidt, P. Glatzel, G.J. Williams, S. Boutet, P.D. Adams, A. Zouni, J. Messinger, N.K. Sauter, U. Bergmann, J. Yano, V.K. Yachandra, Nat. Commun. 4371 (2014) 1–11.

[12] a) T.G. Carrell, A.M. Tyryshkin, G.C. Dismukes, J. Biol. Inorg. Chem. 7 (2002) 2–22;
b) K.A. Åhrling, R.J. Pace, in: K. Satoh, T. Wydrzynski, Govindjee (Eds.), Photosystem II: The Light Driven Water:Plastoquinone Oxidoreductase, Vol. 12, Springer, Dordrecht, The Netherlands 2005, pp. 285–305 (Ch).

[13] a) A. Boussac, J.-J. Girerd, A.W. Rutherford, Biochemistry 35 (1996) 6984–6989;
b) A. Boussac, M. Sugiura, D. Kirilovsky, A.W. Rutherford, Plant Cell Physiol. 46 (2005) 837–842.

[14] a) D.W. Randall, B.E. Sturgeon, J.A. Ball, G.A. Lorigan, M.K. Chan, M.P. Klein, W.H. Armstrong, R.D. Britt, J. Am. Chem. Soc. 117 (1995) 11780–11789;
b) J.M. Pelloquin, K.A. Campbell, D.W. Randall, M.A. Evanchik, V.L. Pecoraro, W.H. Armstrong, R.D. Britt, J. Am. Chem. Soc. 122 (2000) 10926–10942;
c) C. Teutloff, S. Pudollek, S. Keßen, M. Broser, A. Zouni, R. Bittl, Phys. Chem. Chem. Phys. 11 (2009) 6715–6726;
d) S. Pudollek, F. Lendzian, R. Bittl, Biochem. Soc. Trans. 36 (2008) 1001–1004;
e) L.V. Kulik, B. Epel, W. Lubitz, J. Messinger, J. Am. Chem. Soc. 129 (2007) 13421–13435;
f) N. Cox, L. Rapatskiy, J.-H. Su, D.A. Pantazis, M. Sugiura, L. Kulik, P. Dorlet, A.W. Rutherford, F. Neese, A. Boussac, W. Lubitz, J. Messinger, J. Am. Chem. Soc. 133 (2011) 3635–3648;
g) V. Krewald, M. Retegan, N. Cox, J. Messinger, W. Lubitz, S. DeBeer, F. Neese, D.A. Pantazis, Chem. Sci. 6 (2015) 1676–1695, http://dx.doi.org/10.1039/c4sc03720k;
h) L. Rapatskiy, N. Cox, A. Savitsky, W.M. Ames, J. Sander, M.M. Nowaczyk, M. Rögner, A. Boussac, F. Neese, J. Messinger, W. Lubitz, J. Am. Chem. Soc. 134 (2012) 16619–16634.

[15] L. Jin, P. Smith, C.J. Noble, R. Stranger, G.R. Hanson, R.J. Pace, Phys. Chem. Chem. Phys. 16 (2014) 7799–7812.

[16] Y. Kurashige, G.K.-L. Chan, T. Yanai, Nat. Chem. 5 (2013) 660–666.

[17] N. Cox, M. Retegan, F. Neese, D.A. Pantazis, A. Boussac, W. Lubitz, Science 345 (2014) 804–808.

[18] a) G. te Velde, F.M. Bickelhaupt, S.J.A. van Gisbergen, C. Fonseca Guerra, E.J. Baerends, J.G. Snijders, T. Ziegler, J. Comb. Chem. 22 (2001) 931–967;
b) C. Fonseca Guerra, J.G. Snijders, G. te Velde, E. Baerends, Theor. Chem. Accounts 99 (1998) 391–403;
c) ADF2014.01, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, 2014;
d) The PyMOL Molecular Graphics System, Version 1.7.0 Schrödinger, LLC.
e) U. Ayachit, The ParaView Guide: A Parallel Visualization Application, Kitware, 2015.

[19] a) J.P. Perdew, Phys. Rev. B 33 (1986) 8822–8824;
b) A.D. Becke, Phys. Rev. A 38 (1988) 3098–3100;
c) A.D. Becke, J. Chem. Phys. 98 (1993) 5648–5652;
d) E. van Lenthe, E.J. Baerends, J. Comb. Chem. 24 (2003) 1142–1156.

[20] a) E.S. Kadantsev, T. Ziegler, J. Phys. Chem. A 112 (2008) 4521–4526;
b) J.P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. 77 (1996) 3865–3868;
c) J.P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. 78 (1997) 3865;
d) C. Adamo, V. Barone, J. Chem. Phys. 110 (1999) 6158–6169.

[21] S. van der Walt, S.C. Colbert, G. Varoquaux, Comput. Sci. Eng. 13 (2011) 22–30.

[22] a) E. van Lenthe, E.J. Baerends, J.G. Snijders, J. Chem. Phys. 99 (1993) 4597–4610;
b) E. van Lenthe, E.J. Baerends, J.G. Snijders, J. Chem. Phys. 101 (1994) 9783–9792;
c) E. van Lenthe, A. Ehlers, E.-J. Baerends, J. Chem. Phys. 110 (1999) 8943–8953.

[23] B.C. Polander, B.A. Barry, PNAS 110 (2013) 10634–10639.

[24] L. Noodleman, J. Chem. Phys. 74 (1981) 5737–5743.

[25] a) S.R. Cooper, M.J. Calvin, J. Am. Chem. Soc. 99 (1977) 6623–6630;
b) K.-O. Schäfer, R. Bittl, F. Lendzian, V. Barynin, T. Weyhermüller, K. Wieghardt, W. Lubitz, J. Phys. Chem. B 107 (2003) 1242–1250;
c) C. Teutloff, K.-O. Schäfer, S. Sinnecker, V. Barynin, R. Bittl, K. Wieghardt, F. Lendzian, W. Lubitz, Magn. Reson. Chem. 43 (2005) S51–S64;
d) G. Blondin, R. Davydov, C. Philouze, M.-F. Charlot, S. Styring, B. Åkermark, J.-J. Girerd, A. Boussac, J. Chem. Soc. Dalton Trans. (1997) 4069–4074;
e) C. Zhang, C. Chen, H. Dong, J.-R. Shen, H. Dau, J. Zhao, Science 348 (2015) 690–693.

[26] a) M.C.W. Evans, K. Gourovsckaya, J.H.A. Nugent, FEBS Lett. 450 (1999) 285–288;
b) M.C.W. Evans, A.M. Rich, J.H.A. Nugent, FEBS Lett. 477 (2000) 113–117;
c) K.A. Åhrling, M.C.W. Evans, J.H.A. Nugent, R.J. Pace, Biochim. Biophys. Acta 1656 (2004) 66–77;
d) D.A. Force, D.W. Randall, G.A. Lorigan, K.L. Clemens, R.D. Britt, J. Am. Chem. Soc. 120 (1998) 13321–13333.

[27] A. Schweiger, G. Jeschke, Principles of Pulse Electron Paramagnetic Resonance Oxford University Press Oxford United Kingdom, 2001 359–403.

[28] L. Hermosilla, J.M. Garcia de la Vega, C. Sieiro, P. Calle, J. Chem. Theory Comput. 7 (2011) 169–179.

[29] M. Zheng, S.V. Khangulov, G.C. Dismukes, V.V. Barynin, Inorg. Chem. 33 (1994) 382–387.

[30] N. Cox, M. Retegan, F. Neese, D.A. Pantazis, A. Boussac, W. Lubitz, Science 345 (2014) 804–808.

Introduction to Paper IV: *Vibrational Intensities in the Mobile Block Hessian Approximation* 7

Einstein argued that there must be simplified explanations of nature, because God is not capricious or arbitrary. No such faith comforts the software engineer.

Frederick Phillips Brooks, Jr.

7.1 Citation

Terrett, R.; Stranger, R.; Frankcombe, T.; Pace, R.J.; *Vibrational Intensities in the Mobile Block Hessian Approximation*; 19 (**2017**) pp. 6654–6664; doi: 10.1039/C6CP07498G

7.2 Introduction

The MBH intensities methodology emerged from the hypothesis that, as vibrational absorption intensities are proportionate to the squared dipole derivative of a given vibrational mode (Equation 7.1), they should be generalisable to the vibrational modes that emerge from a block Hessian calculation in a manner similar to conventional normal mode analysis.

$$A_i \propto \left\| \frac{\delta \vec{\mu}}{\delta \vec{S}_i} \right\|^2 \tag{7.1}$$

Furthermore, the modes associated with nontrivial motions of mobile blocks were anticipated to be quite low in energy and thus well-separated from highly localised diagnostic modes. Therefore, given appropriate selection of mobile blocks, a large number of degrees of freedom can be eliminated from a molecule without necessarily compromising the accurate reproduction of frequencies and intensities of high-frequency

modes. This elimination of extraneous modes would allow for substantial improvements in computational efficiency with respect to all-atom normal mode analysis. We investigated this hypothesis, with reference to the convergence seen in Partial Hessian Vibrational Analysis (PHVA) frequencies and intensities with respect to expansion of PHVA explicit region size, and the lack of improvement of computational efficiency seen in benchmark calculations we performed in the PHVA approximation.

7.3 My contribution to this work

7.3.1 Development of methodology

I performed a program of initial investigation leading to the development of the method described in the manuscript. This began with investigation of the dipole derivative response of heterodiatomic molecules (namely, hydrogen halides) with respect to numerical differentiation step size. This was undertaken to gain intuition as to the regime of atomic displacements in which linearity is maintained. Briefly, numerical integration is beset by two principal sources of error: floating point error and formula error. If a numerical step size is too small, the digital finite-precision representation of real numbers (typically in the form of so-called *floating point* numbers) becomes significant. This is both in terms of the finite size of the mantissa

and the irregular density of real numbers that can be represented on the floating point number line.[a] Alternately, if the step size is too large, the assumption of approximate linearity of the function under study over the step interval may be invalid (see Figure 7.1 for an illustrative example). Progressive refinement and debugging of the technique expanded to larger, non-linear molecules until consistency with conventional normal mode analysis was achieved.

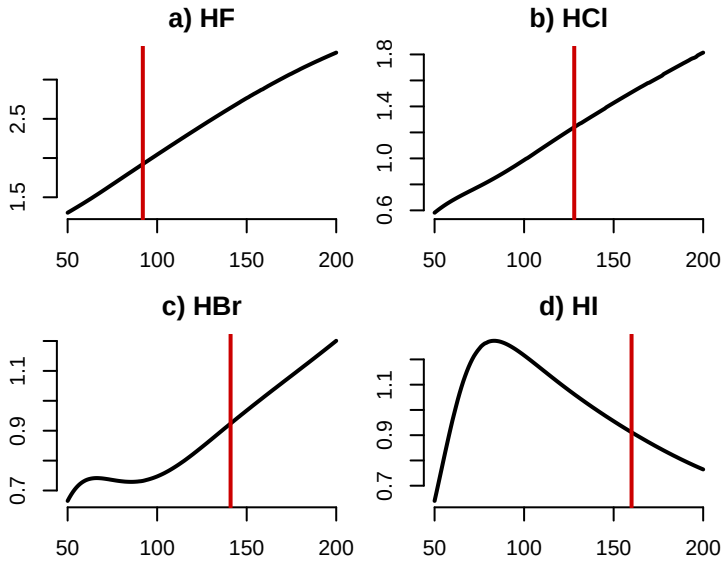


Figure 7.1: Computed dipole moments of **a**) Hydrogen Fluoride, **b**) Hydrogen Chloride, **c**) Hydrogen Bromide, **d**) Hydrogen Iodide with respect to interatomic separation, computed at B3LYP/def2-TZVP level of theory in ORCA. [237] These data demonstrate considerable nonlinearity in response to a single geometric variable. Nominal equilibrium interatomic separations for gas-phase hydrogen halides are indicated in red. [238]

[a] See Goldberg's *tour de force* survey for further information on this enormous topic, which goes largely ignored in computational science. [236]

7.3.2 Implementation of software

I wrote software to coordinate the required calculations to generate vibrational intensities in the MBH formalism. This development was performed in the Python and R programming languages. This software consisted of:

1. Code for extraction of high-resolution normal mode vector data from ADF `TAPE21` files. ADF stores machine-precision normal mode vectors in the `TAPE21` format, for which a parser was written in Python to locate and extract the relevant data and rectify it for use in generation of distorted geometries for dipole derivative calculation.

2. Code for perturbation of molecular geometries, dispatch of ADF calculations and recovery of dipole derivatives from these calculations. This code was written in Python and acts as a wrapper to ADF, perturbing the equilibrium geometry by some multiple of the mass-weighted normal mode, dispatching an ADF calculation as a subtask, and waiting until completion of the calculation. Thereafter, dipole derivatives are numerically obtained by comparing dipole moment vectors in the equilibrium and perturbed geometries in the central difference approximation.

3. Code for convolution, normalisation and visualisation of spectra

thus generated. This software was developed in R [225] and generates a synthetic spectrum from vibrational frequencies and intensities, simulating spectral broadening by modeling each spectral peak as a scaled Dirac function convolved with a Cauchy distribution. From here, difference spectra and integrated intensities can also be derived. This software was additionally ported to the Julia programming language for prospective future use in the iJulia interface. [239]–[241]

7.3.3 Benchmarking of MBH intensities methodology

I benchmarked the MBH intensities technique with respect to a range of molecules under systematic expansion of MBH explicit regions and numeric differentiation step size choice. The species surveyed included small organic molecules with particular diagnostic utility, *keto–enol* tautomers to assess the performance of the technique with respect to FTIR difference spectroscopy,[b] and manganese complexes with thematic relevance to the PSII OEC. These calculations were undertaken to assess the convergent behaviour of the methodology with respect to conventional normal mode calculations, parameterise sources of error in MBH intensities calculations, and assess the computational efficiency of the technique

[b] The simplification of difference spectra being a particular future research target; Chapter 5 discusses our earlier research on all-atom difference spectroscopy calculations involving small molecule models of the PSII OEC.

with respect to normal mode analysis.

7.3.4 Analysis and discussion

I wrote the drafts of the manuscript, contributing analysis, discussion, methodological exposition, theoretical background, illustrations, all data visualisations, and supporting information. Additionally, I implemented reviewer suggestions and amendments.

7.4 Prospective derivative research

At the time of writing, we are preparing a manuscript which details the application of the MBH intensities technique to large crystallographic models of the PSII OEC, as the efficient inspection of difference spectra of residues of interest in crystallographically adapted PSII has been a long-term research goal.

Moreover, as a proof-of-concept, the specific implementation of the MBH intensities technique given in our paper possesses ample opportunities for optimisation, both in terms of ease of use and computational efficiency.

PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2017, **19**, 6654

Vibrational intensities in the mobile block Hessian approximation†

Richard Terrett,[a] Rob Stranger,*[a] Terry Frankcombe[b] and Ronald J. Pace[a]

Received 2nd November 2016,
Accepted 6th February 2017

DOI: 10.1039/c6cp07498g

rsc.li/pccp

Herein we present a proof of concept for the recovery of vibrational intensities from density functional theory vibrational calculations performed in the Mobile Block Hessian (MBH) approximation, which constrains the internal degrees of freedom of designated subsets of a molecule. We compare and contrast the behaviour of this methodology with respect to conventional vibrational calculations, and characterise the performance and accuracy of our method with respect to the size of MBH constrained regions within a variety of species. We demonstrate the viability of this method as a means by which to obtain vibrational intensities for regions of interest within a molecule whilst potentially dramatically reducing computational expense with respect to conventional all-atom vibrational calculations, and discuss caveats for application.

1 Introduction

The modeling of spectroscopic properties of molecules is a major focus of computational chemistry. By exploring the commonality—or lack thereof—between spectroscopic models and experimental results, the structure and behavior of complex chemical systems can be probed.

This is particularly relevant to vibrational spectroscopy, which is able to probe the bonds between atoms. However, the study of the vibrational structure of macromolecules is hampered by the considerable computational expense of calculating the second derivatives of molecular energy with respect to atomic positions (the Hessian matrix, **H**), either analytically or numerically. Concretely, the numerical evaluation of second derivatives in the three-point central difference approximation for a molecule comprised of n atoms in Cartesian space requires $6n + 1$ individual single-point calculations on perturbed geometries, representing positive and negative perturbations of every atom in the molecule along three orthogonal axes. In practice, the potentially onerous nature of vibrational calculations on macromolecules or ligands bound to surfaces or nanoparticles necessitates molecular simulations that adequately compromise between capturing the essential structural and electronic features of the molecule under study at an appropriate level of theory, whilst being computationally feasible on extant hardware.

It is often the case that much of the structure of a molecular model that is required for electronic or structural realism is not of particular interest insofar as vibrational structure is concerned. An example of this is the interest in the spectroscopic response of individual carboxylate moieties coordinating the Mn_4CaO_5 cofactor of Photosystem II (PSII), the holoenzyme responsible for the water oxidation step of oxygenic photosynthesis.[1–5] Within this system, the electronic structure of the cofactor and its attendant coordination environment modulates the geometry of each coordinating carboxylate, however the vibrational structure of the model as a whole is less important than particular diagnostic absorptions localized to individual carboxylates.

In situations where mechanical coupling (so-called Fermi resonance) of vibrational modes exist, artificial isolation of coupled vibrations from one another and observation of the magnitude of attendant level repulsion can provide insight into the relative intensity of this coupling.

Moreover, it is also potentially the case for molecular models that represent some subregion of a larger system (for example, the active site of an enzyme) that constraints may be required at points of truncation for geometric realism. These constraints may give rise to a strictly non-equilibrium structure with high-energy imaginary frequencies. Constraints of this sort are a general feature of calculations peformed on truncated models adapted from crystallographic structures of larger biological systems, where truncation of a model from a larger protein matrix may introduce unrealistic flexibility into a system.

It is for these reasons and others that the class of computational methods commonly referred to as partial-Hessian techniques have been developed. The feature that unifies these methodologies is the treatment of the vibrational structure for only part of a molecule. Typically, this process involves freezing the degrees

[a] *Research School of Chemistry, Australian National University, Canberra, Australia. E-mail: rob.stranger@anu.edu.au*
[b] *School of Physical, Environmental & Mathematical Sciences, UNSW Canberra, Australian Defence Force Academy, Canberra, Australia*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c6cp07498g

of freedom in regions of a molecule that are not interesting to the researcher. The most elementary of these methods is Partial Hessian Vibrational Analysis (PHVA), which neglects a nominated single region of a molecule from consideration whilst calculating the second derivatives of all other atoms. This is achieved by conferring an effective infinite mass upon atoms in the neglected region. This region need not be contiguous on the molecular graph, however only one region is permitted. Whilst the technique grew out of the modeling of the vibrational structure of adsorbates, it has since been generalized.[6] The Vibrational Subsystem Analysis (VSA) technique is thematically related but includes inertial contributions of the neglected region.[7]

An elaboration on PHVA is the Mobile Block Hessian (MBH) method, which subsumes nominated regions of atoms into so-called 'mobile blocks', which have no internal degrees of freedom but 6 external degrees of freedom (namely, three orthogonal rotations and translations in Cartesian space).[8,9] Arbitrary numbers of mobile blocks may be specified, and these blocks can undergo translation or rotation with respect to each other. The computational complexity of MBH is predictable, requiring $6n' + 12m + 1$ single-point calculations for a molecule comprised of n' free atoms and m mobile blocks in the three-point stencil approximation. Depending on the number of atoms contained within mobile blocks, the method may afford significantly shorter calculation times, which is to point out that any block containing more than 2 atoms will result in fewer required single-point sub-calculations. An illustration of the reduction of degrees of freedom in an MBH calculation is given in Fig. 1. A specialised predecessor of the MBH approximation is the Rotation–Translation Blocks (RTB) approximation, which aims to capture only the low-frequency vibrational structure of macromolecules by a coarse-grained approximation of a macromolecule as a series of blocks.[10,11] Ghysels and coworkers provide a theoretic and practical comparison of the PHVA, MBH and VSA techniques.[12]

A recent innovation in the field of approximate vibrational analysis is the development of the mode-tracking and intensity-tracking techniques, which take a different approach with respect to partial Hessian approaches. Mode-tracking efficiently obtains approximate vibrational structure by identifying and refining modes with significant interatomic displacement, proceeding from a guess vibrational structure.[13] Intensity tracking produces spectra capturing significant vibrational intensity without recourse to a guess spectrum.[14–16] Mode- and intensity-tracking are qualitatively distinct from partial- or block-Hessian techniques as they do not invoke the freezing of degrees of freedom, however they do offer substantial potential for acceleration of characterising the diagnostic frequencies of large molecules.

At the time of writing, no software packages were identified that furnished vibrational intensities for MBH calculations.‡ As such, we have implemented our methodology for obtaining these intensities, adapting the general principle of numerical differentiation of dipole derivatives with respect to normal mode displacements in the MBH approximation. We have developed this methodology with an eye to rationalising the vibrational difference structure of metal–ligand interactions in metalloproteins such as Photosystem II, wherein multiple peptide carboxylates and imidazoles ligate a $CaMn_4O_5$ cofactor. This methodology facilitates the vibrational investigation of expansive crystallographically adapted structural models where simulation of the complete vibrational structure would constitute an otherwise prohibitive computational expense on extant computational infrastructure.

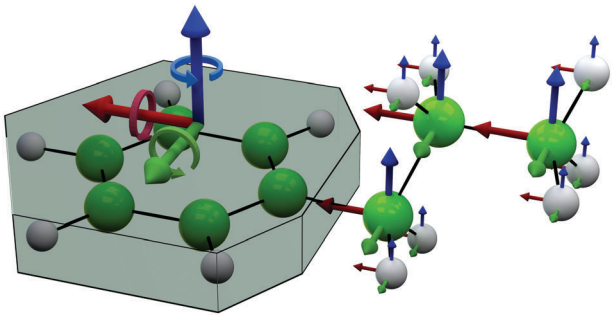


Fig. 1 Schematic representation of MBH region partitioning and the translational and rotational degrees of freedom that arise in this formalism. The volume enclosing the aromatic ring represents a mobile block, in which all internal degrees of freedom are frozen. Due to the lack of rotational symmetry of the block arising from its internal structure, it possesses 3 orthogonal translational and rotational degrees of freedom, indicated by the superimposed coordinate system. All other atoms possess the conventional translational degrees of freedom in Cartesian space indicated.

2 Theory

2.1 Computational methodology

Calculations were performed in the Amsterdam Density Functional (ADF) suite, version 2014.01.[17–19] All calculations were performed using the Becke–Perdew combined exchange/correlation (XC) functional.[20,21] All calculations were performed using an all-electron triple-zeta Slater-type basis set with one polarization function (TZP).[22] No relativistic correction was applied to the electronic energy. Where species were anticipated to possess unpaired spin density, their electronic structures were evaluated in the unrestricted Kohn–Sham formalism. The evaluation of open-shell singlet states (which is to say, antiferromagnetically coupled configurations with $M_S = 0$) was carried out with deliberate polarization of the initial spin density using the ModifyStartPotentials method available in the ADF software.

2.2 Geometry optimizations

Geometries were optimized using a gradient convergence criterion of 1×10^{-4} E_h Å^{-1}. Geometry convergence was performed in two steps, the first using the ADF default convergence criteria and the latter refining this structure with a 10-fold tighter gradient threshold. This tight convergence criterion was chosen to elicit

‡ Whilst our benchmarking of the PHVA methodology as implemented in QChem 4.1 afforded vibrational intensities, it was found that this implementation did not provide any discernible computational advantage over full frequency calculation, only offering the freezing of internal degrees of freedom in a nominated region of a molecule. This functionality is useful for vibrationally isolating areas of interest in a molecule as well as permitting non-equilibrium vibrational analysis, but does not provide the computational acceleration afforded by the MBH methodology.

high-quality minimum energy geometries, reducing the likelihood of negative eigenvalues in the Hessian.

2.3 Mobile block Hessian calculations

MBH calculations were performed to obtain frequencies and normal modes within the MBH formalism, for use in further dipole derivative evaluations to generate corresponding absorption intensities. All calculations used a single mobile block region, regardless of whether the atoms comprising the blocks were contiguous in the molecular graph.

2.4 Region definitions and notation

MBH regions are denoted within the paper by numbers assigned to atoms in each molecular figure. For any given molecule, MBH region R will define all atoms labelled with numbers equal to or less than R as free, with the remainder of the molecule subsumed into a single mobile block. Thus, the definition of successive MBH regions entails a building up of the number of independently mobile atoms in the MBH calculation with respect to its predecessors, and therefore successive increase in the degree of model completeness with respect to a conventional unconstrained frequency calculation.

2.5 Calculation of MBH intensities

We have developed a methodology for recovering the vibrational intensities corresponding to molecular normal modes in the MBH formalism. Noting that the intensity of a vibrational mode is proportionate to the square of the dipole derivative associated with that normal mode distortion,[23] and that the normal mode atomic displacements are reported in the output of ADF MBH calculations, we implemented a program in Python that determines dipole derivatives of a molecule $\delta\vec{\mu}/\delta\vec{S}$ with respect to normal mode coordinates by means of the finite difference method. Concretely, our code dispatches single-point calculations on geometries perturbed along the normal mode coordinates of the vibrational spectrum and records the corresponding change in dipole moment vector.§ A schematic depiction of the MBH intensity recovery methodology is given in Fig. 2.

2.6 Algorithmic description

In the first instance, the vibrational structure of an optimized molecular geometry is evaluated in the MBH approximation, which yields a list of frequencies, a list of nominal atomic masses by which the Hessian matrix is weighted to give the aforementioned frequencies, and a set of normal mode vectors. The intensity A_i of a vibrational mode arising from eigenmode $\vec{S}_i$ of the Hessian is proportional to the square of the derivative of the molecular dipole with respect to atomic displacements along that mode:

$$A_i \propto \left\| \frac{\delta\vec{\mu}}{\delta\vec{S}_i} \right\|^2$$

§ We chose numerical differentiation in Cartesian space owing to the conceptual simplicity of the method, the availability of Cartesian normal mode displacements in high resolution, the lack of atomic polar tensor (APT) data in the MBH output, and the lack of access to ADF internals or source code.

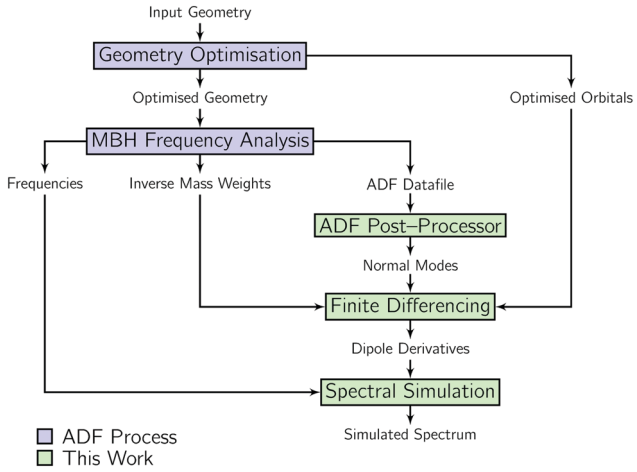


Fig. 2 Flow diagram representing the methodology used to obtain MBH intensities.

Our post-processor extracts these normal mode vectors and performs numerical differentiation of the dipole derivative $\delta\vec{\mu}/\delta\vec{S}_i$ with respect to the normal mode coordinate of the ith vibrational mode. Numerical differentiation is performed with a suitably small step size in the central difference approximation. The normal mode vectors are inverse mass-weighted using the atomic masses previously obtained. We use a conservative multiplicative coefficient of 1×10^{-3} times each displacement vector, where the l^2-norm of each displacement vector produced by the ADF software is ≤ 1.0 Å, however we observed a surprising stability over a range of step sizes from 1×10^{-5} to 5×10^{-1}.

Finally, having obtained a list of vibrational frequencies and dipole moments, the data are converted to vibrational intensities subjected to convolution with a Cauchy–Lorentz kernel to simulate spectral broadening. More detailed implementational detail with respect to the ADF suite is provided in ESI,† S1.

2.7 Post-processing of spectral data and molecular visualisation

Convolution, integration, subtraction and visualization of simulated spectra was performed using scripts written in the R language for statistical computing.[24] Spectral broadening is modeled by the effective convolution of scaled impulse (Dirac) functions convolved with a Cauchy–Lorentz kernel with full width at half maximum (FWHM) of 25 cm^{-1}. The spectral domain ranges from 0 to 4000 cm^{-1} with a granularity of 1 cm^{-1}. Spectra are individually normalized to a maximum intensity of 1.0. Where applicable, integrated spectra are computed numerically by the method of trapezoids, with a granularity of 1 cm^{-1}. Rotational and translational eigenmodes are suppressed by eliminating the six modes with the lowest absolute frequency. We elaborate on the technical aspects of this in our Discussion section. Molecular visualisation was achieved in the PyMol molecular graphics suite and the Blender 3D graphics suite.[25,26]

3 Results

We opted to test the vibrational response of a number of organic species (4-phenylbutanoic acid, heptane, (*Z*)-1-pentenol and pentanal) as well as mononuclear and dinuclear coordination complexes with respect to systematic variance of the MBH region size.

3.1 4-Phenylbutanoic acid

Calculations were performed on 4 different MBH regions of 4-phenylbutanoic acid. This molecule was chosen because it bears three features relevant to systematic analysis of MBH intensity recovery:

(1) A carboxylic acid residue, anticipated to bear symmetric and antisymmetric stretch features in a similar region (1000–2000 cm^{-1}) to carboxylate anions bound to metal centers, such as aspartate residues coordinating the Mn_4CaO_5 active site of Photosystem II. In practice, this is indeed the case.

(2) An aromatic ring, expected to add rich general vibrational structure to the molecule.

(3) An aliphatic chain affording a variable number of methylene repeat units within the MBH region.

Scissor motions of the protons of these methylene units typically couple to carboxylate symmetric stretches. The MBH regions of this molecule were defined as depicted in Fig. 3.

Salient vibrational features were identified in the vibrational spectrum of MBH region 1, including an OH out-of-plane wag (A) around 620 cm^{-1}, a carboxylic acid α-carbon umbrella mode (B) around 760 cm^{-1}, carboxylic acid Cα–Cβ stretching with Cα–OH stretching character and OH wag character (C) around 1100 cm^{-1} roughly analogous to the carboxylate symmetric stretch, a C–OH stretch (D) around 1290 cm^{-1}, and a C═O stretch (E) around 1730 cm^{-1}, analogous to the carboxylate antisymmetric stretch.¶ Aliphatic proton stretches are found in the region of 3000 cm^{-1} (F), with aromatic proton stretches slightly upfield. Finally, the O–H stretch (G) presents at 3600 cm^{-1}. Vibrational structures corresponding to MBH regions 1–4 are presented in Fig. 4.

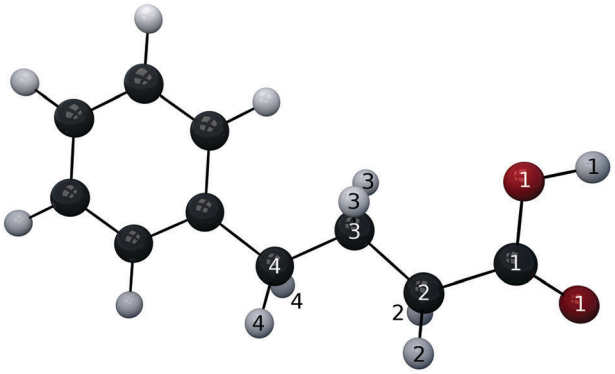


Fig. 3 4-Phenylbutanoic acid model used in MBH study.

¶ Both of these vibrations are not truly symmetric due to the desymmetrizing influence of the attached proton.

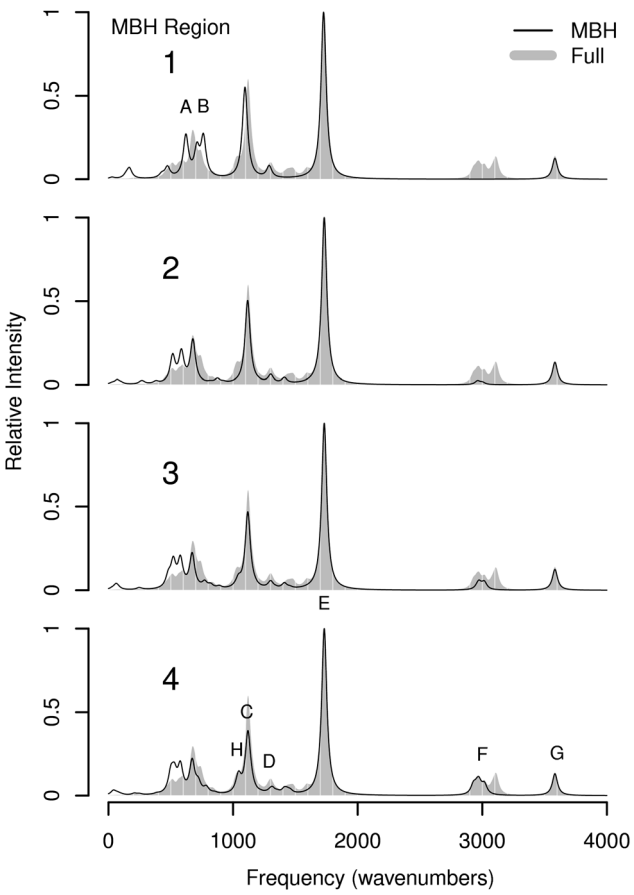


Fig. 4 Vibrational structure of 4-phenylbutanoic acid as a function of MBH region size. Note the appearance of aliphatic and aromatic C–H stretches as region size increases.

In practice, rapid convergence is seen in the carboxylic acid diagnostic region, with peak positions and intensities for the symmetric and antisymmetric stretches as well as the O–H stretch accurately reproduced using a near-minimal MBH region including only the β carbon and substituent protons. The total intensity of the major peak around 1100 cm^{-1} is imperfectly reproduced and indeed decreases with respect to the nominal intensity as a function of MBH region size. The cause of this behavior is unknown, although the appearance of a composite peak (H) around 1030 cm^{-1} replicating the full spectrum suggests that enhanced mechanical resonance with respect to the increased MBH region size elicits an interchange of relative intensity: the modes comprising this peak do exhibit a small degree of –OH wag character.

3.2 Heptane

Heptane was selected as a benchmark species in light of its simple linear aliphatic structure and low variety of bond types, which was predicted to exhibit smooth convergence with respect to MBH region size. The heptane molecule used for benchmarking was optimized in the all-*trans* conformation with no formal symmetry but pseudo-C_{2v} symmetry. The MBH region definitions are given in Fig. 5. The progression of the

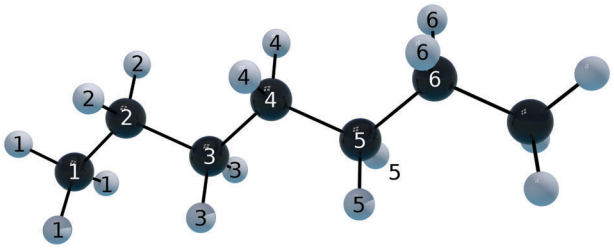


Fig. 5 Pseudo-C_{2v} heptane structure used in MBH analysis.

calculated spectra with respect to the size of the MBH region is quite straightforward, with acceptable convergence behavior (Fig. 6). However, the treatment of the data and its implications for convergence bears some clarification. The vibrational structure of heptane is relatively sparse on account of most of the intensity being present in the C–H stretch region (A) around 3000 cm^{-1}, where the intensity of a large number of comparable methyl and methylene C–H stretches are aggregated into a composite peak. Due to the normalization of the calculated MBH spectra with respect to the full spectrum, the MBH intensity is initially substantially inflated with respect to the full spectrum. This is simply a function of the reduced number of unfrozen C–H bonds in the MBH structures with respect to the full spectrum. The inflation caused by this normalization affects the entire spectrum, and therefore it is observed that peaks down-field of the C–H stretch region are overestimated but converge to the nominal spectrum as the MBH region size (and thus the C–H stretch integrated intensity) increases. Note however that the initially simple vibrational structure of the $\leq$1500 cm^{-1} region is elaborated and peak intensity interchange occurs as the MBH region is expanded. MBH region 5 is anomalous in that the intensity of features in the $\leq$1500 cm^{-1} region is significantly larger that either MBH region 4 or 6, although the relative intensity with respect to the C–H stretch region remains negligible.

3.3 Simulation of difference spectra of keto–enol tautomerism

Calculations were performed to assess the convergence of infra-red difference spectra associated with a simple chemical reaction with respect to MBH region size. Tautomerization of (*Z*)-1-pentenol to pentanal (depicted in Fig. 7a and b) is associated with disappearance of a highly localized O–H stretch (A) around 3720 cm^{-1}, which is reproduced perfectly at all MBH levels, as well as the appearance of a similarly well-localized aldehydic C–H stretch (B) at 2760 cm^{-1} (Fig. 8). The less localized carbonyl C=O stretching mode with attendant minor aldehydic C–H and Cα flexion motion that appears (C) at 1720 cm^{-1} on isomerization is well-reproduced from the second MBH region onwards. It is evident that for this tautomerism it is only necessary to include the alpha carbon and attached protons to achieve essential correspondence with the complete difference spectrum in the $\geq$1000 cm^{-1} region. The quality of correspondence between the MBH difference and complete difference spectra of the tautomerism is extraordinary given that the combined error in both the keto and enol direct

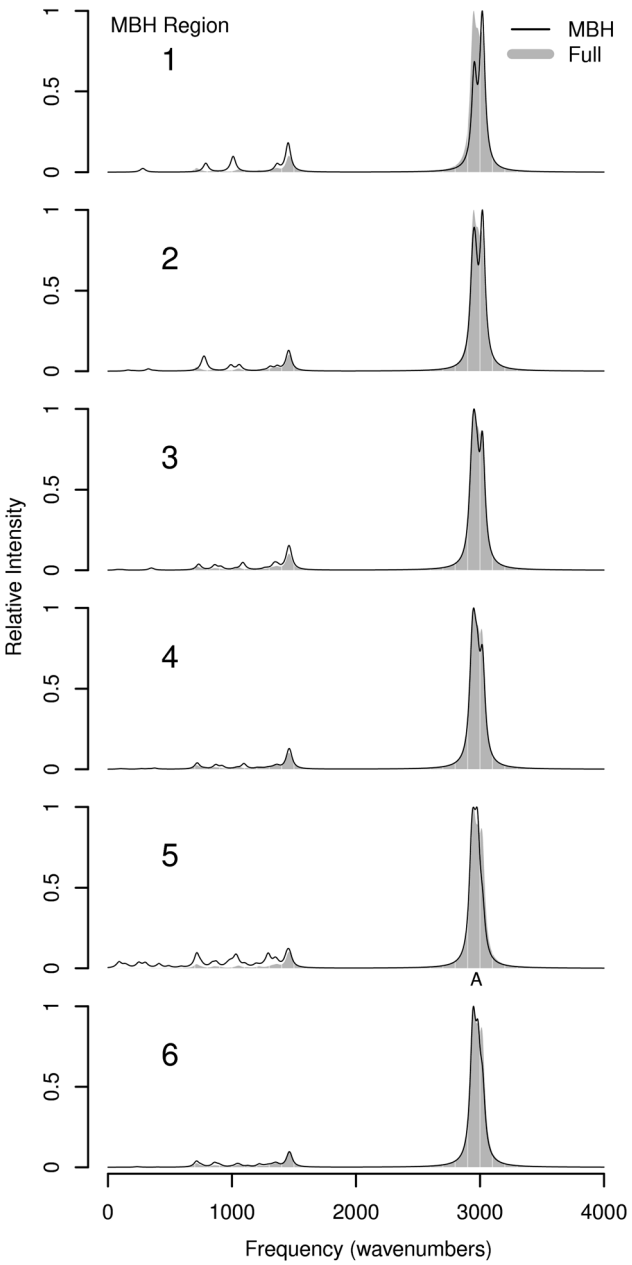


Fig. 6 Vibrational structure of *n*-heptane as a function of MBH region size.

spectra will propagate into the derived difference spectrum. Moreover, the technique appears to be agnostic to the differing number of atoms in MBH regions 1 and 2 between tautomers. Relative difference intensities in the carboxylate diagnostic region are not reproduced perfectly.

3.4 Salpn complex

The salpn complex[27] of Martínez *et al.* was chosen as a benchmark species as it is a mononuclear transition metal complex with non-trivial ligation. It consists of a Mn^{III} ion coordinated by a tetradentate salpn ligand and a pentanoate ligand. The high-spin d^4 electron configuration of the neutral complex

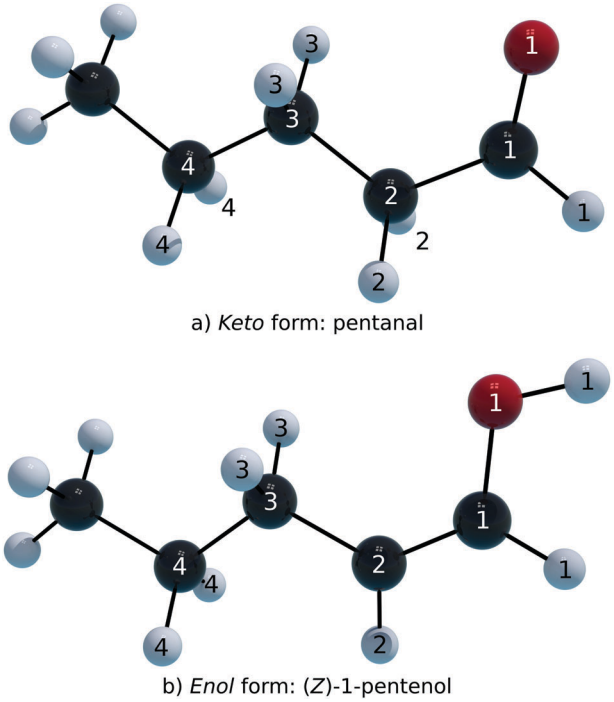


Fig. 7 (a) (*Z*)-1-Pentenol and (b) pentananl, related by keto–enol tautomerization.

implies Jahn–Teller activity, and this is evident in the dissimilar Mn–O bonds of the coordinating pentanoate ligand (202 pm and 243 pm). The construction of the MBH regions of the complex beings from a three-atom region comprised of the carboxylate oxygens and carbon, followed by incremental addition of aliphatic methylene units of the pentanoate and adjacent atoms in the molecular graph beginning from the manganese ion, followed by the salpn phenolic oxygens and secondary amine nitrogens and finally *ipso* carbon atoms in the final region. These region definitions are depicted in Fig. 9.

The vibrational structure thus obtained (Fig. 10) only shows evidence of convergence with the largest region size. This is attributable to the coincidence of the phenolic oxygen–carbon stretches and secondary amine nitrogen–carbon stretches with the carboxylate symmetric and antisymmetric stretch region. Also notable is the sudden appearance of a dominant peak (A) around 920 cm^{-1} in region 3: this mode consists of collective metal–ligand motion and is not present in the largest MBH region, suggesting that it is an artifact of the MBH boundary conditions. Particularly, this mode has a large phenolic C–O stretching component due to the frozen nature of the phenolic carbon atoms. Examination of atomic motions in this frequency regime in region 4 indicates a large number of modes in the approximate range 600 cm^{-1} to 1000 cm^{-1} with relative motions of the nitrogen, phenolic oxygen and salpn ligand backbone atoms with respect to the Mn ion. Notably, modes in this domain do not exhibit large relative motions of carbon and oxygen atoms, which instead are found (B) around 1350 cm^{-1}. A plausible inference is that MBH region 3 does not afford sufficient degrees of freedom to these atoms to furnish a large

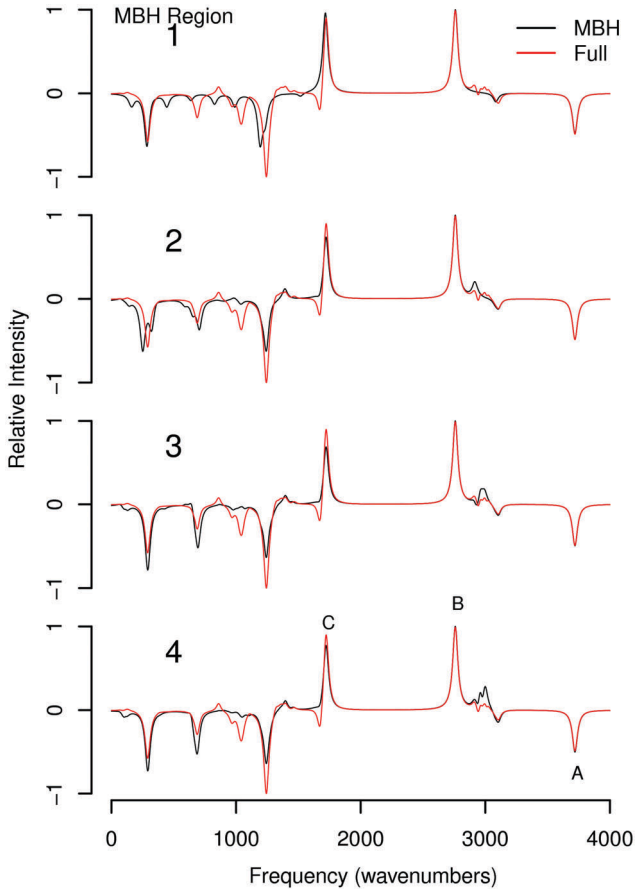


Fig. 8 Difference spectra for keto–enol tautomerism as a function of MBH region size.

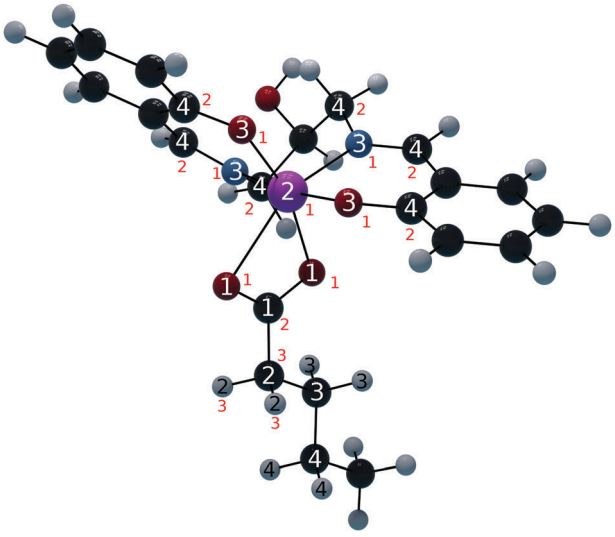


Fig. 9 Salpn complex. Form A region definitions are indicated in black and white, form B region definitions are indicated in red.

number of normal modes of disparate phase distributed over a large region of frequency space.

We note the accumulation of strong dipole derivative contributions in the carboxylate signature region of the salpn complex.

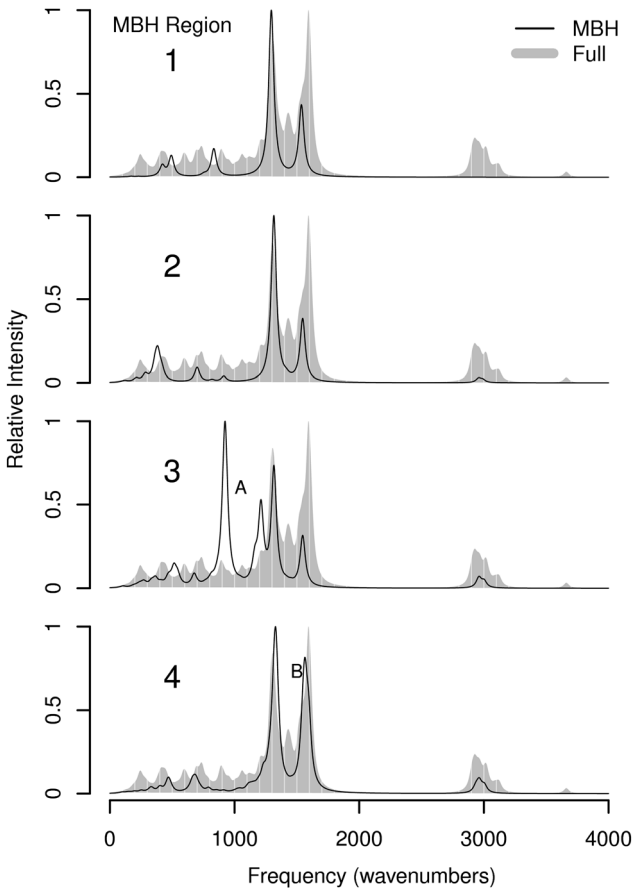


Fig. 10 Vibrational structure of salpn complex as a function of MBH region size.

These contributions are not well-modeled at small region sizes due to the origin of the MBH explicit region being centered on the pentanoate carboxylate moiety. We investigated whether an alternative, metal-centered origin for the salpn MBH explicit region results in faster convergence with respect to systematic contraction of the block. The region definitions and simulated spectra arising from these calculations are reported in Fig. 9 and 11 respectively. This redefinition results in significantly accelerated convergence in the carboxylate diagnostic region, qualitatively reproducing the two major peaks within region 2, which contains the metal center, nitrogen and oxygen atoms ligated to the metal center, relevant salpn carbons and the aliphatic carbonyl and α methylene. Unfortunately, this region definition is almost identical to region 4 in the initial definition, and thus does not confer a great advantage in terms of MBH region size, however it is illustrative as to the effects of region definition choice on the recovered spectrum.

3.5 Dinuclear manganese complex

We selected a Mn_2O_2-centered complex (Fig. 12) that is coordinated by a hexanoate moiety in a μ_1:μ_1 fashion. The coordination of the antiferromagnetically paired oxo-Mn^{III}/Mn^{III} unit is completed by coordination with diethylenetriamine (DETA) ligands. MBH spectra with respect to region size are presented in Fig. 13. Rapid approximate convergence is noted in the

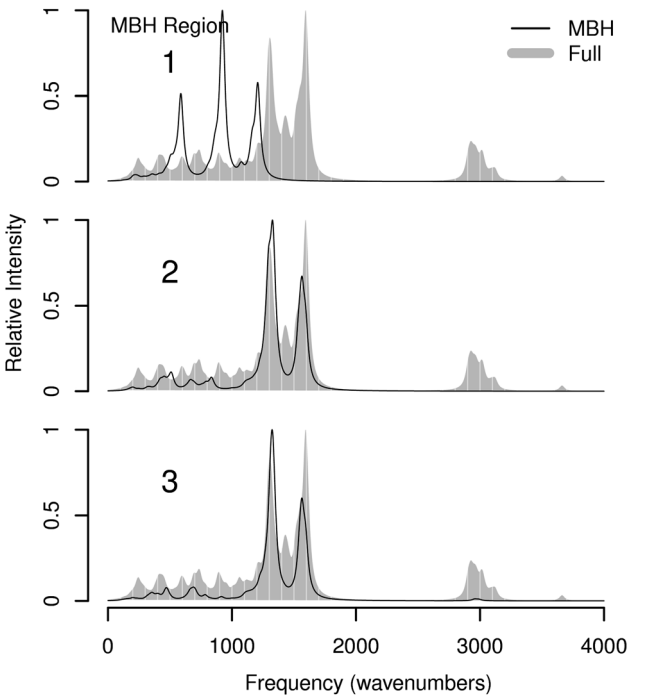


Fig. 11 Vibrational structure of salpn complex as a function of MBH region size, using form B region definitions.

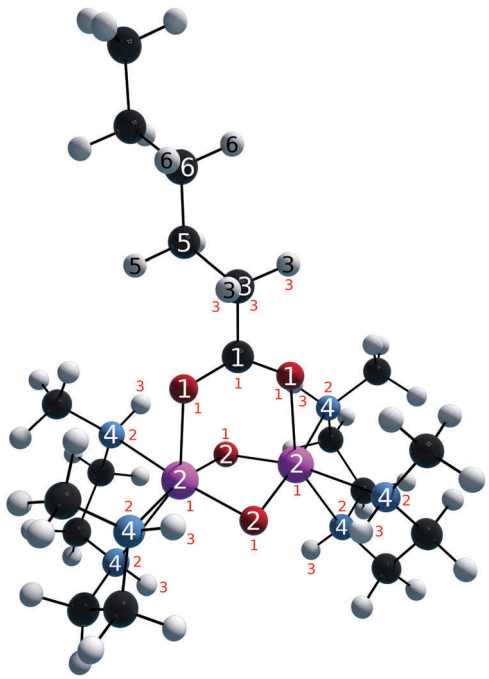


Fig. 12 Dinuclear Mn complex. Form A region definitions are indicated in black and white, form B region definitions are indicated in red.

carboxylate symmetric and antisymmetric stretch region, whilst low-frequency behavior is inconsistently modeled and absorption intensity attributable to modes localized on the DETA ligands is predictably absent on account of its non-inclusion in any region definition.

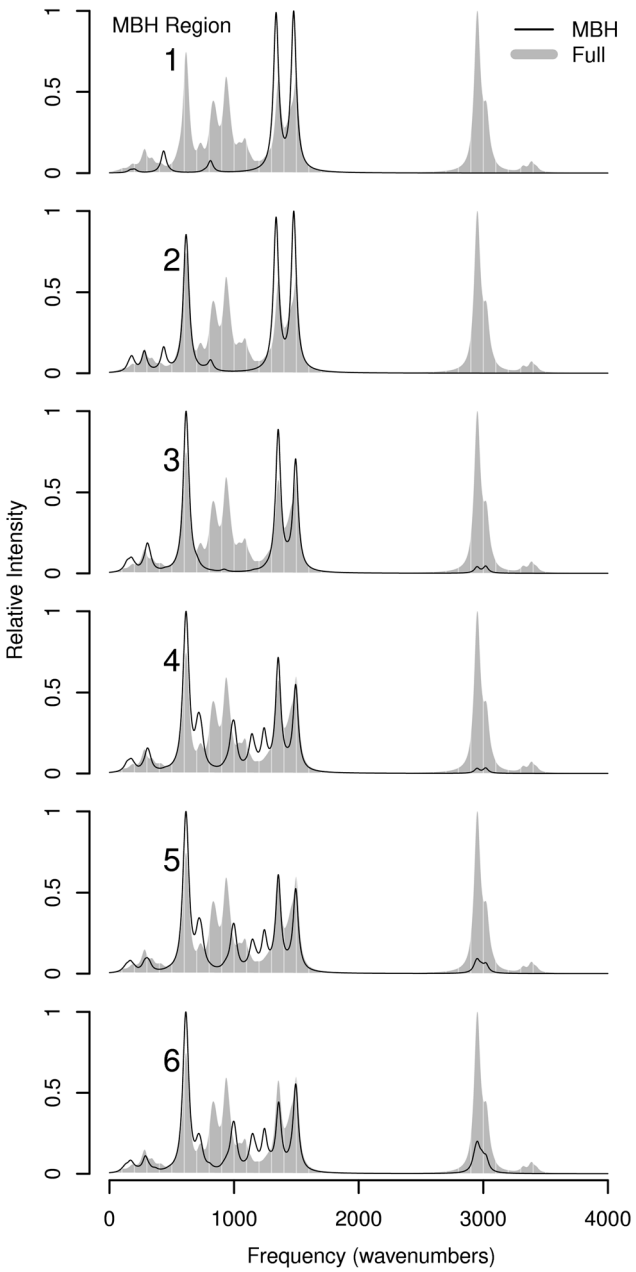


Fig. 13 Vibrational structure of dinuclear manganese complex as a function of MBH region size.

Selecting alternate MBH regions centered on the oxo-manganese butterfly results in acceptable convergence with respect to the full system, as depicted in Fig. 14. It should be noted that naïve normalization of the spectra distorts the similarity of the MBH and full spectra in the smallest region due to the large number of co-located dipole derivative contributors in the C–H stretch region.

3.6 Oxidation–reduction difference spectroscopy in dinuclear manganese complexes

A confounding factor in the direct comparison of MBH and full vibrational spectra of the previously described dinuclear

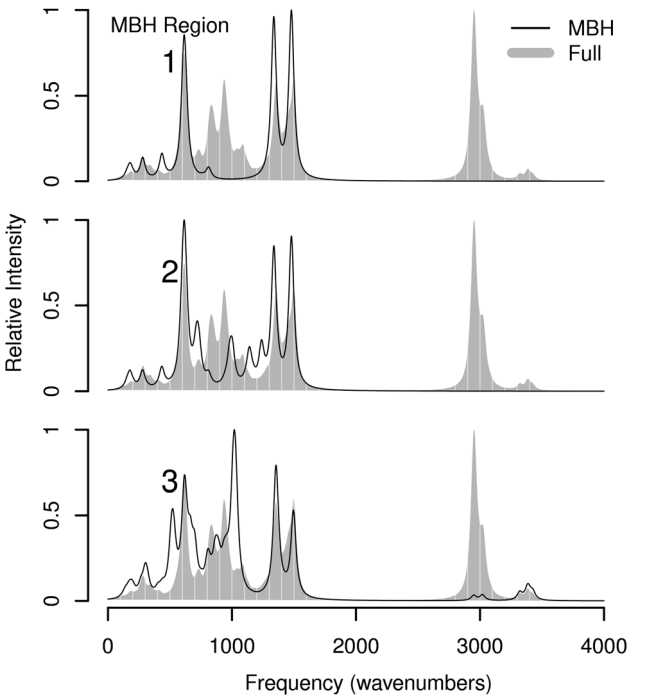


Fig. 14 Vibrational structure of dinuclear manganese complex as a function of MBH region size, using form B region definitions.

manganese complex is the colocation of vibrational intensity in the carboxylate diagnostic region arising from unrelated moieties. As such, it is instructive to instead examine the degree of correspondence between MBH and full difference spectra corresponding to an oxidation state change in one of the Mn ions of the complex, the character of which should be preserved irrespective of the total integrated intensity within the approximation of any particular MBH region.

Reduction of one of the Mn ions in the dinuclear complex from Mn^{III} to Mn^{II} produced difference spectra illustrated in Fig. 15. Within the carboxylate difference spectrum diagnostic region 1100–1700 cm^{-1} (after Debus *et al.*), we find an excellent correspondence between the difference spectra within various MBH regions and that of the full molecule, with evident near-quantitative reproduction of peak shifts associated with the carboxylate symmetric and antisymmetric stretches. Outside of this region (not depicted) correspondence is poor. This is unsurprising given consideration of the conformational changes in the DETA ligands arising from the deactivation of a Jahn–Teller axis. As the DETA ligands are peripheral to the origin of our carboxylate-centered MBH region definitions, they are truncated even in the most expansive MBH models.

3.7 Timing analysis

It is readily observable that the MBH methodology as implemented in ADF is capable of significantly reducing calculation runtime with respect to all-atom numerical frequency calculations, by reducing the number of atomic displacements that need be computed. The number of single-point calculations comprising the MBH calculation can be calculated *a priori*, and

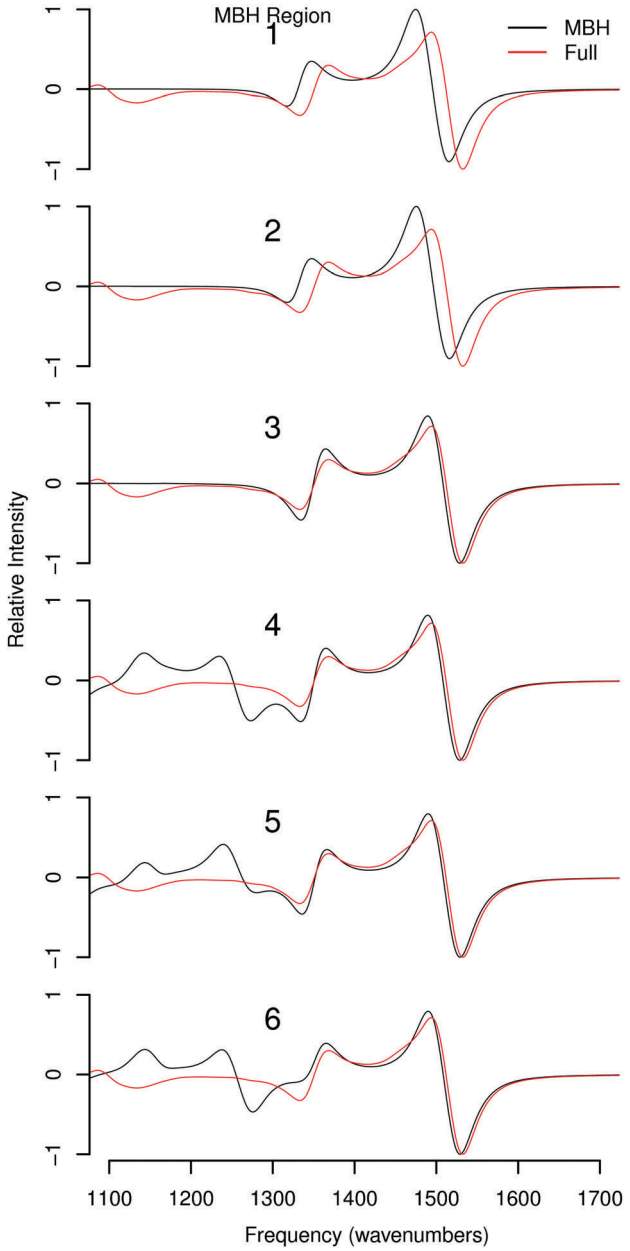


Fig. 15 Difference intensity of dinuclear manganese complex undergoing reduction at a Mn center as a function of MBH region size, using form A region definitions.

therefore the expected computational expense of an MBH calculation with respect to the corresponding full calculation can be estimated with high confidence. Calculation of the dipole derivatives of the MBH calculation in the central difference approximation requires an additional $2i$ single point calculations, where i is the number of normal modes calculated by the MBH calculation. i is a function of the number of free atoms and mobile blocks defined by the MBH calculation. The total runtime of the MBH intensity calculation is therefore the sum of the MBH calculation determining the normal modes, and the finite differencing step deriving the corresponding dipole derivatives.

In sufficiently large systems, MBH intensity calculations provide significant computational advantage over the corresponding full frequency calculation, with a strongly linear relationship between MBH region size and calculation runtime as shown in Fig. 16a, demonstrating relative MBH intensities performance for the dinuclear Mn complex, consisting of 75 atoms. Linear regression indicates that the MBH intensity methodology is intrinsically less efficient than the corresponding full

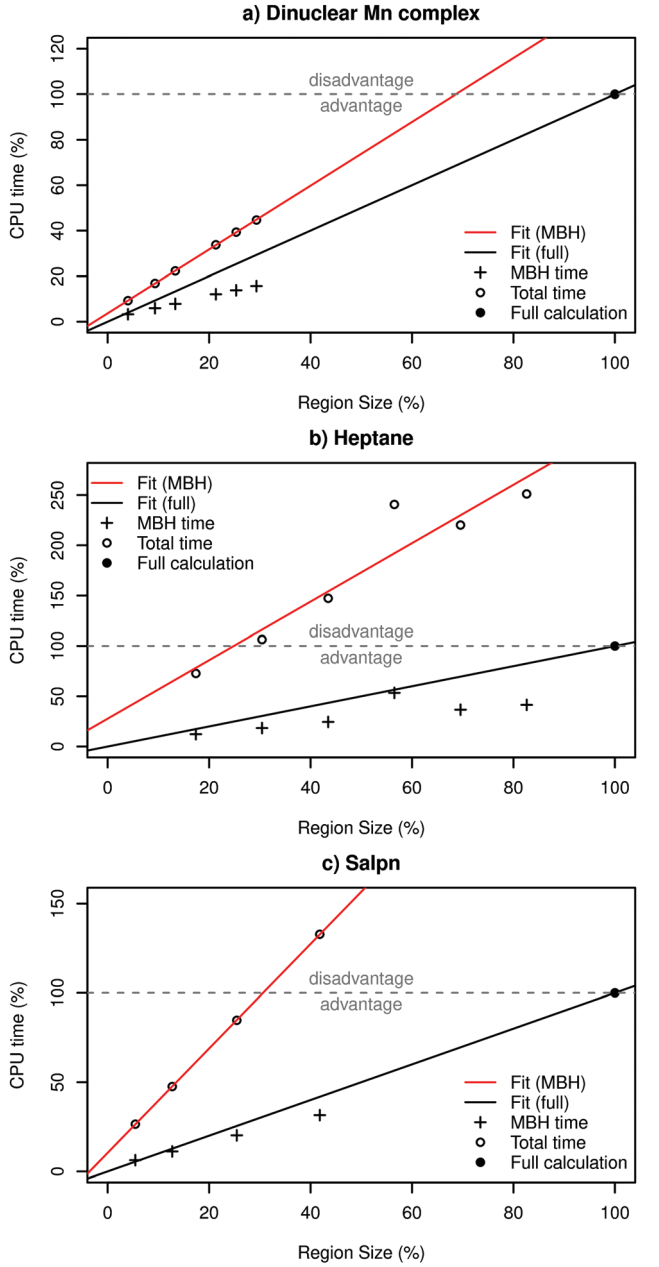


Fig. 16 Relative timing results as a function of MBH region size. Calculations on the dinuclear Mn complex and salpn complex ((a) and (c), respectively) were run on 16 concurrent threads, for both the MBH normal mode and intensity steps. Calculations on heptane (b) analogously employed 4 concurrent threads. Discussion of anomalous CPU-time behaviour is given in ESI,† S2. 'Fit' indicates a linear model regression through the relevant data.

frequency calculation. Therefore, as the MBH region size approaches parity with the full calculation, the MBH intensity methodology is liable to become computationally disadvantageous. For small molecules the overhead of the MBH intensities technique is unlikely to afford any computational advantage, as seen in the case of heptane (Fig. 16b, consisting of 23 atoms). Intrinsic overhead in the MBH normal mode calculation for heptane is compounded by the derivative calculation, producing disappointing performance. Concretely, only the smallest MBH region—consisting of a single methyl moiety—produces a computational advantage over the full calculation. The salpn complex (Fig. 16c, 55 atoms) appears to interpolate this behavior, producing clear albeit underwhelming advantages in runtime with the exception of the largest MBH region.

We find that the discriminant for MBH intensity calculation efficiency with respect to the full numeric frequency calculation is—to a first-order approximation—a function of the number of single-point calculations implicated in the MBH and finite difference calculations. Assuming the central difference approximation and a single mobile block, this number is $12n' + 24$, where n' is the number of free atoms, compared to $6n$ for the conventional frequency calculation, where n is the total number of atoms. Employing a forward difference approximation in the calculation of dipole derivatives would reduce the expected computational expense of the MBH intensities methodology by approximately 50%. The computational overhead exhibited by these calculations is a function of the specific post-processing technique adopted for this proof of concept, however more efficient methods—such as integrated evaluation of dipole derivatives during the MBH calculation—are possible and are anticipated to produce efficiencies that are always advantageous with respect to complete vibrational analyses.

Finally, we found that preconditioning our MBH and intensities calculations by reuse of SCF data between sub-calculations was a prerequisite of efficient calculation with respect to conventional vibrational analysis.

4 Discussion and conclusion

4.1 Low frequency behavior in the MBH formalism

It is noted that the behavior of MBH spectra in both peak position and intensity is relatively unreliable in the $\nu \leq 1000$ cm^{-1} regime. This behavior is attributable to the inadequate modeling of collective modes (such as flexion and breathing modes) with delocalized atomic motions under the imposition of mobile block constraints. Contrariwise, peak intensities are well-replicated in the $\nu \geq 1000$ cm^{-1} regime, although this realism is dependent on attention to the suitability of MBH region definitions.

4.2 Effect of MBH region size around regions of interest

We are particularly interested in measuring the effect of MBH size on the vibrational structure of ligands coordinating metal centers. This work was particularly inspired by the inference of ligand denticity on the basis of vibrational structure by Martínez *et al.*[27] As such, our benchmarks have been constructed with some focus on the behavior of carbonyl moieties with respect to MBH regions containing some number of aliphatic methylene repeat units. This focus is especially concerned with the perturbation of carboxylate symmetric and antisymmetric stretches by Fermi resonance with CH_2 and OH flexions that are relatively proximate in the molecular graph.

We find in the case of aliphatic chains terminating in carboxylate moieties that surprisingly few methylene groups need be explicitly modeled to reach an excellent degree of convergence with respect to a full frequency calculation. This observation has been made with respect to the convergence of carboxylate symmetric and antisymmetric stretch frequencies, however we find that this behavior is generalizable to vibrational intensities.

4.3 Caveats regarding MBH region selection and realism of spectral intensities

Our calculations broadly suggest that the most effective use of MBH intensity recovery is not in approximating the complete vibrational spectrum of a molecule but rather the isolation of particular peaks of interest, in the absence of any background of co-located vibrational intensity. Such calculations furnish useful information regarding the relative intensity of peaks of interest as well as their modulation with respect to geometric and electronic changes. Notably, MBH intensity recovery appears to be a good strategy for the simulation of the difference spectra arising from changes in moieties of interest in a molecule. We find that the presence of vibrational peaks that are co-located with peaks of diagnostic interest may exhibit surprising disparity with complete vibrational spectra in light of the neglect of degrees of freedom giving rise to those peaks.

We noted the presence of transient peaks in some calculations that we attribute to boundary condition effects, namely, the bisection of polarized bonds by the definition of a mobile block.

4.4 Opportunities for parallelism

We note that the single-point calculations on displaced geometries contributing to the calculation of dipole derivatives are totally independent with respect to each other, and therefore the evaluation of these derivatives is an 'embarrassingly parallel' problem. This fact offers further routes to reducing real-time computational expense, although we did not investigate this. This is because most computational chemistry software already leverages fine-grained parallelism to accelerate SCF evaluation. Hierarchical or so-called multi-grain parallelism is therefore liable to introduce unwarranted additional complexity in the provisioning of computational resources without affording a meaningful improvement in performance, excepting cases where heterogeneous hardware provides a logical mapping of the different hierarchical levels of a computational problem into hardware.‖ [28]

‖ An accessible example of effective hierarchical parallelism on heterogeneous hardware is the Folding@Home project (https://folding.stanford.edu/), where a task-parallel problem (conformation space exploration) is allocated to compute nodes in a high-latency fashion (to wit, personal computers of volunteers *via* the Internet), where molecular dynamics calculations (optionally performed with fine-grained parallelism contingent on available CPU or GPU cores) are performed.

5 Conclusion

Herein, we have demonstrated the feasibility of calculating vibrational intensities in a partial-Hessian approximation, permitting the simplification of IR intensity calculations to subsets of a molecule by re-use of pre-calculated properties as well as the calculation of the IR intensity contribution of vibrationally isolated moieties. We parameterize the computational efficiency of the methodology and find it to be highly predictable with respect to the size of the species under study and the size of the region of interest within that molecule. We suggest that the methodology is most useful for the simulation of difference spectroscopy within substructure of a molecule.

Acknowledgements

This research was undertaken with the assistance of resources from the National Computational Infrastructure (NCI), which is supported by the Australian Government.

References

1 M. Strickler, L. Walker, W. Hillier and R. Debus, *Biochemistry*, 2005, **44**, 8571–8577.
2 R. Debus, *Coord. Chem. Rev.*, 2008, **252**, 224–258.
3 W. Chuah, R. Stranger, R. Pace, E. Krausz and T. Frankcombe, *J. Phys. Chem.*, 2014, **118**, 3553–3558.
4 R. Terrett, T. Frankcombe, R. Pace and R. Stranger, *J. Inorg. Biochem.*, 2006, **155**, 101–104.
5 W. Chuah, R. Stranger, R. Pace, E. Krausz and T. Frankcombe, *J. Phys. Chem. B*, 2016, **120**, 377–385.
6 H. Li and J. Jensen, *Theor. Chem. Acc.*, 2001, **107**, 211–219.
7 H. Woodcock, W. Zheng, A. Ghysels, Y. Shao, J. Kong and B. Brooks, *J. Chem. Phys.*, 2008, **129**, 214109.
8 A. Ghysels, D. Van Neck, V. Van Speybroeck, T. Verstraelen and M. Waroquier, *J. Chem. Phys.*, 2007, **126**, 224102.
9 A. Ghysels, D. Van Neck and M. Waroquier, *J. Chem. Phys.*, 2007, **127**, 164108.
10 P. Durand, G. Trinquier and Y.-H. Sanejouand, *Biopolymers*, 1994, **34**, 759–771.
11 F. Tama, F. Gadea, O. Marques and Y.-H. Sanejouand, *Proteins: Struct., Funct., Genet.*, 2000, **41**, 1–7.
12 A. Ghysels, V. van Speybroeck, E. Pauwels, S. Catak, B. Brooks, D. van Neck and M. Waroquier, *J. Comput. Chem.*, 2009, **31**, 994–1007.
13 M. Reiher and J. Neugebauger, *J. Chem. Phys.*, 2003, **118**, 1634–1641.
14 S. Luber, J. Neugebauer and M. Reiher, *J. Chem. Phys.*, 2009, **130**, 064105.
15 S. Luber and M. Reiher, *ChemPhysChem*, 2009, **10**, 2049–2057.
16 K. Kiewisch, S. Luber, J. Neugebauer and M. Reiher, *Chimia*, 2009, **63**, 270–274.
17 G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders and T. Ziegler, *J. Comput. Chem.*, 2001, **22**, 931–967.
18 C. Fonseca Guerra, J. Snijders, G. te Velde and E. Baerends, *Theor. Chem. Acc.*, 1998, **99**, 391–403.
19 SCM, *ADF 2014*, Vrije Universiteit, Theoretical Chemistry, Amsterdam, The Netherlands, 2014.
20 A. Becke, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1988, **38**, 3098–3100.
21 J. Perdew, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1986, **33**, 8822–8824.
22 E. van Lenthe and E. Baerends, *J. Comput. Chem.*, 2003, **24**, 1142–1156.
23 J. Dybal and S. Krimm, *J. Mol. Struct.*, 1988, **189**, 383–392.
24 R Core Team, *R: a language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, 2016.
25 Blender Online Community, *Blender – a 3D modelling and rendering package*, Blender Institute, Amsterdam, The Netherlands, 2016.
26 *The PyMOL Molecular Graphics System, Version 1.8*, Schrödinger, LLC, 2015.
27 D. Martínez, M. Motevalli and M. Watkinson, *Dalton Trans.*, 2010, **39**, 446–455.
28 F. Blagojevic, X. Feng, K. W. Cameron and D. S. Nikolopoulos, Modeling Multi-grain Parallelism on Heterogeneous Multicore Processors: A Case Study of the Cell BE, in *Proceedings of the 3rd International Conference on High-Performance Embedded Architectures and Compilers, Lecture Notes in Computer Science*, Springer, 2008, vol. 4917, pp. 38–52.

Conclusion 8

> Science is what we understand well enough to explain to a computer. Art is everything else we do.
>
> *Donald Ervin Knuth*

8.1 Remarks on Introductory Chapters

Achieving the transition to a sustainable and carbon-neutral energy economy is perhaps one of the most exigent challenges facing humanity.

A potential route to this economy is artificial photosynthesis, wherein the process by which plants fix solar energy by splitting the water

molecule is replicated in artificial chemical systems. Artificial photosynthesis can yield both hydrogen from water, as well as more complex, value-added compounds through Fischer–Tropsch synthesis. Whilst the biological mechanism and infrastructure of photosynthesis is without a doubt exceptionally complex—representing over two billion years of evolutionary refinement—the operative principles of photosynthesis are not beyond comprehension and distillation into substantially simpler functional models. Understanding the photosynthetic mechanism in as much detail as possible is important for facilitating rational design of artificial photosystems.

This thesis has introduced the experimental and theoretical background of photosynthesis in Chapter 1. To a great extent, interest in photosynthesis has been informed by a desire to know how nature can elegantly achieve:

- Electrochemical scission of water under ambient conditions of temperature and pressure, a process that is quite energetically intensive in its extant industrial implementations.
- Controlled recycling of oxidation states in a conservative catalytic cycle.

Whilst water oxidation by reduced metals is a common phenomenon, controlled cyclic reduction of a metal by a light-driven process requires

somewhat more sophistication. The OEC consists of a highly sensitive framework supporting both the oxidation of water and the reduction of four chemically distinct manganese ions, and it has not been until recently that crystallography, magnetic resonance, and other forms of spectroscopy have converged on a coherent picture of the geometry of PSII. Nevertheless, a complete account of the dynamical behaviour and oxidation state disposition of the OEC remains both highly contentious and an invaluable source of insights into biomimetic photoassisted oxidation.

In Chapter 2, a case has been made for the necessity and urgency of development of artificial photosynthesis technology, drawing together several threads:

- Anthropogenic global warming, and its corollary of climate change. Combustion of fossil fuels has been linked to increasing concentrations of greenhouse gases—particularly carbon dioxide—in Earth's atmosphere, which has been causally implicated in excursions in mean global temperatures with respect to historical records since the beginning of the industrial revolution. Climate change is anticipated to be a source of stress on biodiversity, food security, water security, and health.

- Geopolitical instability arising from a lack of energy independence on the part of industrialised nations. The geographical localisation

of petrochemical deposits has been both a direct and indirect cause of war, and is likely to be of great concern in the future.

- Depletion of petrochemical reserves is likely to result in both short-term price shocks and secular increase in oil prices. The popularisation of new modalities for oil extraction, such as tar sand exploitation and hydraulic fracturing (*'fracking'*) raise serious environmental questions, whilst petrochemical exploration and exploitation in the Arctic and Antarctica is a potential threat to the environment in those regions in the long term.

Whilst any of these threats is cause for serious concern, the confluence and interplay of these factors represents a dire prospect for the progress and continuity of complex society as we currently understand it. Transition to sustainable sources of energy—such as are potentially afforded by artificial photosynthesis—is anticipted to ameliorate these stresses. This is by **a**) comprising part of an energy economy that has the potential to be carbon-neutral, thus not contributing CO_2—a greenhouse gas—to the atmosphere, and **b**) De-coupling energy security from access to highly localised petrochemical deposits. Whilst hydrogen economies are a plausible mode of energy storage and transport in a post-peak-oil world, or in the non-legacy world concept popularised by Nocera, *et al.*, present hydrogen generation capabilities are strongly tied to fossil fuel

reforming, which does not solve the problem of petrochemical reliance. A number of credible lines of inquiry towards water splitting as a source of hydrogen fuel are being actively developed in light of this situation, each with attendant technical challenges. Artificial photosynthesis revolving around the biomimesis of PSII is attractive because of the high quantum efficiency of PSII and the extremely mild conditions under which biological photosynthesis is achieved, motivating efforts to understand in detail the structure and mechanism of the PSII OEC.

In Chapter 3, an abridged introduction to computational chemistry was presented, with specific focus on locating Density Functional Theory—the premiere electronic structure theory for intermediate size systems (tens to hundreds of atoms)—within the context of allied methodologies. DFT relies upon a number of approximations (some unique to DFT, others which are shared by electronic structure methods generally) to obtain accurate results in a computationally efficient manner. Limitations on the size of systems that can be analysed by computational techniques speak to the differing algorithmic complexity of those techniques, as well as the speed, parallelism, and memory capacity of extant technology.

8.2 Remarks on Research Publications

The introductory chapters enumerated previously provide a context and vocabulary for understanding the specifics and thematic relations between the scholarly publications that are introduced in Chapters 4–7. These publications are united by a common goal of rationalising particular structural and mechanistic traits of the PSII OEC. The first three of these papers present different computational explorations of the PSII OEC, themselves based on different experimental probes of the OEC's structure and water oxidation metabolism. The fourth presents a novel methodology that was developed with specific intent to answer questions about the photosynthetic apparatus in a computationally expeditious way. This methodology also stands on its own as a general technique for acceleration and simplification of computational vibrational structure investigation.

The publication introduced in Chapter 4 in part sought to resolve the aetiology of structural differences in the OEC upon substitution of the Ca^{2+} ion of the Mn_4CaO_5 heterocubane with Sr^{2+}, by constructing a minimal crystallographically adapted computational model. Sr^{2+} substitution of the PSII OEC is of particular interest because strontium is the only known metal ion that substitutes for calcium in the OEC whilst retaining

substantive water-oxidising activity. Whilst other known metal substitutions of the PSII OEC do not express meaningful water oxidation activity, diffractometric characterisation of their respective OECs, potentially assisted by computational modeling, may be fruitful in parameterising metal ions and ligand geometries germane to artificial photosystem design.

The publication introduced in Chapter 5 elaborated upon elementary studies of small molecule models of part of the PSII OEC, which provided a proof-of-concept of a putative compensatory effect that may account for the observed FTIR difference silence on S-state advance of D1-Asp170, which is observed crystallographically to coordinate the OEC. This publication parameterised the integrated FTIR difference response of a model of D1-Asp170 with respect to conformational and constitutional isomerism. The massively combinatoric approach adopted—as well as the simulation of both frequency and intensity shifts—affords substantial additional insight over the preceding work and enhanced confidence in the feasibility of the nominal compensatory effect.

The publication introduced in Chapter 6 presented magnetic structure data (in the form of broken-symmetry DFT spin ladders) and hyperfine tensor data for a number of models of the PSII OEC, and illustrated the extend to which these data comport with competing models of magnetic coupling and ligand interactions within the functional OEC, imputed

from experimental data. These studies illustrate the degree to which computational studies can act as a critial adjunct to indirect probes of molecular structure, which typically underdetermine geometries. In a similar vein, correlation of X-ray absorption fine structure data (EXAFS) with structural models of the OEC is an important target of future research within the group.

In Chapter 7, a methodology was introduced for obtaining complete vibrational information (*i.e.* both vibrational frequencies and intensities) of substructures of a computational model in the Mobile Block Hessian approximation. This work was undertaken to lay the foundation for vibrational analysis of models of the PSII OEC in a computationally efficient manner. More specifically, the MBH intensities methodology is anticipated to make feasible inspections of FTIR difference intensity corresponding to ligands attached to oxidisable Mn atoms in large, highly realistic OEC models. This endeavour was previously precluded by limited computational resources, however preliminary work has been performed on applying the MBH intensities methodology to crystallographically adapted models of the OEC, with initial performance information suggesting that wide conformational exploration is feasible at relatively modest computational expense.

8.3 Coda

Computational chemistry, tempered by intuition, has proven itself to be in invaluable tool in understanding chemical systems. Nevertheless, the complex electronic structures of bioinorganic enzymatic cofactors have placed extraordinary demands on technology. The exponential depreciation of computational resources leaves much to be hopeful about, regarding the accelerating accessibility of increasingly complete and sophisticated models of nature's *'engines of creation'*.[a]

This is fortunate, because a convergence of circumstances creates a pressing need for new routes to energy security. Through an evolutionary process spanning geological time, plants and other photosynthetic organisms have exquisitely distilled the secrets of water oxidation in the form of the structural and mechanistic details of the PSII OEC. It stands to reason that if we can secure the collective human insight and digital processing power, we can learn them as well.

[a] After K.E. Drexler. [242]

Bibliography

[1] F.–J. Schmitt, E.G. Maksimov, P. Hätti, et al. "Coupling of different isolated photosynthetic light harvesting complexes and CdSe/ZnS nanocrystals via Förster resonance energy transfer". In: *Biochimica et Biophysica Acta - Bioenergetics* 1817.8 (2012), pp. 1461–1470. DOI: 10.1016/j.bbabio.2012.03.030.

[2] W. Martin, J. Baross, D. Kelley, et al. "Hydrothermal vents and the origin of life". In: *Nature Reviews Microbiology* 6 (2008), pp. 805–814. DOI: 10.1038/nrmicro1991.

[3] S. Duperron, S.M. Gaudron, C.F. Rodrigues, et al. "An overview of chemosynthetic symbioses in bivalves from the North Atlantic and Mediterranean Sea". In: *Biogeosciences* 10 (2013), pp. 3241–3267. DOI: 10.5194/bg-10-3241-2013.

[4] M.J. Mottl. "Serpentinization, Abiogenic Methane, and Extremophilic Archaea within the Seafloor". In: NAI Winter School, Jan. 2005 at

University of Hawai'i School of Ocean & Earth Science & Technology, 2005. URL: http://www.ifa.hawaii.edu/UHNAI/NAIweb/presentations/Mariana%20Forearc.pdf.

[5] F. Lartaud, C.T.S. Little, M. de Rafelis, et al. "Fossil evidence for serpentinization fluids fueling chemosynthetic assemblages". In: *Proceedings of the National Academy of Sciences* 108.19 (2011), pp. 7698–7703. DOI: 10.1073/pnas.1009383108.

[6] B. Ménez, V. Pasini, and D. Brunelli. "Life in the hydrated suboceanic mantle". In: *Nature Geoscience* 5 (2012), pp. 133–137. DOI: 10.1038/ngeo1359.

[7] G. Etiope and B. Sherwood Lollar. "Abiotic Methane on Earth". In: *Reviews of Geophysics* 51 (2013), pp. 276–299. DOI: 10.1002/rog.20011/abstract.

[8] O. Müntener. "Serpentine and serpentinization: A link between planet formation and life". In: *Geology* 32.10 (2010), pp. 959–960. DOI: 10.1130/focus102010.1.

[9] D.O.B. Jones, A. Walls, M. Clare, et al. "Asphalt mounds and associated biota on the Angolan margin". In: *Deep-Sea Research I* 194 (2014), pp. 124–136. DOI: 10.1016/j.dsr.2014.08.010.

[10] M. Brüning, H. Sahling, I.R. MacDonald, et al. "Origin, distribution, and alteration of asphalts at Chapopote Knoll, Southern Gulf

of Mexico". In: *Marine and Petroleum Geology* 27 (2009), pp. 1093–1106. DOI: 10.1016/j.marpetgeo.2009.09.005.

[11] I.R. MacDonald, G. Bohrmann, E. Escobar, et al. "Asphalt Volcanism and Chemosynthetic Life in the Campeche Knolls, Gulf of Mexico". In: *Science* 304 (2004), pp. 999–1002. DOI: 10.1126/science.1097154.

[12] H.D. Holland. "The oxygenation of the atmosphere and oceans". In: *Philosophical Transactions of the Royal Society B* 361 (2006), pp. 903–915. DOI: 10.1098/rstb.2006.1838.

[13] S.A. Crowe, L.N. Døssing, N.J. Beukes, et al. "Atmospheric oxygenation three billion years ago". In: *Nature* 501 (2013), pp. 535–539. DOI: 10.1038/nature12426.

[14] J.F. Kasting. "Theoretical constraints on oxygen and carbon dioxide concentrations in the Precambrian atmosphere". In: *Precambrian Research* 34.3 (1987), pp. 205–229. DOI: 10.1016/0301-9268(87)90001-5.

[15] Y. Umena, K. Kawakami, J.–R. Shen, et al. "Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å". In: *Nature* 473.7345 (2011), pp. 55–60. DOI: 10.1038/nature09913.

[16] R.E. Blankenship. "Antenna Complexes". In: *Molecular Mechanisms of Photosynthesis*. International: Blackwell Science, 2002, pp. 61–94.

[17] B. Loll, J. Kern, W. Saenger, et al. "Towards complete cofactor arrangement in the 3.0 Å resolution structure of photosystem II". In: *Nature* 438 (2005), pp. 1040–1044. DOI: 10.1038/nature04224.

[18] R.E. Blankenship. "Reaction Centers". In: *Molecular Mechanisms of Photosynthesis*. International: Blackwell Science, 2002, pp. 95–123.

[19] J. Whitmarsh and Govindjee. "The Photosynthetic Process". In: *Concepts in Photobiology: Photosynthesis and Photomorphogenesis*. Ed. by G.S. Singhal, G. Renger, S.K. Sopory, et al. New Delhi, India/Dordrecht, Netherlands: Narosa Publishers/Kluwer Academic, 1999, pp. 11–51.

[20] B. Kok, B. Forbush, and M. McGloin. "Cooperation of Charges in Photosynthetic O_2 Evolution – I. A Linear Four Step Mechanism". In: *Photochemistry and Photobiology* 11 (1970), pp. 457–475. DOI: 10.1111/j.1751-1097.1970.tb06017.x.

[21] G.M. Cheniae. "A recollection of the development of the Kok–Joliot model of photosynthetic oxygen evolution". In: *Photosynthesis Research* 38 (1993), pp. 225–227. DOI: 10.1007/BF00046748.

[22] R. Emerson and W. Arnold. "A Separation of the Reactions in Photosynthesis by Means of Intermittent Light". In: *The Journal of General Physiology* 15 (1932), pp. 391–420.

[23] R. Emerson and W. Arnold. "The Photochemical Reaction in Photosynthesis". In: *The Journal of General Physiology* 16 (1932), pp. 191–205.

[24] F.L. Allen and J. Franck. "Photosynthetic Evolution of Oxygen by Flashes of Light". In: *Archives of Biochemistry and Biophysics* 58 (1955), pp. 124–143. DOI: 10.1016/0003-9861(55)90100-4.

[25] V. Krewald, M. Retegan, N. Cox, et al. "Metal Oxidation States in Biological Water Splitting". In: *Chemical Science* 6 (2015), pp. 1676–1695. DOI: 10.1039/C4SC03720K.

[26] D.M. Proserpio, R. Hoffmann, and G.C. Dismukes. "Molecular Mechanism of Photosynthetic Oxygen Evolution: A Theoretical Approach". In: *Journal of the American Chemical Society* 114 (1992), pp. 4374–4382. DOI: 10.1021/ja00037a052.

[27] L. I. Krishtalik. "Activation energy of photosynthetic oxygen evolution: an attempt at theoretical analysis". In: *Bioelectrochemistry and Bioenergetics* 23 (1990), pp. 249–263. DOI: 10.1016/0302-4598(90)80014-A.

[28] V.P. Shinkarev. "Oxygen Evolution in Photosynthesis: Simple Analytical Solution for the Kok Model". In: *Biophysical Journal* 85.1 (2003), pp. 435–441. DOI: 10.1016/S0006-3495(03)74488-9.

[29] D.R.J. Kolling, N. Cox, G.M. Ananyev, et al. "What Are the Oxidation States of Manganese Required To Catalyze Photosynthetic Water Oxidation?" In: *Biophysical Journal* 103.2 (2012), pp. 313–322. DOI: 10.1016/j.bpj.2012.05.031.

[30] G. Han, F. Mamedov, and S. Styring. "Misses during Water Oxidation in Photosystem II Are S State-dependent". In: *The Journal of Biological Chemistry* 287.16 (2012), pp. 13422–13429. DOI: 10.1074/jbc.M112.342543.

[31] R.M. Cinco, K.L. McFarlane Holman, J.H. Robblee, et al. "Calcium EXAFS establishes the Mn–Ca cluster in the Oxygen Evolving Complex of Photosystem II". In: *Biochemistry* 41.43 (2002), pp. 12928–12933. DOI: 10.1021/bi026569p.

[32] M. Newville. "Fundamentals of XAFS, Rev. 1.7". In: (2004), pp. 1–43.

[33] J. Yano, Y. Pushkar, P. Glatzel, et al. "High-Resolution Mn EXAFS of the Oxygen-Evolving Complex in Photosystem II: Structural Implications for the Mn_4Ca Cluster". In: *Journal of the American*

Chemical Society 127.43 (2005), pp. 14974–14975. DOI: 10.1021/ja054873a.

[34] M. Suga, F. Akita, K. Hirata, et al. "Native structure of Photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses". In: *Nature* 517 (2015), pp. 99–103. DOI: 10.1038/nature13991.

[35] Z. Jurek, G. Faigel, and M. Tegze. "Dynamics in a cluster under the influence of intense femtosecond hard X-ray pulses". In: *The European Physical Journal D - Atomic, Molecular, Optical and Plasma Physics* 29.2 (2004), pp. 217–229. DOI: 10.1140/epjd/e2004-00033-3.

[36] G.C. Dismukes and Y. Siderer. "Intermediates of a polynuclear manganese center involved in photosynthetic oxidation of water". In: *Proceedings of the National Academy of Sciences* 78.1 (1981), pp. 274–278.

[37] O. Hanson, R. Aasa, and T. Vänngård. "The Origin of the Multiline and $g = 4.1$ Electron Paramagnetic Resonance Signals from the Oxygen-Evolving System of Photosystem II". In: *Biophysical Journal* 51.5 (1987), pp. 825–832.

[38] R.J. Debus. "The manganese and calcium ions of photosynthetic oxygen evolution". In: *Biochimica et Biophysica Acta* 1102.3 (1992), pp. 269–352. DOI: 10.1016/0005-2728(92)90133-M.

[39] P.W. Anderson. "Antiferromagnetism. Theory of Superexchange Interaction". In: *Physical Review* 79.2 (1950), pp. 350–357. DOI: 10.1103/PhysRev.79.350.

[40] L. Jansen and R. Block. "Superexchange and Superconductivity". In: *Physica A* 161.3 (1989), pp. 385–431. DOI: 10.1016/0378-4371(89)90434-2.

[41] D.G. Nocera. "The Artificial Leaf". In: *Accounts of Chemical Research* 45.5 (2011), pp. 767–776. DOI: 10.1021/ar2003013.

[42] J. Pieper, T. Hauß, A. Buchsteiner, et al. "Temperature- and Hydration-Dependent Protein Dynamics in Photosystem II of Green Plants Studied by Quasielastic Neutron Scattering". In: *Biochemistry* 46 (2007), pp. 11398–11409. DOI: 10.1021/bi700179s.

[43] A. Melis. "Photosystem-II damage and repair cycle in chloroplasts: what modulates the rate of photodamage *in vivo*?" In: *Trends in Plant Science* 4.4 (1999), pp. 130–135. DOI: 10.1016/S1360-1385(99)01387-4.

[44] W.F. Ruettinger, D.M. Ho, and G.C. Dismukes. "Protonation and Dehydration Reactions of the $Mn_4O_4L_6$ Cubane and Synthesis and Crystal Structure of the Oxidized Cubane $[Mn_4O_4L_6]^+$: A Model for the Photosynthetic Water Oxidizing Complex". In: *Inorganic Chemistry* 38.6 (1999), pp. 1036–1037. DOI: 10.1021/ic981145y.

[45] R. Chakrabarty, S.J. Bora, and B.K. Das. "Synthesis, Structure, Spectra and Electrochemical Properties, and Catalytic Use of Cobalt(III)-Oxo Cubane Clusters". In: *Inorganic Chemistry* 46 (2007), pp. 9450–9462. DOI: 10.1021/ic7011759.

[46] R. Brimblecombe, G.F. Swiegers, G.C. Dismukes, et al. "Sustained Water Oxidation Photocatalysis by a Bioinspired Manganese Cluster". In: *Angewandte Chemie International Edition* 47 (2008), pp. 7335–7338. DOI: 10.1002/anie.200801132.

[47] D. Lutterman, Y. Surendranath, and D.G. Nocera. "A Self-Healing Oxygen Evolving Catalyst". In: *Journal of the American Chemical Society Communications* 131 (2009), pp. 3838–3839. DOI: 10.1021/ja900023k.

[48] R.K. Hocking, R. Brimblecombe, L.-Y. Chang, et al. "Water-oxidation catalysis by manganese in a geochemical-like cycle". In: *Nature Chemistry* 3 (2011), pp. 461–466. DOI: 10.1038/nchem.1049.

[49] R.K. Hocking, R. Malaeb, W.P. Gates, et al. "Formation of a Nanoparticulate Brinessite-Like Phase in Purported Molecular Water Oxidation Catalyst Systems". In: *ChemCatChem* (2014), pp. 2028–2038. DOI: 10.1002/cctc.201400066.

[50] F.F. Marafatto, M.L. Strader, J. Gonzalez-Holguera, et al. "Rate and mechanism of the photoreduction of birnessite (MnO_2) nanosheets".

In: *Proceedings of the National Academy of Sciences* 112.15 (2015), pp. 4600–4605. DOI: 10.1073/pnas.1421018112.

[51] R. Chakrabarty and B.K. Das. "Epoxidation of α-pinene catalysed by tetrameric cobalt(III) complexes". In: *Journal of Molecular Catalysis A: Chemical* 223 (2004), pp. 39–44. DOI: 10.1016/j.molcata.2004.02.028.

[52] P. Sarmah, R. Chakrabarty, P. Phukan, et al. "Selective oxidation of alcohols catalysed by a cubane–like Co(III) oxo cluster immobilised on porous organomodified silica". In: *Journal of Molecular Catalysis A: Chemical* 268 (2007), pp. 36–44. DOI: 10.1016/j.molcata.2006.11.042.

[53] R. Chakrabarty, B.K. Das, and J.H. Clark. "Enhanced selectivity in green catalytic epoxidation using a supported cobalt complex". In: *Green Chemistry* 9 (2007), pp. 845–848. DOI: 10.1039/b618862a.

[54] R. Chakrabarty, D. Kalita, and B.K. Das. "Catalytic oxidation of *p*-xylene in water by cobaltIII cubane cluster". In: *Polyhedron* 26 (6 2007), pp. 1239–1244. DOI: 10.1016/j.poly.2006.10.014.

[55] R. Chakrabarty, P. Sarmah, B. Saha, et al. "Catalytic Properties of Cobalt(III)-Oxo Cubanes in TBHP Oxidation of Benzylic Alcohols". In: *Inorganic Chemistry* 48 (2009), pp. 6371–6379. DOI: 10.1021/ic802115n.

[56] N.S. McCool, D.M. Robinson, J.E. Sheats, et al. "A Co_4O_4 "Cubane" Water Oxidation Catalyst Inspired by Nature". In: *The Journal of the American Chemical Society* 133.30 (2011), pp. 11446–11449. DOI: 10.1021/ja203877y.

[57] M. Natali, M. Orlandi, S. Berardi, et al. "Photoinduced Water Oxidation by a Tetraruthenium Polyoxometalate Catalyst: Ion-pairing and Primary Processes with $Ru(bpy)_3^{2+}$ Photosensitiser". In: *Inorganic Chemistry* 51 (2012), pp. 7324–7331. DOI: 10.1021/ic300703f.

[58] S. Berardi, G. La Ganga, M. Natali, et al. "Photocatalytic Water Oxidation: Tuning Light-Induced Electron Transfer by Molecular Co_4O_4 Cores". In: *Journal of the American Chemical Society* 134.27 (2012), pp. 11104–11107. DOI: 10.1021/ja303951z.

[59] J.G. McAlpin, T.A. Stich, C.A. Ohlin, et al. "Electronic Structure Description of a $[Co(III)_3Co(IV)O_4]$ Cluster: A Model for the Paramagnetic Intermediate in Cobalt-Catalyzed Water Oxidation". In: *Journal of the American Chemical Society* 133.39 (2011), pp. 15444–15452. DOI: 10.1021/ja202320q.

[60] R. Brimblecombe, A. Koo, G.C. Dismukes, et al. "Solar Driven Water Oxidation by a Bioinspired Manganese Molecular Catalyst".

In: *Journal of the American Chemical Society* 132 (2010), pp. 2892–2894. DOI: 10.1021/ja910055a.

[61] M. Grätzel. "Recent Advances in Sensitized Mesoscopic Solar Cells". In: *Accounts of Chemical Research* 42.11 (2009), pp. 1788–1798. DOI: 10.1021/ar900141y.

[62] G. Boschloo and A. Hagfeldt. "Characteristics of the Iodide/Triiodide Redox Mediator in Dye-Sensitized Solar Cells". In: *Accounts of Chemical Research* 42.11 (2009), pp. 1819–1826. DOI: 10.1021/ar900138m.

[63] P.J. Cameron, L.M. Peter, S.M. Zakeeruddin, et al. "Electrochemical studies of the Co(III)/Co(II)(dbbip)$_2$ redox couple as a mediator for dye-sensitized nanocrystalline solar cells". In: *Coordination Chemistry Reviews* 248 (2004), pp. 1447–1453. DOI: 10.1016/j.ccr.2004.02.010.

[64] Z. Zhang, P. Chen, T.N. Murakami, et al. "The 2,2,6,6-Tetramethyl-1-piperidinyloxy Radical: An Efficient, Iodine-Free Redox Mediator for Dye-Sensitized Solar Cells". In: *Advanced Functional Materials* 18 (2008), pp. 3410–346. DOI: 10.1002/adfm.200701041.

[65] S.M. Zakeeruddin and M. Grätzel. "Solvent–Free Ionic Liquid Electrolytes for Mesoscopic Dye–Sensitized Solar Cells". In: *Advanced Functional Materials* 19 (2009), pp. 2187–2202. DOI: 10.1002/adfm.200900390.

[66] S. Rani and R.M. Mehra. "ZnO solid-state dye sensitized solar cells using composite electrolyte of poly(3-hexylthiophene-2,5-diyl) and carbon nanotubes". In: *International Journal of Renewable Energy* 1 (2009), pp. 033109-1–033109-12. DOI: 10.1063/1.3156004.

[67] J. Li, T. Osasa, Y. Hirayama, et al. "Solid-State Dye-Sensitized Solar Cells Using Poly[2-methoxy-5-(2-ethylhexyloxy)-1,4-phenylene-vinylene] as a Hole-Transporting Material". In: *Japanese Journal of Applied Physics* 45 (2006), pp. 8728–8732. DOI: 10.1143/JJAP.45.8728.

[68] P. Meredith, B.J. Powell, J. Riesz, et al. "Broadband Photon-harvesting Biomolecules for Photovoltaics". In: *Artificial Photosynthesis*. Ed. by A.F. Collings and C. Critchley. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2005. DOI: 10.1002/3527606742.ch3.

[69] J.P. Dekker and R. Van Grondelle. "Primary charge separation in Photosystem II". In: *Photosynthesis Research* 63 (2000), pp. 195–208. DOI: 10.1023/A:1006468024245.

[70] M. Grätzel. "Photoelectrochemical cells". In: *Nature* 414 (2001), pp. 338–344. DOI: 10.1038/35104607.

[71] M. Pastore, E. Mosconi, F. De Angelis, et al. "A Computational Investigation of Organic Dyes for Dye-Sensitized Solar Cells: Benchmark, Strategies, and Open Issues". In: *Journal of Physical Chemistry C* 114 (2010), pp. 7205–7212. DOI: 10.1021/jp100713r.

[72] A.J. Heeger. *Printable Organic Solar Cells—Challenges and Opportunities in Technology Transfer from Lab to Market*. 2010. URL: https://www.natureasia.com/pdf/ja-jp/events/photonics/a-1.pdf.

[73] W.J. Youngblood, S.–H.A. Lee, K. Maeda, et al. "Visible Light Water Splitting Using Dye–Sensitized Oxide Semiconductors". In: *Accounts of Chemical Research* 42.12 (2009), pp. 1966–1973. DOI: 10.1021/ar9002398.

[74] M. Grätzel. "The artificial leaf, bio-mimetic photocatalysis". In: *caTTech* 3 (1999), pp. 3–17.

[75] Intergovernmental Panel on Climate Change (IPCC) Working Groups. "Synthesis Report—Summary for Policymakers". In: *IPCC Fifth Assessment Report (AR5)*. Ed. by R.K. Pachauri and L.A. Meyer. Geneva, Switzerland: IPCC, 2014, pp. 1–32.

[76] M. Rhein, S.R. Rintoul, S. Aoki, et al. "Observations: Ocean". In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by T.F. Stocker, D. Qin, G.-K. Plattner, et al. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 2013. Chap. 3, pp. 255–315.

[77] D.G. Vaughan, J.C. Comiso, I. Allison, et al. "Observations: Cryosphere". In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by T.F. Stocker, D. Qin, G.-K. Plattner, et al. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 2013. Chap. 4, pp. 317–382.

[78] A. Donohoe and D.S. Battisti. "Atmospheric and Surface Contributions to Planetary Albedo". In: *Journal of Climate* 24 (2011), pp. 4402–4418. DOI: 10.1175/2011JCLI3946.1.

[79] J.K. Bishop and W.B. Rossow. "Spatial and Temporal Variability of Global Surface Solar Irradiance". In: *Journal of Geophysical Research* 96.9 (1991), pp. 16839–16858. DOI: 10.1029/91JC01754.

[80] O. Boucher, D. Randall, P. Artaxo, et al. "Clouds and Aerosols". In: *Climate Change 2013: The Physical Science Basis. Contribution*

of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Ed. by T.F. Stocker, D. Qin, G.-K. Plattner, et al. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 2013. Chap. 7, pp. 571–657.

[81] N.R. Bates, M.H.P. Best, K. Neely, et al. "Detecting anthropogenic carbon dioxide uptake and ocean acidification in the North Atlantic Ocean". In: *Biogeosciences* 9 (2012), pp. 2509–2522. DOI: 10.5194/bg-9-2509-2012.

[82] S. Schmidtko, L. Stramma, and M. Visbeck. "Decline in global oceanic oxygen content during the past five decades". In: *Nature* 542 (2017), pp. 335–339. DOI: 10.1038/nature21399.

[83] G.W. vanLoon and S.J. Duffy. *Environmental Chemistry, 2e*. Oxford, 2006, pp. 164–181.

[84] F. Rubel and M. Kottek. "Observed and projected climate shifts 1901–2100 depicted by world maps of the Köppen–Geiger climate classification". In: *Meteorologische Zeitschrift* 19.2 (2010), pp. 135–141. DOI: 10.1127/0941-2948/2010/0430.

[85] T.P. Hughes, J.T. Kerry, M. Álvarez-Noriega, et al. "Global warming and recurrent mass bleaching of corals". In: *Nature* 543.7645 (2017), pp. 373–377. DOI: 10.1038/nature21707.

[86] R.Z. Poore, R.S. Williams Jr., and C. Tracey. *Sea Level and Climate*. 2011. URL: https://pubs.usgs.gov/fs/fs2-00/pdf/fs002-00_williams_508.pdf.

[87] G.A. Meehl and C. Tebaldi. "More Intense, More Frequent, and Longer Lasting Heat Waves in the 21st Century". In: *Science* 305.5686 (2004), pp. 994–997. DOI: 10.1126/science.1098704.

[88] A.L. Westerling, H.G. Hidalgo, D.R. Cayan, et al. "Warming and Earlier Spring Increase Western U.S. Forest Wildfire Activity". In: *Science* 313.5789 (2006), pp. 940–943. DOI: 10.1126/science.1128834.

[89] N.-Y. Kang and J.B. Elsner. "Trade-off between intensity and frequency of global tropical cyclones". In: *Nature Climate Change* 5 (2015), pp. 661–665. DOI: 10.1038/NCLIMATE2646.

[90] Commonwealth Science and Industry Research Organisation and Australian Government Bureau of Meteorology. *State of the Climate 2016*. Tech. rep. Commonwealth of Australia, 2016, pp. 1–22. URL: http://www.bom.gov.au/state-of-the-climate/State-of-the-Climate-2016.pdf.

[91] Australian Government Bureau of Meteorology. *Climate Change and variability: Tracker: Australian timeseries graphs*. 2017. URL:

http://www.bom.gov.au/climate/change/index.shtml#tabs=Tracker&tracker=timeseries.

[92] GISTEMP Team. *GISS Surface Temperature Analysis (GISTEMP)*. 2015. URL: http://data.giss.nasa.gov/gistemp/.

[93] J. Hansen, R. Ruedy, M. Sato, et al. "Global Surface Temperature Change". In: *Reviews of Geophysics* 48 (2010), pp. 1–29. DOI: 10.1029/2010RG000345.

[94] G.W. vanLoon and S.J. Duffy. *Environmental Chemistry, 2e*. Oxford, 2006, pp. 171–178.

[95] C. Sagan. *Cosmos*. MacDonald, 1981, pp. 96–102.

[96] F. Xi, S.J. Davis, P. Ciais, et al. "Substantial global carbon uptake by cement carbonation". In: *Nature Geoscience* 9 (2016), pp. 880–883. DOI: 10.1038/ngeo2840.

[97] Intergovernmental Panel on Climate Change (IPCC) Working Group III. "Industry". In: *IPCC Fifth Assessment Report (AR5)*. Ed. by R.K. Pachauri and L.A. Meyer. Geneva, Switzerland: IPCC, 2014, pp. 1–72.

[98] IPCC. "Summary for Policymakers". In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by T.F. Stocker, D. Qin, G.-K. Plattner, et al. Cambridge, United

Kingdom and New York, NY, USA: Cambridge University Press, 2013, pp. 3–29.

[99] C. MacFarling Meure, D. Etheridge, C. Trudinger, et al. "Law Dome CO_2, CH_4 and N_2O ice core records extended to 2000 years BP". In: *Geophysical Research Letters* 33.14 (2006), p. L14810. DOI: 10.1029/2006GL026152.

[100] C.D. Keeling, S.C. Piper, R.B. Bacastow, et al. "Atmospheric CO_2 and $^{13}CO_2$ exchange with the terrestrial biosphere and oceans from 1978 to 2000: observations and carbon cycle implications". In: *A History of Atmospheric CO_2 and its effects on Plants, Animals, and Ecosystems*. Ed. by J.R. Ehleringer, T.E. Cerling, and M.D. Dearing. New York: Springer Verlag, 2005, pp. 83–113. URL: http://scrippsco2.ucsd.edu/data/atmospheric_co2.

[101] Office of the Under Secretary of Defense for Policy. *National Security Implications of Climate-Related Risks and A Changing Climate*. Tech. rep. United States Department of Defense, 2015, pp. 1–14. URL: http://archive.defense.gov/pubs/150724-congressional-report-on-national-implications-of-climate-change.pdf.

[102] Office of the Director of National Intelligence. *Global Trends 2030: Alternative Worlds*. Tech. rep. United States National Intelligence

Council, 2012, pp. 32–37. URL: https://www.dni.gov/files/documents/GlobalTrends_2030.pdf.

[103] "Quantitative risk assessment of the effects of climate change on selected causes of death, 2030s and 2050s". In: (2014). Ed. by S. Hales, S. Kovats, S. Lloyd, et al., pp. 1–16. URL: http://apps.who.int/iris/bitstream/10665/134014/1/9789241507691_eng.pdf.

[104] M.P. McCarthy, M.J. Best, and R.A. Betts. "Climate change in cities due to global warming and urban effects". In: *Geophysical Research Letters* 37.9 (2010), pp. 1–5. DOI: 10.1029/2010GL042845.

[105] J.A. Merrigan. "Expected Energy Usage in the United States in the Year 2000". In: Sunlight to Electricity. Cambridge: The MIT Press, 1975. Chap. 1.

[106] Bureau of Public Affairs Office of the Historian. "Oil Embargo, 1973–1974". In: *Milestones: 1969–1976*. United States Department of State, 2017. URL: https://history.state.gov/milestones/1969-1976/oil-embargo.

[107] J.A. Paul. *Great Power Conflict over Iraqi Oil: The World War I Era*. 2002. URL: https://www.globalpolicy.org/component/content/article/185-general/40479-great-power-conflict-over-iraqi-oil-the-world-war-i-era.html.

[108] U.S. Congress, Committee on International Relations, Special Subcommittee on Investigations. *Oil Fields as Military Objectives: A Feasibility Study*. Washington D.C.: Congressional Research Service, 94th Congress, 1st Session, 1975, pp. 1–39. URL: https://www.mtholyoke.edu/acad/intrel/Petroleum/fields.htm.

[109] N. Chomsky. "The Responsibility of Intellectuals". In: *The New York Review of Books* (1967). URL: http://chomsky.info/19670223/.

[110] N. Chomsky. "Foreign Policy and the Intelligentsia". In: *"Human Rights" and American Foreign Policy*. Spokesman, 1978.

[111] United States Bureau of Labor Statistics. *CPI Inflation Calculator*. URL: http://data.bls.gov/cgi-bin/cpicalc.pl?cost1=2.4&year1=1990&year2=2016.

[112] T.C. Hayes. "CONFRONTATION IN THE GULF; The Oilfield Lying Below the Iraq–Kuwait Dispute". In: *The New York Times* (1990). URL: http://www.nytimes.com/1990/09/03/world/confrontation-in-the-gulf-the-oilfield-lying-below-the-iraq-kuwait-dispute.html.

[113] F.D. Roylance. "Burning oil wells could be disaster, Sagan says". In: *The Baltimore Sun* (1991). URL: http://articles.baltimoresun.

com/1991-01-23/news/1991023131_1_kuwait-saddam-hussein-sagan.

[114] R.C. Beckman and C.H. Schofield. "Defining EEZ Claims from Islands: A Potential South China Sea Change". In: *The International Journal of Marine and Coastal Law* 29 (2014), pp. 193–243. DOI: 10.1163/15718085-12341321.

[115] C. Ward. "The Maritime Boundaries of East Timor: the Role of International Law". In: Northern Territory Bar Association Conference, 2014. URL: http://ntba.asn.au/wp-content/uploads/The-Maritime-Boundaries-of-East-Timor.pdf.

[116] C. Duffy. "East Timor accuses Australia of spying for economic gain". In: *Australian Broadcasting Corporation* (2013). URL: http://www.abc.net.au/7.30/content/2013/s3900429.htm.

[117] United States Energy Information Administration. *Monthly Energy Review*. 2017. URL: https://www.eia.gov/totalenergy/data/monthly/.

[118] S.O. Dean. "Overview of Magnetic Fusion". In: *Prospects for Fusion Power*. Ed. by S.O. Dean. Pergamon, 1981, pp. 1–7.

[119] J.I. Lunine. "Titan as an analog of Earth's past and future". In: *The European Physical Journal Conferences* 1 (2009), pp. 267–274. DOI: 10.1140/epjconf/e2009-00926-7.

[120] J.R. Schramski, D.K. Gattie, and J.H. Brown. "Human domination of the biosphere: Rapid discharge of the earth-space battery foretells the future of humankind". In: *Proceedings of the National Academy of Sciences* 112.31 (2015), pp. 9511–9517. DOI: 10.1073/pnas.1508353112.

[121] R.G. Miller and S.R. Sorrel. "The future of oil supply". In: *Philosophical Transactions of the Royal Society A* 372.20130179 (2014), pp. 1–27. DOI: 0.1098/rsta.2013.0179.

[122] M.K. Hubbert. "Nuclear Energy and the Fossil Fuels". In: *Spring Meeting of the Southern District Division of Production, American Petroleum Institute*. Vol. 95. 1956, pp. 1–57.

[123] R.W. Bentley, R.H. Booth, J.D. Burton, et al. "Perspectives on the Future of Oil". In: *Energy Exploitation and Exploration* 18.2 (2000), pp. 1–60. DOI: 10.1260/0144598001492076.

[124] Office for Research and Development. *Hydraulic Fracturing for Oil and Gas: Impacts from the Hydraulic Fracturing Water Cycle on Drinking Water Resources in the United States (Executive Summary)*. Tech. rep. United States Environmental Protection Agency, 2016, pp. 1–50. URL: https://www.epa.gov/sites/production/files/2016-12/documents/hfdwa_executive_summary.pdf.

[125] C.A. Green, P.P. Styles, and B.J. Baptie. *Review and Recommendations for Induced Seismic Mitigation*. Tech. rep. Preese Hall Shale Gas Fracturing, 2012, pp. 1–26. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/15745/5075-preese-hall-shale-gas-fracturing-review.pdf.

[126] A.D. Charpentier, J.A. Bergerson, and H.L. MacLean. "Understanding the Canadian oil sands industry's greenhouse gas emissions". In: *Environmental Research Letters* 4 (2009), pp. 1–11. DOI: 10.1088/1748-9326/4/1/014005.

[127] K.P. Timoney and P. Lee. "Does the Alberta Tar Sands Industry Pollute? The Scientific Evidence". In: *The Open Conservation Biology Journal* 3 (2009), pp. 65–81.

[128] S. Pezard, A. Tingstad, K. Van Abel, et al. *Maintaining Arctic Cooperation with Russia*. Tech. rep. RAND Corporation, 2017, pp. 34–37. URL: 10.7249/RR1731.

[129] Australian Antarctic Division. *Mining in Antarctica*. 2012. URL: http://www.antarctica.gov.au/about-antarctica/environment/geology/mining.

[130] Secretariat of The Antarctic Treaty. *Protocol on Environmental Protection to The Antarctic Treaty*. Tech. rep. 1998. URL: http://www.ats.aq/documents/recatt/Att006_e.pdf.

[131] J.J. Romm. *The Hype About Hydrogen*. Island Press, 2004, p. 3.

[132] E. Kuperminc. *Aluminum Smelting*. 2003. URL: https://www.alcoa.com/global/en/about_alcoa/pdf/Smeltingpaper.pdf.

[133] *Hydrogen Fact Sheet: Hydrogen Production — Steam Methane Reforming (SMR)*. 2006.

[134] C. Wang and R.G. Prinn. "Impact of Emissions, Chemistry, and Climate on Atmospheric Carbon Monoxide: 100-year Predictions from a Global Chemistry–Climate Model". In: *Chemosphere - Global Change Science* (1998), pp. 73–81. DOI: 10.1016/S1465-9972(99)00016-1.

[135] C. E. G. Padró and V. Putsche. *Survey of the Economics of Hydrogen Technologies*. Tech. rep. National Renewable Energy Laboratory, 1999.

[136] Energy Efficiency and Renewable Energy Group. *Hydrogen Production – Natural Gas Reforming*. Tech. rep. U.S. Department of Energy, 2008.

[137] K.B. Anderson, K. Carrado-Gregar, C.L. Marshall, et al. *United States Patent Nº. 6,668,763 B2: Process for in-situ Production of*

Hydrogen (H_2) by Alcohol Decomposition for Emission Reduction for Internal Combustion Engines. 2003.

[138] A.V. Pattekar, M.V. Kothare, S.V. Karnik, et al. "A microreactor for *in-situ* hydrogen production by catalytic methanol reforming". In: *Proceedings of the 5th International Conference on Microreaction Technology (IMRET 5)*. 2001. DOI: 10.1007/978-3-642-56763-6_34.

[139] M.W. Kanan, Y. Surendranath, and D.G. Nocera. "Cobalt–phosphate oxygen-evolving compound". In: *Chemical Society Reviews* 38 (2009), pp. 109–114. DOI: 10.1039/B802885K.

[140] J.B. Gadhe and R.B. Gupta. "Hydrogen production by methanol reforming in supercritical water: Catalysis by *in-situ*–generated copper nanoparticles". In: *International Journal of Hydrogen Energy* 32 (2007), pp. 2374–2381. DOI: 10.1016/j.ijhydene.2006.10.050.

[141] W.L. Mao and H.-K. Mao. "Hydrogen storage in molecular compounds". In: *Proceedings of the National Academy of Sciences* 101.3 (2004), pp. 708–710. DOI: 10.1073/pnas.0307449100.

[142] D. Henwood and J. David Carey. "*Ab initio* investigation of molecular hydrogen physisorption on graphene and carbon nanotubes".

In: *Physical Review B* 75 (2007), 245413(1–10). DOI: 10.1103/PhysRevB.75.245413.

[143] W.J. Fan, R.Q. Zhang, B.K. Teo, et al. "Prediction of energetically optimal single-walled carbon nanotubes for hydrogen physisorption". In: *Applied Physics Letters* 95 (2009), 013116(1–3). DOI: 10.1063/1.3158597.

[144] B.D. Adams and A. Chen. "The role of palladium in a hydrogen economy". In: *Materials Today* 14.6 (2011), pp. 282–289. DOI: 10.1016/S1369-7021(11)70143-2.

[145] M.J. Sonter. "The Technical and Economic Feasibility of Mining The Near-Earth Asteroids". In: *Acta Astronautica* 41.4–10 (1998), pp. 637–647. DOI: 10.1016/S0094-5765(98)00087-3.

[146] T.J. Frankcombe. "Calcium Borohydride for Hydrogen Storage: A Computational Study fo $Ca(BH_4)_2$ Crystal Structures and the CaB_2H_x Intermediate". In: *Journal of Physical Chemistry* 114 (2010), pp. 9503–9509. DOI: 10.1021/jp1014109.

[147] D.A.J. Rand and R.M. Dell. *Hydrogen Energy - Challenges and Prospects*. The Royal Society of Chemistry, 2008, p. 109.

[148] B. Yildiz and M.S. Kazimi. "Efficiency of hydrogen production systems using alternative nuclear energy technologies". In: *International Journal of Hydrogen Production* 31 (2006), pp. 77–92. DOI: 10.1016/j.ijhydene.2005.02.009.

[149] J. Sigurvinsson, C. Mansilla, P. Lovera, et al. "Can high temperature steam electrolysis function with geothermal heat?" In: *International Journal of Hydrogen Energy* 32 (2007), pp. 1174–1182. DOI: 10.1016/j.ijhydene.2006.11.026.

[150] R.J. Pace. "An Integrated Artificial Photosynthesis Model". In: *Artificial Photosynthesis*. Ed. by A.F. Collings and C. Critchley. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2005.

[151] T.V. Bommaraju, P.J. Orosz, and E.A. Sokol. *Brine Electrolysis*. Tech. rep. Process Technology Optimization, Inc., 2007.

[152] P.A. D'Itri and F.M. D'Itri. "FORUM: Mercury Contamination: A Human Tragedy". In: *Environmental Management* 2.1 (1978), pp. 3–16. DOI: 10.1007/BF01866442.

[153] A. Steinfeld. *Concentrated Solar Energy: Power - Fuels - Storage (Bat-Sheva Seminar)*. 2007. URL: http://www.weizmann.ac.il/conferences/ASEO/Steinfeld.pdf.

[154] G. Vicens. *Renewable hydrogen production. The role of Solar Thermal Water Splitting. (M.Sc. Thesis)*. Lund University, 2012, pp. 27–28.

[155] L.E. Brecher, S. Spewock, and C.J. Warde. "The Westinghouse Sulfur Cycle for the Thermochemical Decomposition of Water". In: *International Journal of Hydrogen Energy* 2.1 (1977), pp. 7–15. DOI: 10.1016/0360-3199(77)90061-1.

[156] G.H. Farbman. "Hydrogen Production by The Westinghouse Sulfur Cycle Process: Program Status". In: *International Journal of Hydrogen Energy* 4.2 (1979), pp. 111–122. DOI: 10.1016/0360-3199(79)90045-4.

[157] F. Jomard, J.P. Feraud, and J.P. Caire. "Numerical modeling for preliminary design of the hydrogen production electrolyzer in the Westinghouse hybrid cycle". In: *International Journal of Hydrogen Energy* 33.4 (2008), pp. 1142–1152. DOI: 10.1016/j.ijhydene.2007.12.052.

[158] D.A.J. Rand and R.M. Dell. *Hydrogen Energy - Challenges and Prospects*. The Royal Society of Chemistry, 2008, pp. 138–139.

[159] J. Petrovic and G. Thomas. "Reaction of Aluminum with Water to Produce Hydrogen". In: (2008), pp. 1–26. URL: https://www1.eere.energy.gov/hydrogenandfuelcells/pdfs/aluminium_water_hydrogen.pdf.

[160] Y. Amano. *The Fukushima Daiichi Accident*. Vienna, Austria: International Atomic Energy Agency, 2015, p. 57. URL: http:

//www-pub.iaea.org/MTCD/Publications/PDF/Pub1710-ReportByTheDG-Web.pdf.

[161] M. Leyse and C. Paine. *Preventing Hydrogen Explosions In Severe Nuclear Accidents: Unresolved Safety Issues Involving Hydrogen Generation And Mitigation*. Tech. rep. Natural Resources Defense Council, 2014, pp. 1–49. URL: https://www.nrdc.org/file/4317.

[162] J. Simons. "Theoretical Chemistry, 1e". In: (2003), p. 116.

[163] P.W. Atkins. *Quanta: A Handbook of Concepts, 1e*. Oxford University Press, 1974, p. 142.

[164] H. Nakashima and H. Nakatsuji. "Solving the Schrödinger equation for helium atom and its isoelectronic ions with the free iterative complement interaction (ICI) method". In: *Journal of Chemical Physics* 127 (2007), pp. 224104-1–224104-14. DOI: 10.1063/1.2801981.

[165] J. Simons. "Theoretical Chemistry, 1e". In: (2003), pp. 276–278.

[166] R. Car and M. Parrinello. "Unified Approach for Molecular Dynamics and Density-Functional Theory". In: *Physical Review Letters* 55.22 (1985), pp. 2471–2474. DOI: 10.1103/PhysRevLett.55.2471.

[167] M.M. Waldrop. "More than Moore". In: *Nature* 530 (2016), pp. 144–148. DOI: 10.1038/530144a.

[168] Top500.org. *Supercomputer Statistics Sublist Generator*. 2017. URL: https://www.top500.org/statistics/sublist/.

[169] J.C. McCallum. *Memory Prices 1957–2017*. 2017. URL: http://www.jcmit.com/memoryprice.htm.

[170] D.J. Becker, T. Sterling, D. Savarese, et al. "Beowulf: A Parallel Workstation for Scientific Computation". In: *International Conference on Parallel Processing '95*. 1995.

[171] G.M. Amdahl. "Validity of the single processor approach to achieving large scale computing capabilities". In: AFIPS spring joint computer conference, 1969. URL: http://www-inst.eecs.berkeley.edu/~n252/paper/Amdahl.pdf.

[172] M. Karplus and A. McCammon. "Molecular dynamics simulations of biomolecules". In: *Nature Structural Biology* 9.9 (2002), pp. 646–652. DOI: 10.1038/nsb0902-646.

[173] The Pande Group. *Folding@Home statistics (2016-03-09)*. 2016. URL: https://folding.stanford.edu/home/.

[174] Top500.org. *Tianhe-2 (MilkyWay-2): National University of Defense Technology*. 2015. URL: http://www.top500.org/featured/top-

systems/tianhe-2-milkyway-2-national-university-of-defense/.

[175] Nvidia Corporation. *NVIDIA's next Generation CUDA architecture: Fermi*. 2009.

[176] J. Sanders and E. Kandrot. *CUDA BY EXAMPLE: An Introduction to General-Purpose GPU Programming*. Addison–Wesley, 2010, pp. 1–11.

[177] J.A. Butts, B. Batson, J.C. Chao, et al. "The Anton 2 Chip: A 2nd Generation ASIC for Molecular Dynamics Simulation". In: Presented at Hot Chips 2014, 14. URL: http://www.hotchips.org/wp-content/uploads/hc_archives/hc26/HC26-11-day1-epub/HC26.11-1-High-Performance-epub/HC26.11.130-Anton-2-Butts-Shaw-Shaw-Res-Search.pdf.

[178] Y. Gu and M.C. Herbordt. "FPGA-Based Multigrid Computation for Molecular Dynamics Simulations". In: Presented at 2007 International Symposium on Field-Programmable Custom Computing Machines, 2007, pp. 117–126. URL: https://www.bu.edu/caadlab/fccm07_MG.pdf.

[179] M. Chiu and M. Herbordt. "Towards Production FPGA-Accelerated Molecular Dynamics: Progress and Challenges". In: Presented at High-Performance Reconfigurable Computing Technology

and Applications 2010, 2010. URL: www.ncsa.illinois.edu/Conferences/HPRCTA10/presentations/chiu.pdf.

[180] M. Wielgosz, G. Mazur, M. Makowski, et al. "Analysis of the Basic Implementation Aspects of Hardware-Accelerated Density Functional Theory Calculations". In: *Computing and Informatics* 29 (2010), pp. 989–1000. URL: http://www.cai.sk/ojs/index.php/cai/article/viewFile/125/104.

[181] M. Wielgosz, E. Jamro, P. Russek, et al. "FPGA Implementation of the Exchange-Correlation Potential Calculation for DFT". In: *Automatyka* 15.3 (2011), pp. 485–498. URL: http://journals.bg.agh.edu.pl/AUTOMATYKA/2011-03/Auto_2011_3_36.pdf.

[182] Theoretical and Computational Biophysics Group. *Modeling Biological Processes at Hybrid Resolutions*. 2013. URL: https://www-s.ks.uiuc.edu/Research/cgfolding/.

[183] T. Sommerfeld. "Lorentz Trial Function for the Hydrogen Atom: A Simple, Elegant Exercise". In: *Journal of Chemical Education* 88 (2011), pp. 1521–1524. DOI: 10.1021/ed200040e.

[184] P. Hohenberg and W. Kohn. "Inhomogeneous Electron Gas". In: *Physical Review B* 136.3 (1964), pp. 864–871. DOI: 10.1103/PhysRev.136.B864.

[185] J.P. Perdew, A. Ruzsinsky, J. Constantin L.A. Sun, et al. "Some Fundamental Issues in Ground-State Density Functional Theory: A Guide for the Perplexed". In: *Journal of Chemical Theory and Computation* 5.4 (2009), pp. 902–908. DOI: 10.1021/ct800531s.

[186] J. Simons. "Theoretical Chemistry, 1e". In: (2003), pp. 208–212.

[187] I.N. Levine. *Quantum Chemistry, 4e*. Prentice Hall, 1991, p. 521.

[188] M. Swart, F.M. Bickelhaupt, and M. Duran. *DFT2016 Poll*. 2016. URL: http://www.marcelswart.eu/dft-poll/news2016.pdf.

[189] M.G. Medvedev, I.S. Bushmarinov, J. Sun, et al. "Density functional theory is straying from the path toward the exact functional". In: *Science* 355 (2017), pp. 49–52. DOI: 10.1126/science.aah5975.

[190] G.F. Giuliani and G. Vignale. *Quantum Theory of the Electron Liquid*. Cambridge, 2008, p. 70.

[191] I.N. Levine. *Quantum Chemistry, 4e*. Prentice Hall, 1991, p. 524.

[192] C.C. Cramer. "Essentials of Computational Chemistry". In: (2003), pp. 242–248.

[193] J. Tao, J.P. Perdew, V.N. Staroverov, et al. "Climbing the Density Functional Ladder: Nonempirical Meta–Generalized Gradient

Approximation Designed for Molecules and Solids". In: *Physical Review Letters* 91.14 (2003), pp. 146401-1–146401-4. DOI: 10.1103/PhysRevLett.91.146401.

[194] R. Baer and D. Neuhauser. "A density functional theory with correct long-range asymptotic behavior". In: *Physical Review Letters* 94.4 (2005), pp. 043002-1–043002-4. DOI: 10.1103/PhysRevLett.94.043002.

[195] S. Grimme. "Density functional theory with London dispersion corrections". In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 1.2 (2011), pp. 211–222. DOI: 10.1002/wcms.30.

[196] M. Born and J.R. Oppenheimer. "On the Quantum Theory of Molecules". In: *Annalen der Physik* 84 (1927), pp. 457–484. DOI: 10.1007/978-3-662-44593-8_1.

[197] I.N. Levine. *Quantum Chemistry, 4e*. Prentice Hall, 1991, p. 506.

[198] S. Pisana, M. Lazzeri, . Casiraghi C, et al. "Breakdown of the adiabatic Born–Oppenheimer approximation in graphene". In: *Nature Materials* 6 (2007), pp. 198–202. DOI: 10.1038/nmat1846.

[199] L. Noodleman. "Valence bond description of antiferromagnetic coupling in transition metal dimers". In: *Journal of Chemical Physics* 74 (1981), pp. 5737–5743. DOI: 10.1063/1.440939.

[200] L. Genovese, A. Neelov, S. Goedecker, et al. "Daubechies wavelets as a basis set for density functional theory pseudopotential calculations". In: *Journal of Chemical Physics* 129 (2008), pp. 014109-1–014109-14. DOI: 10.1063/1.2949547.

[201] J. Simons. "Theoretical Chemistry, 1e". In: (2003), pp. 190–191.

[202] C.C. Cramer. "Essentials of Computational Chemistry". In: (2003), pp. 156–164.

[203] C.C. Cramer. "Essentials of Computational Chemistry". In: (2003), pp. 166–167.

[204] P.J. Mohr, B.N. Taylor, and D.B. Newell. "CODATA recommended values of the fundamental physical constants: 2010". In: *Reviews of Modern Physics* 84.4 (2012), pp. 1527–1605. DOI: 10.1103/RevModPhys.84.1527.

[205] P. Pyykkö. "Relativistic Effects in Structural Chemistry". In: *Chemical Reviews* 88 (1988), pp. 563–594. DOI: 10.1021/cr00085a006.

[206] E. van Lenthe, E.J. Baerends, and J.G. Snijders. "Relativistic regular two-component Hamiltonians". In: *Journal of Chemical Physics* 99 (1992), pp. 4597–4610. DOI: 10.1063/1.466059.

[207] M.L. Cowan, B.D. Bruner, N. Huse, et al. "Ultrafast memory loss and energy redistribution in the hydrogen bond network of liquid

H_2O". In: *Nature* 434.7030 (2005), pp. 199–202. DOI: 10.1038/nature03383.

[208] T. van der Post, C.–S. Hsieh, M. Okuno, et al. "Strong frequency dependence of vibrational relaxation in bulk and surface water reveals sub-picosecond structural heterogeneity". In: *Nature Communications* 6.8384 (2015), pp. 1–7. DOI: 10.1038/ncomms9384.

[209] J. Tomasi, N. Mennucci, and R. Cammi. "Quantum Mechanical Continuum Solvation Models". In: *Chemical Reviews* 105.8 (2005), pp. 2999–3094. DOI: 10.1021/cr9904009.

[210] M. Svensson, S. Humbel, R.D.J. Froese, et al. "ONIOM: A Multilayered Integrated MO + MM Method for Geometry Optimizations and Single Point Energy Predictions. A Test for Diels–Alder Reactions and $Pt(P(t\text{–}Bu)_3)_2 + H_2$ Oxidative Addition". In: *Journal of Physical Chemistry* 100 (1996), pp. 19357–19363. DOI: 10.1021/jp962071j.

[211] M. Swart and F.M. Bickelhaupt. "QUILD: QUantum-Regions Interconnected by Local Descriptions". In: *Journal of Computational Chemistry* 29.5 (2007), pp. 724–735. DOI: 10.1002/jcc.20834.

[212] J. Simons. "Theoretical Chemistry, 1e". In: (2003), pp. 105–115.

[213] F.H.M. Koua, Y. Umena, K. Kawakami, et al. "Structure of Sr-substituted photosystem II at 2.1 Å resolution and its implications

in the mechanism of water oxidation". In: *Proceedings of the National Academy of Science* 110.10 (2013), pp. 3889–3894. DOI: 10.1073/pnas.1219922110.

[214] D.F. Ghanotakis, G.T. Babcock, and C.F. Yocum. "Structure of the oxygen-evolving complex of Photosystem II: Calcium and lanthanum compete for sites on the oxidizing side of Photosystem II which control the binding of water-soluble polypeptides and regulate the activity of the manganese complex". In: *Biochimica et Biophysica Acta - Bioenergetics* 809.2 (1985), pp. 173–180. DOI: 10.1016/0005-2728(85)90060-X.

[215] J.S. Vrettos, A. Stone, and G.W. Brudvig. "Quantifying the Ion Selectivity of the Ca^{2+} Site in Photosystem II: Evidence for Direct Involvement of Ca^{2+} in O_2 formation". In: *Biochemistry* 40 (2001), pp. 7937–7945. DOI: 10.1021/bi010679z.

[216] K.R. Long, Van Gosen, B.S., et al. *The Principal Rare Earth Elements Deposits of the United States—A Summary of Domestic Deposits and a Global Perspective*. United States Geological Survey (U.S. Department of the Interior), 2010, p. 5. URL: http://pubs.usgs.gov/sir/2010/5220/downloads/SIR10-5220.pdf.

[217] F. Zeng, Y. An, L. Ren, et al. "Effects of Lanthanum and Calcium on Photoelectron Transport Activity and the Related Protein Complexes in Chloroplast of Cucumber Leaves". In: *Biological Trace Element Research* 77 (2000), pp. 83–91. DOI: 10.1385/BTER:77:1:83.

[218] R.C. Weast. "Crystal Ionic Radii of the Elements". In: *Handbook of Chemistry and Physics*. Ed. by R.C. Weast. Boca Raton, Florida: CRC Press, 1983–1984, F170.

[219] R.J. Debus, M.A. Stricker, L.M. Walker, et al. "No evidence from FTIR difference spectroscopy that aspartate-170 of the D1 polypeptide ligates a manganese ion that undergoes oxidation during the S0 to S1, S1 to S2, or S2 to S3 transitions in photosystem II". In: *Biochemistry* 44.5 (2005), pp. 1367–1374. DOI: 10.1021/bi047558u.

[220] H.–A. Chu, R.J. Debus, and G.T. Babcock. "D1-Asp170 Is Structurally Coupled to the Oxygen Evolving Complex in Photosystem II As Revealed by Light-Induced Fourier Transform Infrared Difference Spectroscopy". In: *Biochemistry* 40.7 (2001), pp. 2312–2316. DOI: 10.1021/bi0022994.

[221] M.A. Strickler, W. Hillier, and R.J. Debus. "No Evidence from FTIR Difference Spectroscopy that Glutamate-189 of the D1 Polypeptide Ligates a Mn Ion that Undergoes Oxidation During the S_0

to S_1, S_1 to S_2, or S_2 to S_3 Transitions in Photosystem II". In: *Biochemistry* 45.29 (2006), pp. 8801–8811. DOI: 10.1021/bi060583a.

[222] M.A. Strickler, L.M. Walker, W. Hillier, et al. "No Evidence from FTIR Difference Spectroscopy That Aspartate-342 of the D1 Polypeptide Ligates a Mn Ion That Undergoes Oxidation during the S_0 to S_1, S_1 to S_2, or S_2 to S_3 Transitions in Photosystem II". In: *Biochemistry* 46.11 (2007), pp. 3151–3160. DOI: 10.1021/bi062195e.

[223] H.A. Chu, W. Hillier, and R.J. Debus. "Evidence that the C-terminus of the D1 polypeptide of photosystem II is ligated to the manganese ion that undergoes oxidation during the S1 to S2 transition: an isotope-edited FTIR study". In: *Biochemistry* 43.11 (2004), pp. 3152–3166. DOI: 10.1021/bi035915f.

[224] W.Y. Chuah, R. Stranger, R.J. Pace, et al. "Ab Initio Modeling of the Effect of Oxidation Coupled with H_nO Deprotonation on Carboxylate Ligands in Mn/Ca Clusters". In: *Journal of Physical Chemistry B* 118.13 (2014), pp. 3553–3558. DOI: 10.1021/jp500362q.

[225] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2016. URL: https://www.R-project.org/.

[226] M. Retegan and D.A. Pantazis. "Interaction of methanol with the oxygen-evolving complex: atomistic models, channel identification, species dependence, and mechanistic implications". In: *Chemical Science* 7 (2016), pp. 6463–6476. DOI: 10.1039/c6sc02340a.

[227] Brookhaven National Laboratory National Nuclear Data Center. "Current Version of Nuclear Wallet Cards, $Z = 25$". In: (2012). URL: http://www.nndc.bnl.gov/nudat2/wcbyz.jsp?z=25.

[228] A. Carrington and A.D. McLachlan. *Introduction to Magnetic Resonance*. 1st ed. New York: Harper and Row, 1967, p. 167.

[229] A. Carrington and A.D. McLachlan. *Introduction to Magnetic Resonance*. 1st ed. New York: Harper and Row, 1967, pp. 99–103.

[230] J. Autschbach, S. Patchkovskii, and B. Pritchard. "Calculation of Hyperfine Tensors and Paramagnetic NMR Shifts Using the Relativistic Zeroth-Order Regular Approximation and Density Functional Theory". In: *Journal of Chemical Theory and Computation* 7.7 (2011), pp. 2175–2188. DOI: 10.1021/ct200143w.

[231] J. Autschbach. "Magnitude of Finite-Nucleus-Size Effects in Relativistic Density Functional Computations of Indirect NMR Nuclear Spin–Spin Coupling Constants". In: *ChemPhysChem* 10.13 (2009), pp. 2274–2283. DOI: 10.1002/cphc.200900271.

[232] E.A. Schmitt, L. Noodleman, E.J. Baerends, et al. "LCAO Xα calculation of the magnetic exchange interactions in a manganese $Mn^{IV}Mn_3{}^{III}$ cubane complex: relevance to the water oxidation center of photosystem II". In: *Journal of the American Chemical Society* 114.15 (1992), pp. 6109–6119. DOI: 10.1021/ja00041a032.

[233] J. Ahrens, Gevici B., and C. Law. *ParaView: An End-User Tool for Large Data Visualization*. Elsevier, 2005. ISBN: 978-0123875822.

[234] U. Ayatchit. *The ParaView Guide: A Parallel Visualization Application*. Kitware, 2015. ISBN: 978-1930934306.

[235] D.B. Ennis, G. Kindlman, I. Rodriguez, et al. "Visualization of Tensor Fields Using Superquadric Glyphs". In: *Magnetic Resonance in Medicine* 53 (2005), pp. 169–176. DOI: 10.1002/mrm.20318.

[236] D. Goldberg. "What Every Computer Scientist Should Know About Floating-Point Arithmetic". In: *ACM Computing Surveys* 23.1 (1991), pp. 5–48. DOI: 10.1145/103162.103163.

[237] F. Neese. "The ORCA program system". In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 2.1 (2012), pp. 73–78. ISSN: 1759-0884. DOI: 10.1002/wcms.81. URL: http://dx.doi.org/10.1002/wcms.81.

[238] G. Alyward and T. Findlay. *SI Chemical Data 5e*. Milton, Australia: Wiley, 2002, p. 120.

[239] J. Bezanson, S. Karpinski, V.B. Shah, et al. "Julia: A Fast Dynamic Language for Technical Computing". In: (Sept. 2012). arXiv: 1209.5145 [cs.PL].

[240] J. Bezanson, A. Edelman, S. Karpinski, et al. "Julia: A Fresh Approach to Numerical Computing". In: (Nov. 2014). arXiv: 1411.1607 [cs.MS].

[241] J. Bezanson, J. Chen, S. Karpinski, et al. "Array operators using multiple dispatch: a design methodology for array implementations in dynamic languages". In: *ARRAY'14 Proceedings of ACM SIGPLAN International Workshop on Libraries, Languages, and Compilers for Array Programming*. New York, NY, USA: ACM, 2014, pp. 56–61. DOI: 10.1145/2627373.2627383. arXiv: 1407.3845 [cs.PL].

[242] K.E. Drexler. "Engines of Creation". In: (2007).