

Substrate-Induced Perturbations to the Active  
Site Electronic Structure of AA10 Lytic  
Polysaccharide Monooxygenases

Peter J. Lindley

PhD

University of York

Chemistry

September 2022

# Acknowledgements

The last four years have been tremendously formative for me as both a scientist and as a person. I would like to thank so many people from whom I have learned so much - both in and out of the lab. Firstly, I would like to thank my PhD supervisors (Paul Walton, Gideon Davies and Alison Parkin) and the University of York chemistry department for providing me the opportunity to partake in such an exciting PhD program at a prestigious university. I have been fortunate to work alongside some truly fantastic scientific minds. A particular thank you goes to Paul Walton who has continually stretched and supported me throughout my project. – I will greatly miss our scientific group meetings. I would also like to thank all of my collaborators for their support in my work and for including me in theirs; it has been a pleasure to work alongside you and I have learned a lot from them.

Moreover, I would like to thank Dr Luisa Ciano and Dr Alessandro Paradisi for being such fabulous mentors for me. You both went above and beyond to guide me and make me feel welcome in the group. Thank you for teaching me SO much and making the journey enjoyable along the way. I miss you both like mad and wish you all the best in your futures. Martino, thank you for being my best friend throughout the PhD. You have helped me so much in the lab over the years – I mean you even managed to get me to do a reflux for crying out loud! You're a fantastic chemist and again I am lucky to have worked with you. But by far the biggest thing, you helped me so much outside of work. You have been a great climbing buddy, chess player and pub partner. You seem to spot when I am struggling and always spare the time to make me laugh and pick me up. Thanks for everything! - And sorry for dropping you on belay... I should have paid more attention!

I would like to thank entire my family for supporting me through these challenging years. Particularly my mum, Dawn, and my stepdad, Phil. You guys have been my biggest cheerleaders from the beginning, and I would not be where I am now without you. Mum, you're perhaps the most resilient and headstrong human I have ever met, and I believe it is a snippet of those genes that have helped me to get here! Phil has also shown me an immense amount of support and has made such an effort to get me to believe in myself as much as he does. You guys have been my rock so thank you. I would also like to thank my mother and father-in-law, Rob and Lenore Klassen (and the late family cat, Minou). They both have been unbelievably supportive of me and made these last years in York some of the best of my life so far. They rather unexpectedly hosted me during the Cov-19 pandemic where their support, laughter and cooking made some extremely challenging times truly special.

Finally, I would like to thank my wife, Dani. You are truly wonderful and have been there for me during these stressful times. You are so supportive, and I am honestly not sure if I could have got through it without you. Thank you for always making me laugh and making me feel loved. You make my highs so much higher and my lows a whole lot less low. You have patiently listened for hours and hours of science totally irrelevant to your own work. But at least now you somehow find yourself knowing more about EPR spectroscopy than most chemistry graduates. Thanks for being my rock and I now look forward to all our adventures together in our coming years. I love you.

# Abstract

The following PhD thesis contains a collection of published papers and draft manuscripts that explore the variable behaviour of AA10 lytic polysaccharide monooxygenase (LPMO) enzymes using a range of spectroscopic techniques and theoretical methods. It primarily focusses on the changes that occur at the active sites of AA10 LPMOs caused by the binding of their polysaccharide substrates. We highlight significant alterations to the copper ligand field of AA10 LPMOs after binding crystalline  $\beta$ -chitin using techniques such as EPR spectroscopy and DFT calculations. Principally, we discovered that substrate binding is responsible for driving a change in electronic structure that is commensurate with the loss of a ligand (water) leading to 4-coordinate LPMO-Cu(II) complexes. We propose this imposed 4-coordinate geometry forms the basis of a substrate-coupled  $O_2$  activation pathway among AA10 LPMOs.

This putative substrate-coupled  $O_2$  activation mechanism was then investigated further by using the azide ion ( $N_3^-$ ) to mimic the electronic behaviour of the superoxide ion ( $O_2^-$ ). On this basis, we prepared a range of LPMO-Cu(II)- $N_3$  complexes at the active site of a chitin-active AA10 LPMO, the spectroscopic characteristics of which, by analogy, shed light on the elusive LPMO-Cu(II)- $O_2$  species. DFT calculations revealed that this reduced coordination number leads to the generation of more potent oxidising intermediates when compared to the 5-coordinate geometry in the resting state of the enzymes. Such a finding is significant to the field of LPMOs, reinforcing the hypothesis of a coupling mechanism existing between substrate binding and  $O_2$  activation. Such a mechanism explains how AA10 LPMOs perform highly selective C-H bond activations by only producing the most reactive intermediates when the substrate is correctly positioned at the enzyme active site.

# Table of Contents

<b>Acknowledgements</b> .....	2
<b>Abstract</b> .....	3
<b>List of Abbreviations</b> .....	7
<b>Declaration</b> .....	9
<b>Structure of Thesis</b> .....	9
<b>1 Introduction</b> .....	10
<b>1.1 Bioethanol Production from Biomass Saccharification</b> .....	10
<b>1.2 Polysaccharide Structure</b> .....	11
<b>1.2.1 Cellulose</b> .....	14
<b>1.2.2 Hemicellulose</b> .....	15
<b>1.2.3 Lignin</b> .....	16
<b>1.2.4 Pectin</b> .....	18
<b>1.2.5 Chitin and Chitosan</b> .....	18
<b>1.3 Synergistic Enzyme Cocktails to Degrade Lignocellulosic Biomass</b> .....	21
<b>1.3.1 Glycosidic Hydrolases</b> .....	22
<b>1.3.1 Lytic Polysaccharide Monooxygenases</b> .....	24
<b>1.4 Copper Sites in Biology</b> .....	28
<b>1.4.1 Type I ‘Blue’ Copper Proteins</b> .....	30
<b>1.4.2 Type II Copper Proteins</b> .....	31
<b>1.4.3 Type III Copper Proteins</b> .....	34
<b>1.5 The ‘Histidine Brace’ – An Alternative Copper Site</b> .....	36
<b>1.5.1 The AA9s</b> .....	39

1.5.2	The AA10s .....	41
1.5.3	The AA11s .....	43
1.5.4	The AA13s .....	44
1.5.5	The AA15s .....	46
1.5.6	The AA16s .....	48
1.5.7	The AA17s .....	48
1.6	LPMO Protein Preparation .....	49
1.6.1	Periplasmic Expression .....	51
1.6.2	Cytoplasmic Expression .....	52
1.6.3	Protein Purification .....	52
1.7	Electron Paramagnetic Resonance (EPR) Spectroscopy .....	54
1.7.1	EPR Theory .....	54
1.7.1.1	The Electronic <i>g</i> -factor .....	55
1.7.1.2	Hyperfine Coupling – the Effect of <i>Nuclear Spin</i> .....	60
1.7.1.3	The Effective and Spin Hamiltonian .....	62
1.7.2	The EPR Experiment .....	63
1.7.2.1	Continuous Wave (cw) EPR .....	64
1.7.2.2	Advanced EPR Techniques (Pulsed and HYSORE) .....	66
1.7.2.3	Simulation of EPR Spectra .....	68
1.8	Computational and Theoretical Studies .....	72
1.8.1	Level of Theory - Functionals and basis sets .....	73
1.8.2	Geometry Optimizations .....	74
1.8.3	EPR Property Calculations .....	77
1.9	On the LPMO Catalytic Mechanism .....	79
1.9.1	Oxygenase vs Peroxygenase Mechanism .....	81

1.9.2	Computational Insights .....	82
1.9.3	Stopped-flow Spectrophotometry.....	87
	References .....	90
2	Research Papers .....	97
3	Mapping the protonation states of the histidine brace in an AA10 lytic polysaccharide monooxygenase using CW-EPR spectroscopy and DFT calculations .....	99
4	Mechanistic basis of substrate–O <sub>2</sub> coupling within a chitin-active lytic polysaccharide monooxygenase: An integrated NMR/EPR study .....	134
5	On the Mechanism of Substrate-O <sub>2</sub> Coupling in Lytic Polysaccharide Monooxygenases: EPR, X-Ray Crystallography, and DFT Studies of Azide Binding to a Chitin-Active LPMO.....	203
6	Concluding Remarks .....	237
7	Appendix of Research Papers .....	239
7.1	Secreted pectin monooxygenases drive plant infection by pathogenic oomycetes	239
7.2	C-type cytochrome-initiated reduction of bacterial lytic polysaccharide monooxygenases.....	246

# List of Abbreviations

AA – Auxiliary activity	ESI MS – Electrospray ionisation mass spectrometry
$B$ – Magnetic field	FGE – Formylglycine generating enzyme
$B_e$ - Intrinsic electron magnetic field	FT-IR - Fourier-transform infrared
$B_0$ - Applied magnetic field	$g_e$ – Free electron g-factor
$B_{tot}$ - Total magnetic field	GGA – Generalised gradient approximation
CAPS – N-cyclohexyl-3-aminopropanesulfonic acid	GHG – Greenhouse gas
CASSCF – Complete active space self-consistent field theory	GH – Glycosidic hydrolase
CAZy or CAZymes - Carbohydrate active enzymes	HAA – Hydrogen atom abstraction
CBM – Chitin binding module	HAT – Hydrogen atom transfer
CBH – Cellobiohydrolases	HEPES – 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
CDH – Cellobiose dehydrogenase	HF – Hartree-Fock
CHES – N-Cyclohexyl-2-aminoethanesulfonic acid	HFC – Hyperfine coupling
CW – Continuous wave	HFX – Hartree-Fock exchange
CYP – Cytochrome P450	HILIC – Hydrophilic interaction chromatography
CPCM – Conductor-like polarizable continuum model	HOMO – Highest occupied molecular orbital
DELFT – DFT enhanced ligand field theory	HSQC – Heteronuclear single quantum coherence
DFT – Density functional theory	HYSCORE – Hyperfine sub-level correlation spectroscopy
EC – Enzyme classification	IPTG - Isopropyl $\beta$ -D-1-thiogalactopyranoside
EG - Endoglucanase	ISC – Intersystem crossing
EDTA – Ethylenediaminetetraacetic acid	ISSET – Inner-sphere electron transfer
ENDOR – Electron nuclear double resonance spectroscopy	ITC – Isothermal calorimetry
EPR – Electron paramagnetic resonance	$J$ – Total angular momentum quantum number
ESEEM – Electron spin echo envelope modulation	JT – Jahn Teller
	LB – Luria Bertani

LDA – Local density approximation

LFT – Ligand field theory

LMCT – Ligand to metal charge transfer

LPMO – Lytic polysaccharide monooxygenase

LRET – Long-range electron transfer

LUMO – Lowest unoccupied molecular orbital

MALDI-TOF – Matrix assisted laser desorption ionisation – time of flight

mCPBA - meta-Chloroperoxybenzoic acid

MD – Molecular dynamics

MES – 2-(N-morpholino)ethanesulfonic acid

MM – Molecular mechanics

MO – Molecular orbital

NMR – Nuclear magnetic resonance

NOE – Nuclear Overhauser effect

NOESY – Nuclear Overhauser effect spectroscopy

OD<sub>600</sub> – Optical density (600 nm)

OSET – Outer-sphere electron transfer

PASC – Phosphoric acid swollen cellulose

PB – Peisach Blumberg

PCET – Proton coupled electron transfer

PES – Potential energy surface

PHM –Peptidylglycine alpha-amidating monooxygenase

PMO – Polysaccharide monooxygenase

PRE – Paramagnetic relaxation enhancement

PTM – Post translational modification

QM – Quantum mechanics

QM-MM – Quantum mechanics – molecular mechanics

RAMO – Redox active molecular orbital

ROS – Reactive oxygen species

S – Spin angular momentum quantum number

SDS-PAGE – Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

SEC – Size exclusion chromatography

SH – Spin-Hamiltonian

SOC ( $\lambda$ ) – Spin orbit coupling

SOMO – Singly occupied molecular orbital

SUMO - Small Ubiquitin-like Modifier

SHFC - Superhyperfine coupling

SVD – Singular value decomposition

TB – Terrific broth

TD-DFT – Time dependent density functional theory

TEMPO - 2,2,6,6-tetramethylpiperidine- N - oxyl

TEV - Tobacco Etch Virus

TM – Transition metal

VT – Variable temperature

XRD – X-ray diffraction

1G – First generation

2G – Second generation

$\alpha^2$  – Covalency parameter

$h$  – Planck’s constant

$\mu_B$  – Bohr magneton

$\nu$  - Frequency

Z<sub>eff</sub> – Effective nuclear charge



# Declaration

I declare that the work presented in this thesis is my own for which I am the sole author unless otherwise stated. Before each chapter, I have provided a declaration outlining the contributions of all co-authors for the associated journal article. This work has not previously been presented for an award at this or any other University.

# Structure of Thesis

This PhD thesis comprises of 4 main chapters. The first chapter is an introduction which discusses the underlying chemistry relating to the field of LPMO enzymes. The core topics range from polysaccharide structure to polysaccharide-degrading enzymes and an overview of the key copper sites present in biology. Chapters 2-4 detail original research papers to which I have made major contributions.

Chapter 2 is an original first-author research article (2022, *Faraday Discussions*, RSC) detailing the unique electronic structures adopted by the bacterial enzyme, *BAA10*, with varying pH. This work uses a combination of EPR spectroscopy and DFT calculations to highlight three active site geometries with the pH range of 6-12.5 - undergoing two single deprotonation steps of active site ligands.

Chapter 3 contains an original research article (2020, *Proc. Natl. Acad. Sci. U.S.A*) from multiple co-authors (Gaston Courtade/Luisa Ciano joint 1<sup>st</sup>) which outlines a detailed study into substrate induced perturbations to the active site electronic structure. This research article draws upon the expertise of multiple researchers who helped to unravel how the binding of chitinous substrate to *BAA10* enzyme gives rise to an altered ligand field which has significant implications to the overall catalytic mechanism.

Chapter 4 is an original first-author research article (near-final draft manuscript) which builds on the aforementioned '2020, *Proc. Natl. Acad. Sci. U.S.A*' paper investigating the substrate-induced perturbations to the active site electronic structure of an AA10 LPMO, *BaAA10*. This paper uses a small molecule surrogate for  $O_2^-$ , the azide anion ( $N_3^-$ ), to mimic the electronic behaviour of a superoxide ligand to investigate the changes that occur to the active site upon oxygen activation. Using a combination of EPR spectroscopy, DFT calculations and UV-Vis spectroscopy, we reveal that AA10 LPMOs produce more potent oxidising intermediates when bound to their natural substrates.

# 1 Introduction

## 1.1 Bioethanol Production from Biomass Saccharification

Polysaccharides including cellulose, chitin, and starch are macro biopolymers made up of many repeating sugar units. They are often large carbohydrate molecules harnessing chemical potential. Due to their usefulness in structural rigidity, they are utilised ubiquitously throughout nature. Carbohydrates have diverse applications in biology such as water transportation in plants (transpiration), cell wall materials for plants/fungi, and exoskeletons for insects/crustaceans. The widespread use of these polysaccharides (in conjunction with their comparative chemical inertness) leads to a large amount of organic biomass produced by nature. Cellulosic biomass, in fact, makes up the largest source of organic carbon on the planet with a production rate of *ca.* 1.5 trillion ( $10^{12}$ ) tonnes annually.<sup>1</sup> Biomass comprising of these materials takes considerable time to decompose; typically mediated by suitable enzyme-equipped insects, fungi and bacteria.<sup>2-4</sup> The enormous, widespread abundance of these materials along with their energetic potential make lignocellulosic biomass an attractive candidate for valorisation. Moreover, saccharification of polysaccharides into mixed sugars offers a means of producing biofuel *via* fermentation.

Issues such as global warming, finite fossil fuels and energy security are among the greatest global challenges in the present day. Producing biofuel from this kind of biomass is advantageous since it is considered a 'waste' material. Biofuel produced in this way is known as "second generation" (2G) biofuel. 2G biofuel has advantages over the "first generation" (1G) biofuel alternative, in which food crops such as sugarcane and corn are grown purposefully for fuel production. Though the starch-rich substrates required for 1G biofuel production are more tractable towards both chemical processing and biological attack, land mass ordinarily used to grow food crops must be sacrificed to 'grow' fuel. As such, 1G biofuels are a frequently criticised solution to the growing global fuel crises, known as the "food *versus* fuel" argument.<sup>5,6</sup>

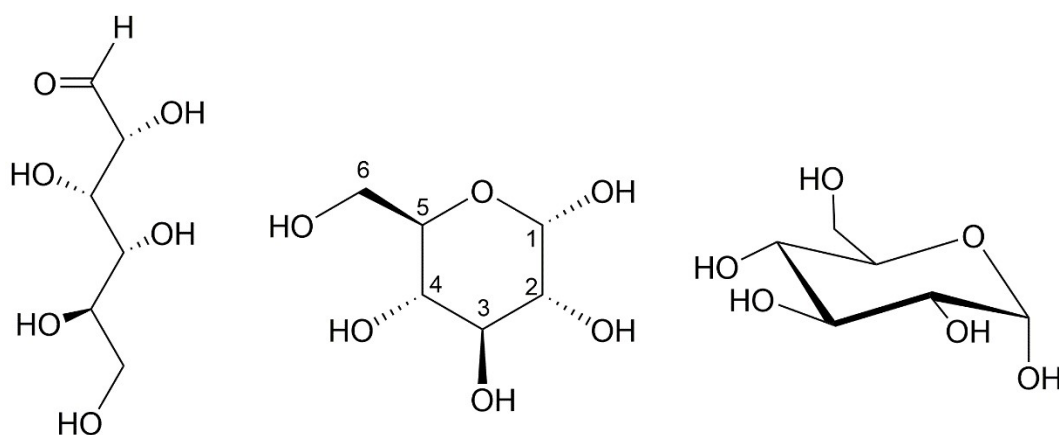
The relative ease of producing 1G biofuel from feedstocks such as sugarcane or corn has led to them being well implemented around the world. The steps involved in making biofuel from sugarcane are rather simple, with the basic crushing of the crop in water to extract sucrose, which is purified to raw sugar and later fermented into ethanol. Corn requires the additional pre-treatment step of hydrolysing the starches using amylase enzymes – adding further cost and difficulty to the process.<sup>7</sup> Brazil, one of the largest producers of both sugarcane and corn, are a world leader in the production

and utilisation of bioethanol in this manner.<sup>8</sup> This has been a largely positive step in relieving our dependence on fossil fuels. These processes have gone so far, that liquid biofuels now comprise 25% of the total transport fuel consumed in Brazil. However, the global picture still remains largely bleak with fossil fuels still making up some 80% of the global energy supply.<sup>9</sup> The need to move to low-carbon, sustainable fuel alternatives is now of paramount importance, with lignocellulosic biomass being a front runner in suitable replacements. This is most pertinent in the transport sector where our reliance on crude-oil derived petroleum and diesel products is exceptionally high. While it is considerably easier to produce 1G biofuel compared to 2G, production of biofuel using food-crop feedstocks is likely to be largely banned in the EU in the coming years. This has resulted from findings showing there are little-to no reductions in greenhouse gas emissions when producing fuel in this way. Furthermore, sacrificing increasing land space for 1G crops simultaneously places food, biodiversity and natural ecosystems under increased risk.<sup>10</sup> Food prices have been rapidly increasing since around 2006 correlated with the rising population.<sup>11</sup> Consequently, a growing number of people are finding themselves unable to afford sufficient food raising questions over the ethics of redirecting food crops for 1G biofuel.

Lignocellulosic biomass (comprising of cellulose, hemicellulose, and lignin) has additional benefits as biofuel feedstock as it typically derives from waste materials. Lignocellulosic material can stem from agricultural, forestry and food waste; requiring no additional land space to be dedicated toward growing fuel crops.<sup>11</sup> Lignocellulosic biomass from 2G feedstocks has a far greater fossil fuel displacement potential than 1G and results in much reduced greenhouse gas (GHG) emissions.<sup>12</sup> However, despite its obvious desirability, converting 2G biofuel feedstocks into fermentable reducing sugars is substantially more difficult compared to that of 1G.<sup>13</sup> In order to be useful for sugar saccharification, 2G biofuel feedstocks often require low lignin levels as the lignin has shown to significantly hinder the release of glucose units from the material. Typically, this waste material is processed into biofuel *via* multiple steps, including: thermochemical pre-treatment, enzymatic hydrolysis, fermentation, and finally distillation and dehydration.<sup>13</sup> The pre-treatment steps are necessary to maximize the surface area available for enzymatic attack, so that optimal saccharification can take place. This crystallinity stems from extensive hydrogen bonding networks forming between polysaccharide chains leading to insoluble materials that are resistant to hydrolysis. This resistance to degradation is the principal reason that 1G biorefineries have been more quickly implemented around the world, thus avoiding the difficulties of lignocellulose deconstruction.

## 1.2 Polysaccharide Structure

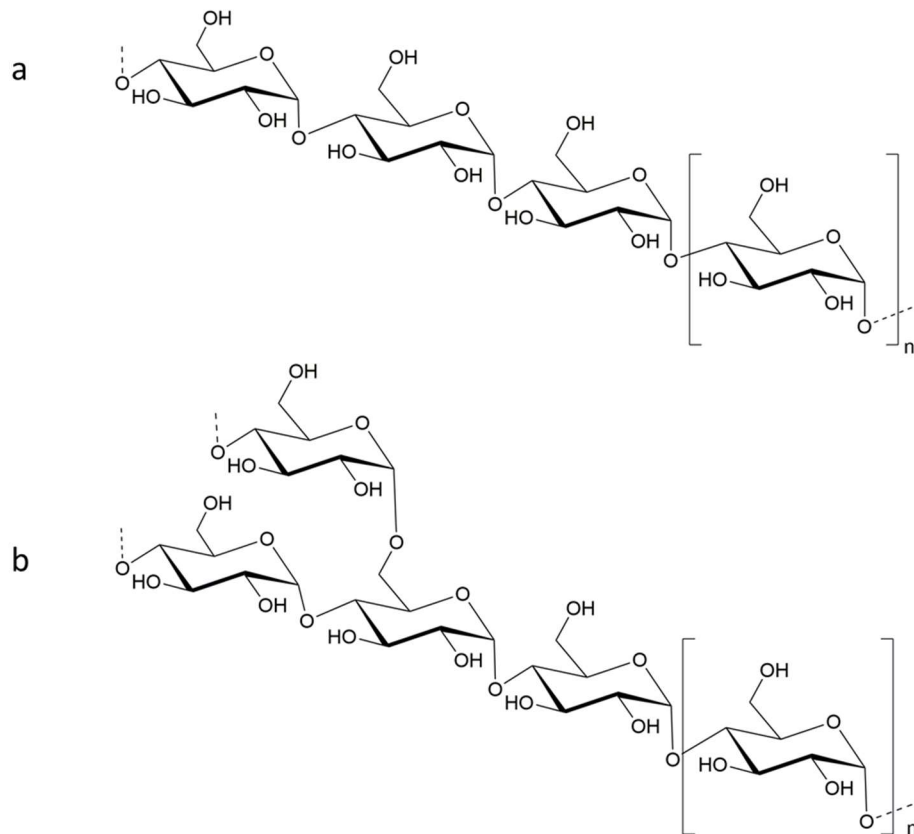
Polysaccharides are known by numerous different names, the most common alternatives including sugars and carbohydrates. Carbohydrates were defined as such as they commonly possess a structure consisting of just carbon, oxygen, and hydrogen. The ratio of hydrogen to oxygen is usually 2:1, making the overall chemical formula commensurate with a *hydrated* form of carbon given by the generic chemical formula:  $C_m(H_2O)_n$ . Single saccharide units can have different numbers of carbon atoms as part of their structure, with common numbers being 5 and 6, leading to pentose and hexose sugars, respectively. Saccharides can exist either in open-chain form (left, **Figure 1**) or in cyclic form (middle and right, **Figure 1**). In solution the vast majority of sugars, such as glucose, exist in the cyclic form (usually > 99%). Once cyclised, the sugar can adopt several structural configurations including boat, twist boat and chair, with the latter usually being the lowest energy and thus most stable (right, **Figure 1**).<sup>14</sup> Polysaccharide synthesis occurs *via* condensation reactions of monosaccharides whereby the removal of a water molecule leads to the formation of a glycosidic bond. Polysaccharides in biology serve a variety of diverse functions. These span from recognition processes (recognition of pathogen glycans is the cornerstone of innate host defence)<sup>15</sup>, to structural applications and energy storage.<sup>16</sup>



**Figure 1. Open chain structure of glucose (left), stereochemical view of the cyclic structure of glucose with numerical carbon labelling (middle) and glucose adopting the energetically favourable “chair” conformation (right).**

*Polysaccharides* are macromolecules consisting of many repeating sugar (saccharide) units *via* glycosidic linkages. Factors including the identity of the repeating sugar unit, the degree of branching and the stereochemistry all have profound influences on the polysaccharides overall macro properties. Starch for example is energy storage carbohydrate comprised of two types of *alpha*-glucan polymers, amylose, and amylopectin (**Figure 2 a and b**, respectively). Starches are archetypal

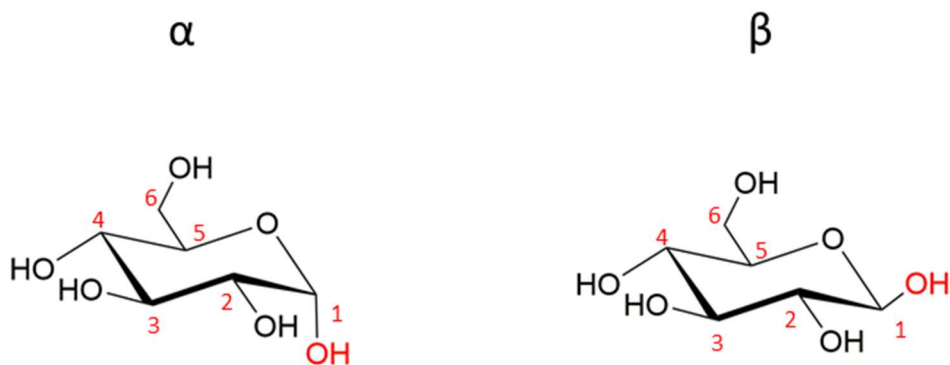
polysaccharides made up of repeating glucose units joined by  $\alpha$  1-4 glycosidic linkages. Amylopectin differs by containing branches created by forming 1-6 glycosidic linkages. A polysaccharide can be constituted of the same repeating saccharide monomer, known as a *homopolysaccharides*, or of alternating/different monosaccharide units, known as *heteropolysaccharides*, opening up the possibility of an unfathomable number of unique carbohydrates.<sup>17</sup> The specific synthesis of even a short (< 10 sugar units) polysaccharide is a major challenge for a synthetic chemist, yet is readily performed on an enormous scale in biological systems.



**Figure 2. Skeletal structure of the two major components of starch, amylose (a) and amylopectin (b).**

### 1.2.1 Cellulose

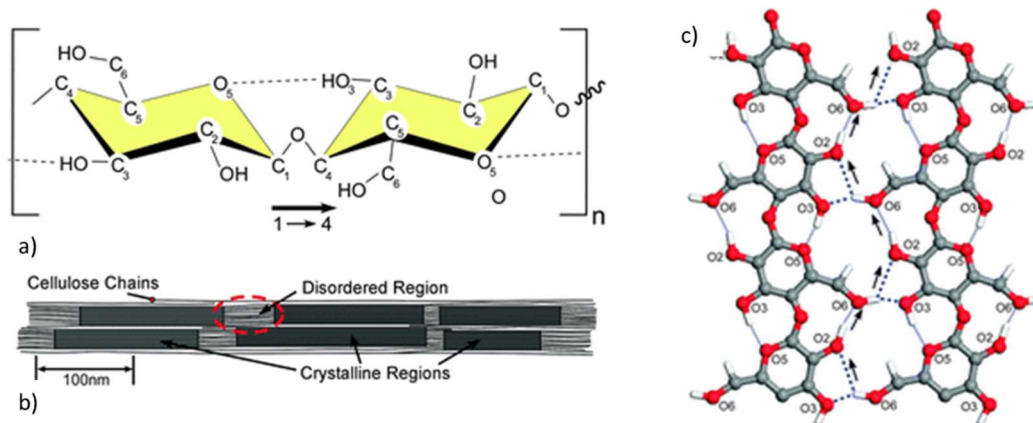
Cellulose is a polymer of glucose joined together by  $\beta$  1-4 glycosidic bonds. It is the most abundant polysaccharide and indeed the largest form of organic carbon on the planet.<sup>18</sup> It is mostly found among species of plant, comprising the major component of cell walls of plants and wood. Cellulose was first isolated from plant matter by a French scientist, Anselme Payen, in the late 1830s who discovered that despite possessing an analogous chemical composition to starch, the two materials possess greatly differing structures and properties.<sup>18</sup> This principal difference between these two carbohydrates is that starch is comprised of  $\alpha$ -glucose, whereby the hydroxyl group on the anomeric carbon is on the same side as the  $-\text{CH}_2\text{OH}$  group, and cellulose is comprised of  $\beta$ -glucose units, where the hydroxyl group on the anomeric carbon is on the opposite side of the  $-\text{CH}_2\text{OH}$  group (**Figure 3**). The isomers of the sugar units such as these are known as *anomers* as they differ solely on the orientation of the constituents on the anomeric carbon. Pure anomeric mixtures of sugars in solution spontaneously undergo tautomerism to reach an equilibrium position containing mixtures of both anomers, with glucose typically existing as a mixture of *ca.* one third alpha to two thirds beta.<sup>19</sup> Indeed, the exact equilibrium position of this polysaccharide tautomerism is dependent on the relative constituents present on the sugars and is dictated by the anomeric effect.<sup>20</sup>



**Figure 3. Skeletal diagram of alpha and beta glucose (left and right, respectively). Numerical labels represent common saccharide nomenclature, where 1 is the anomeric carbon. Oxygen atom shown in red highlights the different orientations adopted by the two anomers.**

The  $\beta$ -1-4 linked arrangement of glucose in cellulose leads to the formation of linear polysaccharide chains *without* branching. The lack of branching often leads to the formation of highly crystalline materials, differing considerably to the starch tautomer. These large and highly ordered materials are usually heterogeneous or totally insoluble, making them unsuitable as energy storage units for organisms, but exceptionally useful in structural applications. The crystallinity of the material

inhibits the action of most hydrolytic enzymes due to diffusion/mass transfer limitations and is a principal reason why cellulose is unused as an energy source for many organisms. The stark heterogeneity of the carbohydrate makes it largely intractable for hydrolysis *via* enzymatic attack. In the absence of energy storage applications, cellulose makes an ideal material for the construction of fibrous long chains (**Figure 4b**). With cellulose forming linear chains without branching, it allows for significant degrees of hydrogen bonding to occur both between the same chain (intramolecular) and neighbouring chains (intermolecular) (**Figure 4c**).

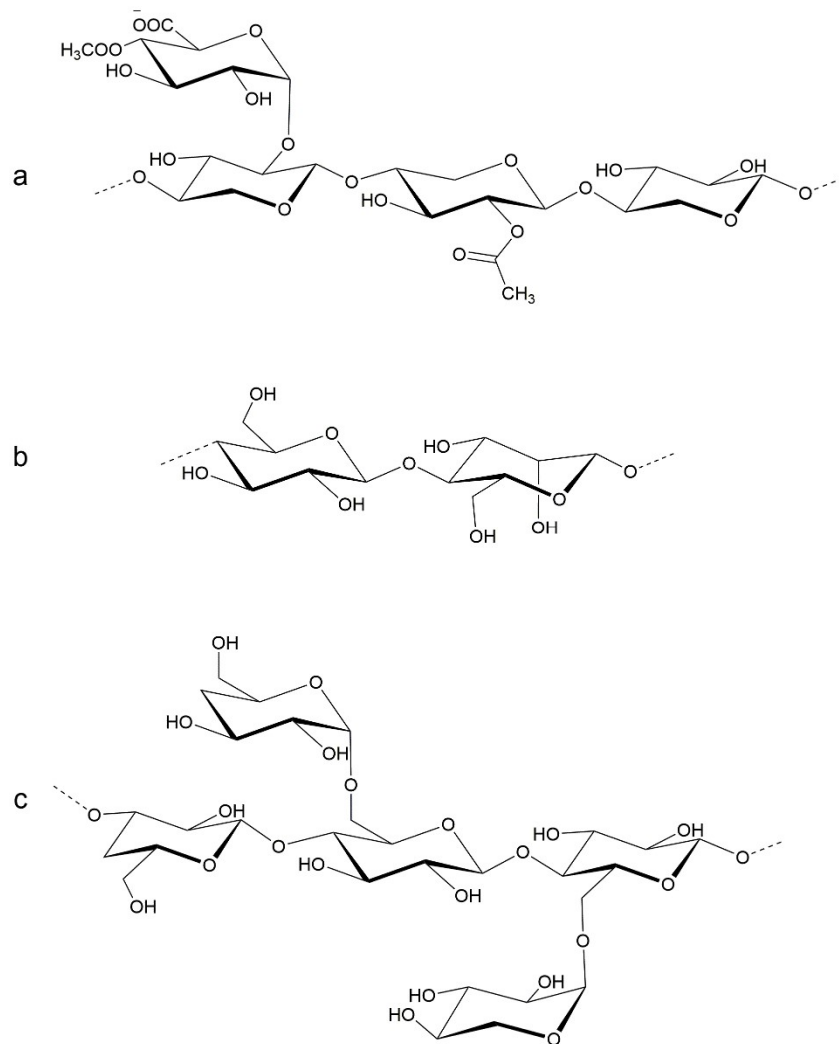


**Figure 4. Schematic of cellulose repeating unit highlighting the  $\beta$ -1-4 glycosidic bond (a), cellulose microfibril diagram highlighting both crystalline and amorphous regions (b) and a cooperative intermolecular hydrogen bond network between two cellulose chains (c). Figure adapted from Moon *et al.*<sup>21</sup>**

### 1.2.2 Hemicellulose

Hemicelluloses are the second largest component of all plant cell walls and function cooperatively with cellulose. They possess some key differences to cellulose; for example, the average chain length is significantly shorter (usually 500-300 units in length compared to a typical 7000-15,000 in cellulose).<sup>22</sup> They are heteropolysaccharides containing different sugar units and aligned in varying orientations. As such, the precise chemical nature of hemicelluloses varies significantly between different plant species. Some common examples of hemicelluloses include xylans, glucomannans and xyloglucans (**Figure 5 a, b and c**, respectively). Due to the alternating sugar units and high degree of branching, hemicellulose is an amorphous material.<sup>22, 23</sup> As a result, hemicelluloses are usually more easily broken down than cellulose, commensurate with their less ordered macrostructure which lends

itself more readily to enzymatic attack. Hemicellulose polysaccharides operate in conjunction with cellulose to help strengthen cell walls and form more intricate, 3-dimensional structures.

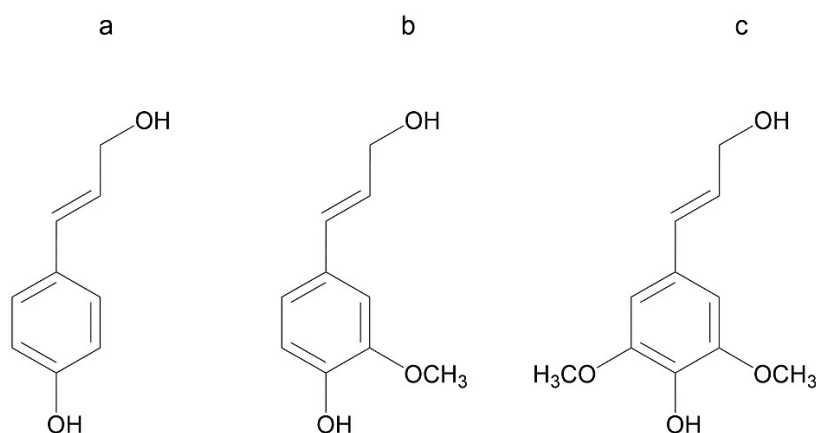


**Figure 5. Skeletal structures of hemicellulose examples, xylan (a), glucomannan (b) and xyloglucan (c).**

### 1.2.3 Lignin

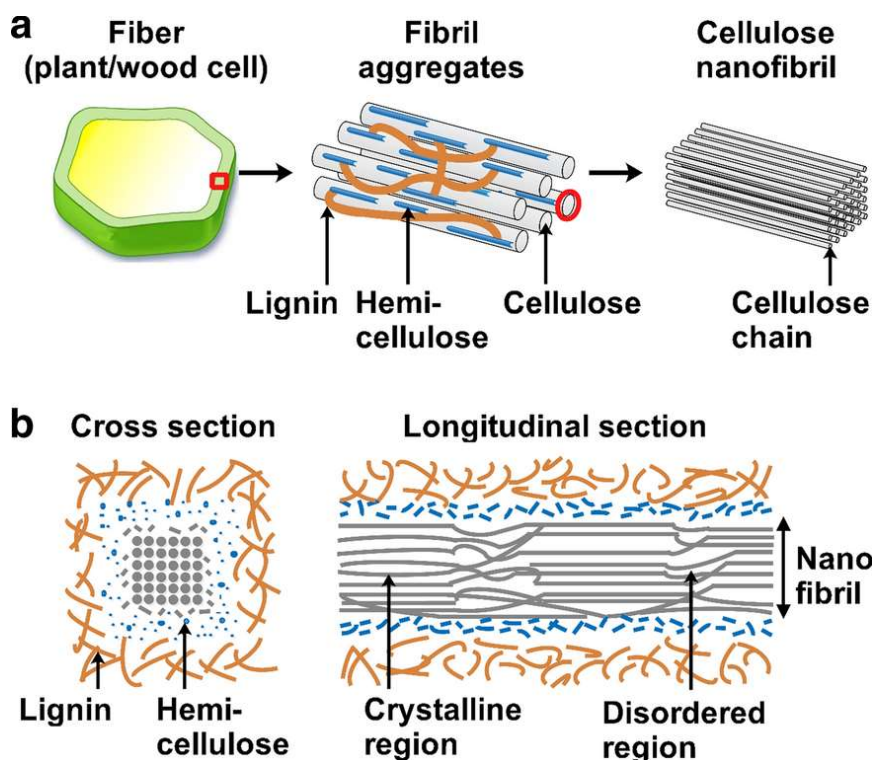
Lignin is another abundant plant material that helps add mechanical strength to the plant cell wall. Unlike the previously mentioned cell wall materials, these are phenylpropanoid polymers rather than polysaccharides as they do not contain saccharide units. They are constructed from phenolic building blocks, usually *para*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol (**Figure 6 a, b and c**, respectively).





**Figure 6. Lignin building blocks: *p*-coumaryl alcohol (a), coniferyl alcohol (b) and sinapyl alcohol (c).**

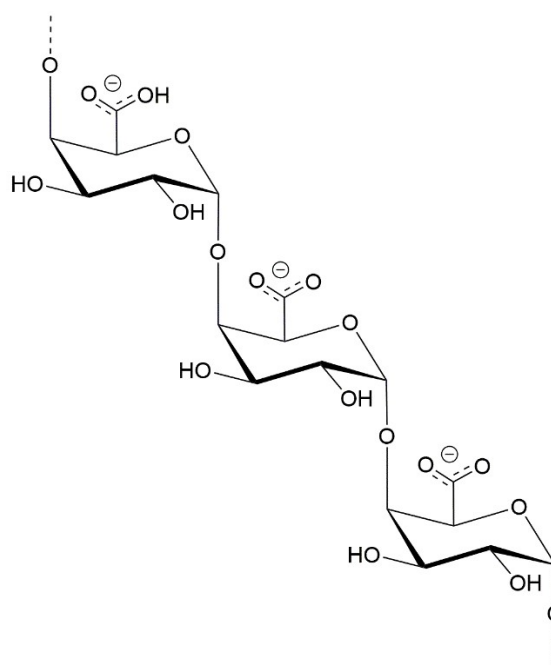
These large phenolic polymers are highly cross-linked and attach different cellulose and hemicellulose units together by covalent bonds. They are a large component of lignocellulosic waste, comprising up to 25% of the materials mass.<sup>24</sup> The aromaticity of lignin adds to the heterogeneity of lignocellulose and makes the material impermeable and resistant to microbial attack/oxidative stress.<sup>24</sup> This resistance to both chemical and biological deconstruction is the principal reason that 2G biofuel feedstocks usually require biomass which have low levels of lignin in order for the processing to be industrially/economically viable. Together, cellulose, hemicellulose and lignin make up the vast majority of lignocellulosic biomass and form the majority of complex structures in plants (**Figure 7**).<sup>25</sup>



**Figure 7. Schematic of the microstructure of plant/wood cellulose fibres showing cellulose fibrils and the associated lignin and hemicellulose (a) and a cross section and longitudinal section of a cellulose nanofibril within a lignin and hemicellulose matrix (b). Figure adapted from Seddigi *et al.*<sup>25</sup>**

#### 1.2.4 Pectin

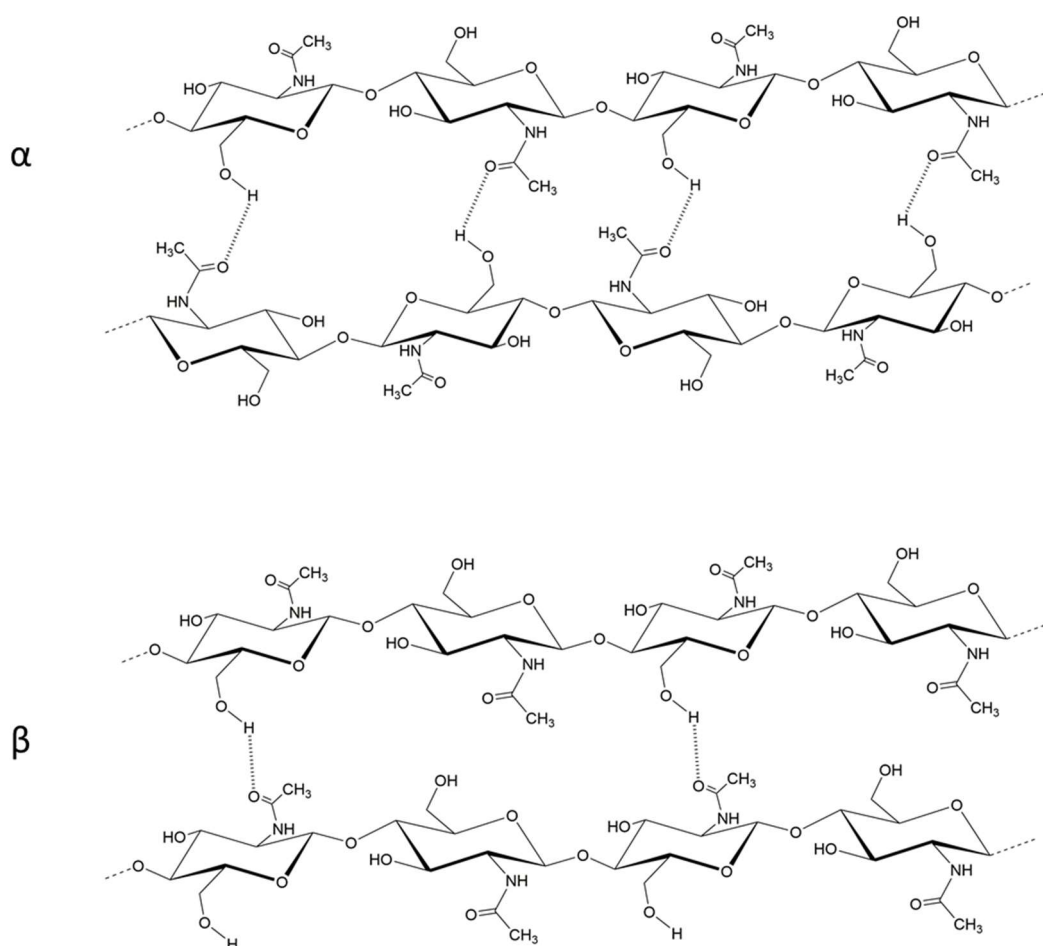
Pectin is a group of polysaccharides found again as a common constituent of plant cell walls. It is the first polysaccharide discussed so far that contains a charged group (-COOH, pKa 3.51) as part of the carbohydrate backbone.<sup>26</sup> Its major component is the homopolysaccharide,  $\alpha$  1-4 linked D-polygalacturonic acid (**Figure 8**).<sup>27</sup> Pectin has a plethora of applications spanning from a plant cell wall components to thickening agents in foodstuffs. Due to the repeating carboxylate groups, pectin is much more hydrophilic than lignin and the previously discussed polysaccharides. The hydrophilicity of the charged polysaccharide means it is better solvated in polar protic solvents (such as water) than many of the other carbohydrates. In fact, solutions incorporating pectin are known to form gels, particularly when in the presence of calcium ions. This the principal reason why pectin is utilised in industry as a thickening agent in foodstuffs including sauces, jellies and jams.<sup>27</sup> Typically, pectin is a substrate much more tractable to degradation than cellulose, as it is more easily permeated by enzymes and its breakdown can lead to the evolution CO<sub>2</sub> providing a great entropic driving force.



**Figure 8. Skeletal structure of  $\alpha$  1-4 polygalacturonic acid, the primary component of pectin.**

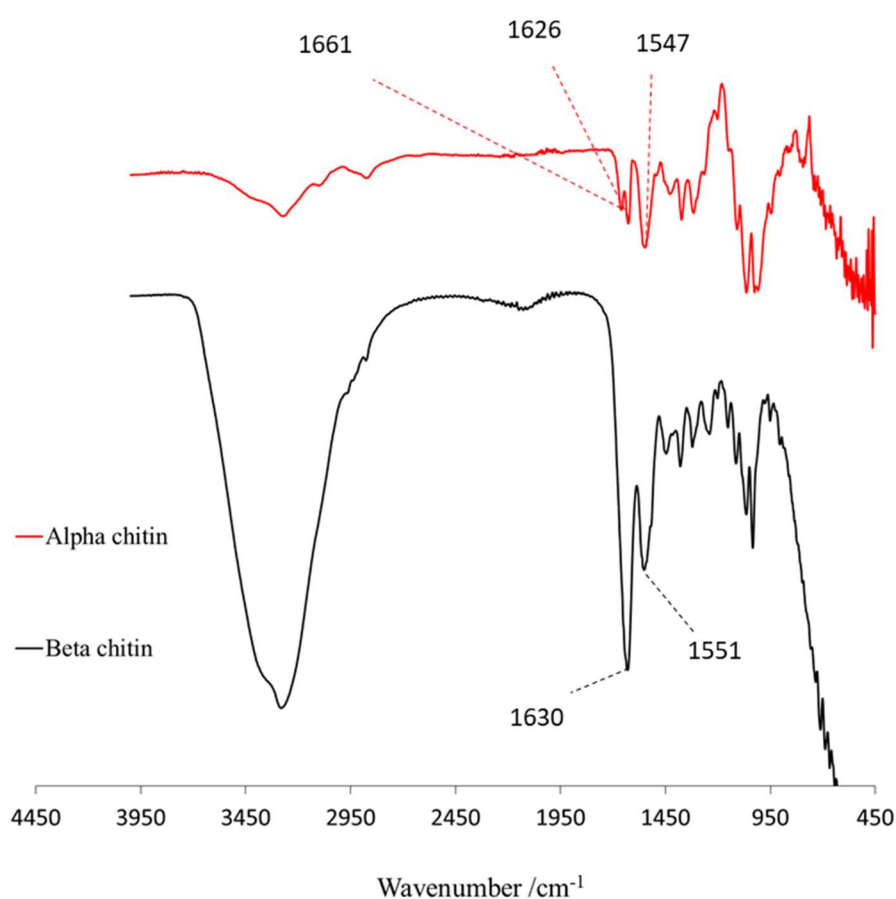
#### 1.2.5 Chitin and Chitosan

Chitin is the second-most abundant polysaccharide on earth behind cellulose stemming from its common implementation in nature.<sup>28</sup> The polysaccharide consists of repeating units of  $\beta$  1-4 linked 2-acetamido-2-deoxy-b-D-glucose. Like cellulose, it is heterogeneous and difficult to solubilise.<sup>29</sup> As such, It is found commonly as an exoskeleton material for insects and crustaceans as well as making up a large component of cell walls in fungi. Three main polymorphic forms of chitin exist, namely: alpha, beta, and gamma. The  $\alpha$  allomorph, commonly found in crab shells, possesses chitin fibrils that are aligned in an antiparallel fashion, forming hydrogen bonds between chains *via* amide and hydroxyl functional groups (top, **Figure 9**). The  $\beta$  allomorph, commonly found in squid-pens, has the chitin chains aligned in parallel, again chains are held together by favourable hydrogen networks between the polymers (bottom, **Figure 9**). The  $\gamma$  allomorph can exhibit different orientations of the polysaccharide chains commonly showing more disorder. The macro properties of the  $\gamma$  form of chitin have shown to possess much greater similarity with  $\alpha$  chitin compared to  $\beta$ .<sup>30</sup>



**Figure 9. Alpha and beta allomorphs of chitin. Strands arranged antiparallel (alpha, top) and strands arrange in parallel (beta, bottom).**

Alpha and beta chitin present different spectroscopic properties owing to their differing intermolecular hydrogen bonding networks. The particular polymorph of chitin can be determined using Fourier-transform infrared (FT-IR) spectroscopy where the amide stretching frequencies differ in energy due to H-bonds between the polysaccharide strands, as shown in **Figure 10**.<sup>31</sup>  $\alpha$  chitin is the most abundant form and is identifiable by three discrete amide stretching frequencies: approximately 1650, 1620 and 1550  $\text{cm}^{-1}$  (top in **Figure 10**). The  $\beta$  form of chitin is often more crystalline than  $\alpha$ , with common sources including the pens from both species of squid and cuttlefish. This allomorph contains just two major IR active amide stretching frequencies centred around 1630 and 1551  $\text{cm}^{-1}$  (bottom in **Figure 10**). Chitin shows many parallels to cellulose in that it is recalcitrant to degradation and requires a diverse arsenal of lytic enzymes to deconstruct. As such, this material is also a highly abundant chemical feedstock with potential for valorisation.



**Figure 10. FT-IR spectrum of shrimp-shell (alpha) chitin (red) and squid-pen (beta) chitin (bottom) highlighting key amide stretching frequencies used to distinguish the two.**

The deacetylated form of chitin, known as chitosan, is also a widespread carbohydrate in nature; employed in applications such as components crustacean exoskeletons or for bird feathers.<sup>32</sup> Chitin itself can be readily converted into its de-acetylated derivative either *via* enzymatic treatments or

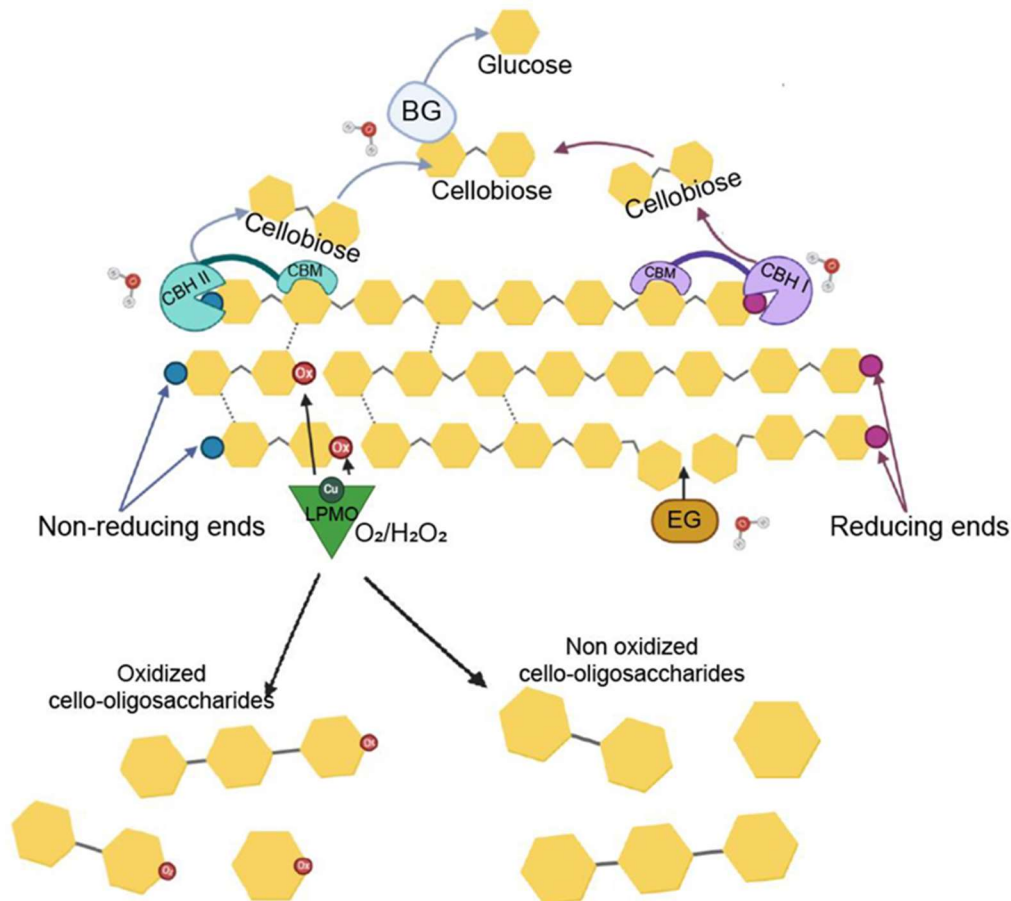
chemical treatments-typically necessitating a strong base such as sodium hydroxide.<sup>33</sup> When converted into chitosan, the original source material loses its crystallinity due to the disruption of H-bond networks previously formed with the acetyl groups. Moreover, chitosan is usually positively charged polysaccharide due to the -NH<sub>2</sub> groups making it more hydrophilic than chitin. Considering the reduced crystallinity and increased hydrophilicity, chitosan is more tractable to degradation and chemical transformation than is chemical predecessor.

Practical implementation of chitosan materials has been successful, with there being many applications in the biomedical industry. Due to the biochemical properties, biodegradability, and non-toxicity, they have been found to make great materials as grafting scaffolds and bandaging.<sup>34</sup> Clever engineering of these materials has also seen the use of chitosan as drug delivery vehicle for drugs and vaccines *via* microencapsulation.<sup>35</sup> Industrial solutions such as these, stemming from renewable waste materials, are vital in achieving a circular economy in which we are less reliant on chemicals deriving from fossil fuel products.

### 1.3 Synergistic Enzyme Cocktails to Degrade Lignocellulosic Biomass

The recalcitrance of lignocellulosic/chitinous biomass means a combination of approaches are usually required for its deconstruction. Industry utilises a range of thermochemical pre-treatments to disrupt the cell wall matrix and remove lignin from the materials. Common pre-treatment methods include: dilute sulfuric acid, liquid hot water, and steam explosion.<sup>36</sup> The pre-treatment steps help increase the surface area of the polysaccharide available for biological attack. Enzymes are necessary to splinter the crystallinity of the cellulose feedstocks as the harsh physical conditions required for the same purpose would be unsuitable for industrial viability (320 °C and 25 mPa in water).<sup>13</sup> Since the chemical composition of carbohydrates are so diverse, there are a wide range of enzymes that are active on polysaccharides to facilitate their synthesis, decomposition and modification.<sup>37</sup> Carbohydrate active enzymes (CAZymes) are collected into their respective classifications based on functionality and sequence similarity in the CAZy database ([www.cazy.org](http://www.cazy.org)).<sup>38</sup> A synergistic mix of these enzymes is referred to as an “enzyme cocktail”. These enzymes cocktails are utilised in biorefineries depolymerise the cellulose fibres into mixed sugars. This combined, synergistic approach to polysaccharide deconstruction was inspired from nature where many species of bacteria, fungi and insects use a diverse enzymatic toolbox of CAZymes to manipulate or deconstruct these materials. These enzymatic mixtures can be deployed as both extracellular and intracellular; where they have roles in breaking down recalcitrant biomass outside of the organism, or for nutrient acquisition inside

of the organism., respectively The most common of these are glycosidic hydrolases (GHs), which are usually divided further into families based on which substrate they are active on – e.g. cellulases, chitinases and amylases.<sup>39</sup> These GH enzymes all share a common function in that they break the glycosidic bond of polysaccharides *via* a hydrolytic mechanism requiring a water molecule to act as a nucleophile. As plant biomass contains vast mixtures of cellulose, pectin, lignin, etc., many different enzymes of required for its complete deconstruction. Crystalline cellulose alone, requires a plethora of enzymes to depolymerise fully into monomeric glucose units (**Figure 11**).

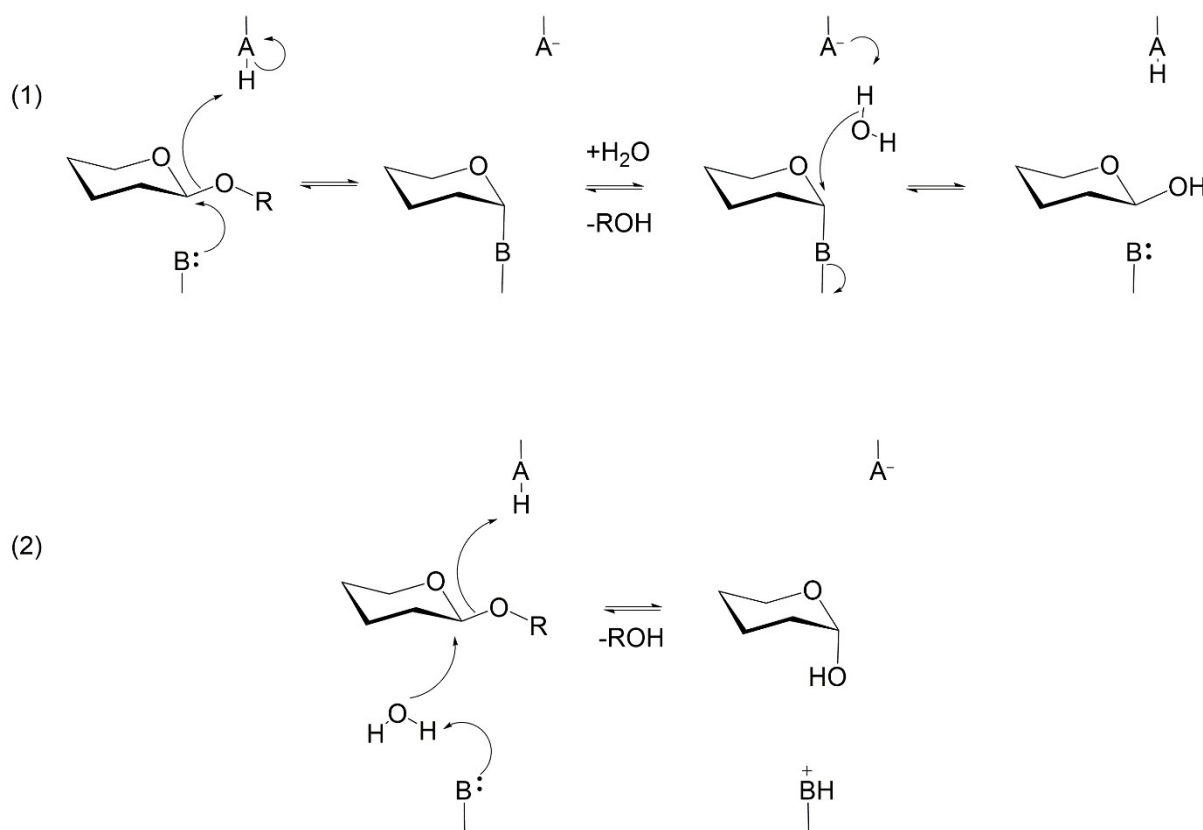


**Figure 11. Schematic of cellulose deconstruction facilitates by an enzyme cocktail. Figure adapted by Calderaro *et al.*<sup>40</sup> Endoglucanases (EG) hydrolyse the glycosidic bond of amorphous cellulose. Cellobiohydrolases hydrolyse the polysaccharide chain into cellobiose units at the reducing end (CBH I) and non-reducing end (CBH II). Lytic polysaccharide monooxygenases (LPMOs) oxidatively cleave the glycosidic bond in crystalline regions of the cellulose introducing new chain breaks and producing oxidised oligosaccharides.**

### 1.3.1 Glycosidic Hydrolases

With carbohydrates being the most abundant biomolecules, there is a requirement for an equally vast number of enzymes to handle their deconstruction. It is thought that these CAZymes constitute

between 1-2% of the genome of *any* organism on earth.<sup>37</sup> The primary constituent of these enzymatic cocktails are the glycosidic hydrolases (GHs) which perform hydrolysis of glycosidic bonds as the etymology would suggest. GHs proceed through a general acid catalysed mechanism mediated by two critical amino acid residues, a proton donor and a nucleophile/base.<sup>41</sup> The overwhelming majority of these key catalytic residues are aspartic acid and glutamic acid. Two main mechanistic pathways exist whereby the enzymatic basic group can either act as the nucleophile resulting in retained stereochemistry (retaining mechanism, **Figure 12**) or as a base to deprotonate a water molecule which subsequently performs nucleophilic substitution on the polysaccharide, inverting the stereochemistry (inverting mechanism, **Figure 12**). These enzymes are crucial for the vast majority of organisms on earth as they allow for the efficient transformation of large polysaccharides into more bioavailable and soluble sugars.



**Figure 12. Polysaccharide hydrolysis performed glycosidic hydrolases retaining and inverting mechanism (1 and 2, respectively). Where A and B are the key acidic and basic/nucleophilic residues of the glycosidic hydrolase enzyme.**

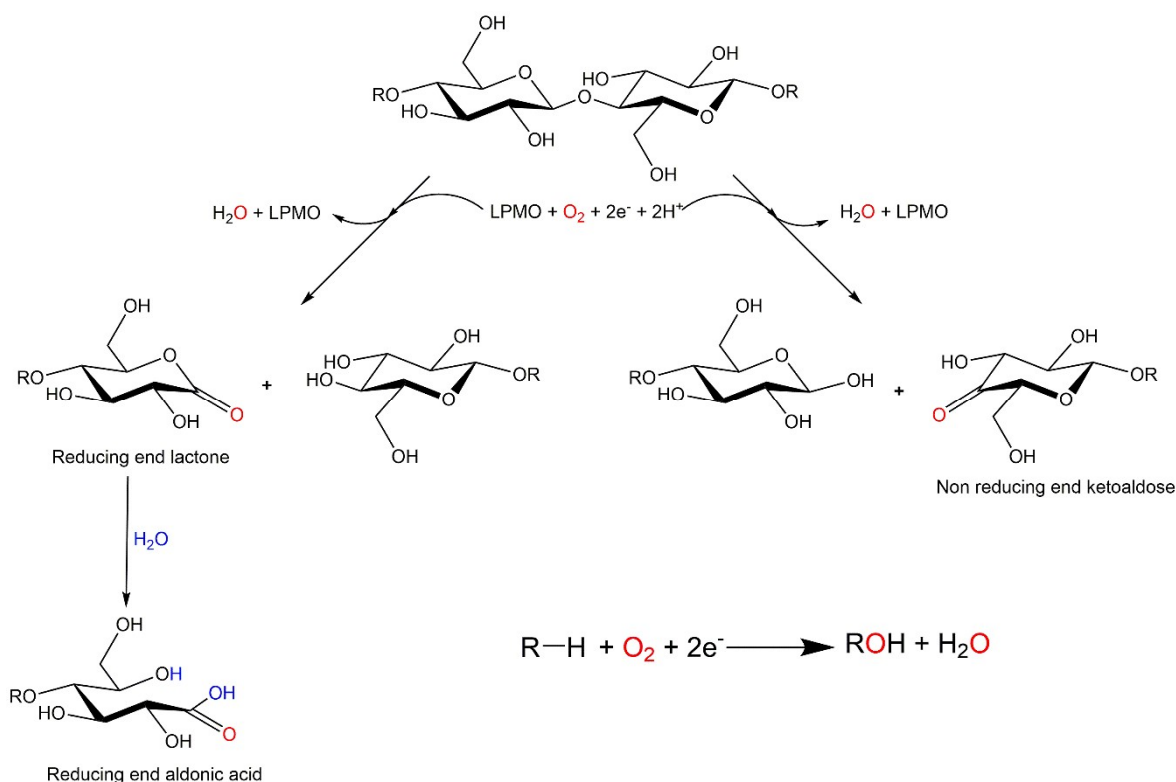
The separation between the two crucial acid/base residues appears to be the key determining factor in dictating the retaining mechanism *versus* the inverting mechanism. For the GHs that proceed through the retaining mechanism, a typical distance between the carboxylate groups is between 5-6 Å. The inverting mechanism, however, necessitates GH enzymes with a larger separation of the key

residues of around 9-10 Å.<sup>42</sup> These two mechanisms have a monopoly on carbohydrate deconstruction, but require amorphous, non-crystalline substrates in order to function. For a more in depth discussion of GH enzymes structure and function, I direct the reader to two detailed reviews by Lairson *et al.*<sup>43</sup> and Davies *et al.*<sup>41</sup>. To maximise the effectiveness of GH enzymes, the crystalline regions of carbohydrate material must be disrupted either by physical/chemical treatment or by other CAZy enzymes that show activity on recalcitrant polysaccharides.

### 1.3.1 Lytic Polysaccharide Monooxygenases

In 1950 a landmark paper was published by Reese *et al.* which outlined the activity of cellulolytic organisms on cellulose laying the platform for understanding the enzymatic toolboxes used by organisms to process cellulose.<sup>44</sup> In recent years, Lytic Polysaccharide Monooxygenases, or LPMOs, have become some of the most hotly studied and promising enzymes in relation to biomass saccharification. Their activity was first discovered in 2010 by Vaaje-Kolstad *et al.* on a bacterial enzyme from *Serratia marcescens*, SmAA10, active on chitin.<sup>45</sup> In these early studies, LPMOs were identified to *oxidatively* cleave glycosidic linkages of various polysaccharides; notably different to the known hydrolytic mechanisms of enzymes mentioned previously. Due to this key reactivity difference, LPMOs are categorised into different Auxiliary Activity (AA) families within the CAZy database ([www.CAZy.org](http://www.CAZy.org)). Since their original discovery, eight distinct LPMO families have been discovered and categorised into separate AA numbers (AA9-11, AA13-17).<sup>45-52</sup> The oxidations performed by LPMOs typically occur at either the C1 or C4 carbon of the polysaccharides. However, a recent study has shown an LPMO capable of also oxidizing at the C6 position of the polysaccharide.<sup>53</sup> These enzymes are sometimes referred to as just “polysaccharide monooxygenases” or “PMOs” as it has been suggested that the lysis step following oxidation is not under the control of the enzyme.<sup>54</sup> The oxidations performed by LPMOs introduce chain breaks in recalcitrant polysaccharide materials allowing for more efficient hydrolysis by GH cocktails. LPMOs activity on crystalline regions of carbohydrate biomass was a major finding in the context of making 2<sup>nd</sup> generation biofuel industrially viable. The oxygenation reaction of cellulose performed by LPMOs can be seen below in **Figure 13**.





**Figure 13. Reaction scheme of cellulose oxygenation performed by LPMOs and an external electron donor leading to lactones (left) and ketoaldose (right).**

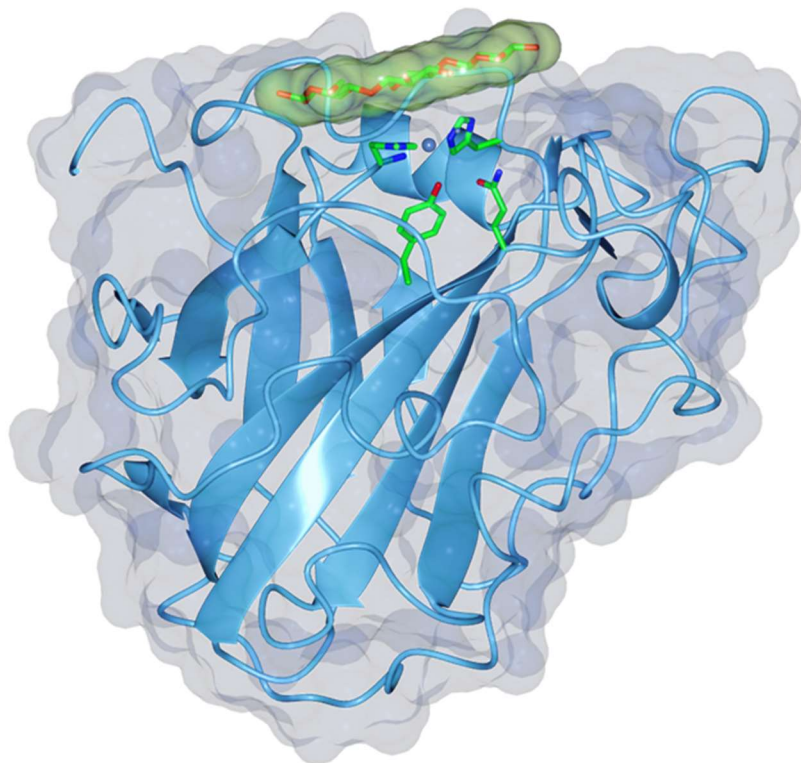
LPMOs are a class of metalloenzymes that are encoded into the genomes of a diverse array of organisms such as bacteria, fungi, and insects.<sup>50</sup> The LPMO metal cofactor was initially mischaracterised in the literature as a sodium ion.<sup>55</sup> However, a landmark study in 2011 by Quinlan *et al.* saw the correct identification of the metal to be a *copper* ion.<sup>46</sup> The following year, Beeson *et al.* reported another key finding that an oxygen atom from  $\text{O}_2$  is inserted into the oxidised products following LPMO catalysis through a clever  $^{18}\text{O}_2$  labelling experiment.<sup>56</sup> Here, they introduced  $^{18}\text{O}_2$  in an activity assay leading to a series of prominent +2 Da peaks in the product mass spectra. This confirmed LPMOs *oxygenase* activity utilising  $\text{O}_2$  as a co-substrate. They also proved the efficacy of cellobiose dehydrogenase (CDH) as an effective reducing partner for LPMO activity.

Metal cofactors are common across biology with typical examples including first row transition metals (TMs) such as iron, copper, or manganese. The main benefit of first row TMs is the vast abundancy/bioavailability of them compared to the precious metals used routinely in synthetic studies to perform similar chemistry. The incorporation of metal ions into proteins allows for much improved diversity in chemical reactivity than would be available with just the amino acid building blocks alone. Metal cofactors usually exist as cations, supported by electron donating ligands. Of the twenty

naturally occurring amino acids, many of them have side chains that facilitate metal binding with common ones including: histidine, aspartic/glutamic acid, cysteine, methionine, and tyrosine. The different metals utilised by nature are known to perform a wide range of roles from reductions and oxidations to long-range electron transfers (LRET).<sup>57, 58</sup> Some of the most commonly studied enzymes that perform difficult C-H bond activations incorporate an iron-heme moiety akin to those of Cytochrome P450 (CYP) enzymes.<sup>59</sup> However, LPMOs provide an example of a mono-copper metalloenzyme capable of performing oxidations of strong C-H bonds (*ca.* 100 kcal/mol).

LPMOs presumably proceed through a totally different mechanism as the same high valent oxygen intermediates used in CYP enzymes (iron(IV)-oxo) are not available to copper enzymes due to the “oxo wall”.<sup>60</sup> The oxo wall refers to the finding that late period transition metals (more than 4 d electrons) are incapable of forming oxo complexes. This is in keeping with ligand field theory which states that for a strong interaction between metal and oxo there can be no more than two  $\pi$ -antibonding electrons. For late row transition metals such as Cu or Ni, this would require exceptionally high and thermodynamically challenging oxidation states.<sup>61</sup> The LPMO oxidative mechanism, therefore, requires the consideration of alternative reactive intermediates and reaction pathways.

LPMOs are relatively small proteins with a typical mass of *ca.* 20 kDa, around half the average size of proteins in most organisms.<sup>62</sup> LPMOs contain some conserved features in their tertiary structure such as a core  $\beta$ -sandwich domain and adjacent  $\alpha$ -helical region which typically lies close to protein surface.<sup>63</sup> The metal binding site consists of two conserved histidine residues, one of which being the N-terminal residue (His1), coordinating a single copper ion in a T-shaped geometry known as the “histidine brace”.<sup>46</sup> This brace of histidine residues forms an intriguing coordination sphere for the copper ion as the amino (-NH<sub>2</sub>) group of the proteins N-terminus is involved as a ligand. This results in a chelating ligand which is relatively uncommon in metalloenzymes. LPMOs have been shown to have poor metal binding properties to many common metals including: Mg<sup>2+</sup>, Ca<sup>2+</sup>, Mn<sup>2+</sup>, Co<sup>2+</sup>, Ni<sup>2+</sup>, or Zn<sup>2+</sup>. However, isothermal calorimetry (ITC) experiments have shown LPMOs to bind stoichiometrically (1:1) to copper with high binding constants ( $K_D < 1$  nM).<sup>46</sup> This finding suggests that the LPMO active site has evolved to be an exceptionally good coordinator of copper and that the ligand field is arranged optimally to coordinate a single copper ion. Presumably, the chelation effect of the N-terminal histidine contributes to the strong binding energies between LPMO and copper, akin to the chelation effect seen for iron in siderophores.<sup>64</sup> An example structure of a copper containing LPMO (*Lentinus similis* AA9, LsAA9) bound to a short cellulose oligosaccharide can be seen in (**Figure 14**, PDB:5ACF).<sup>65</sup>

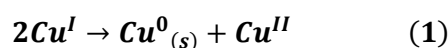


**Figure 14. Cu-LsAA9-Celotriose crystal structure from PDB:5ACF<sup>65</sup> highlighting core  $\beta$ -sandwich (blue arrows) and  $\alpha$ -helical region (blue helix). Key active site residues and copper ion are coloured according to atom type. Electron density map shown for enzyme (grey) and for celotriose unit (green).**

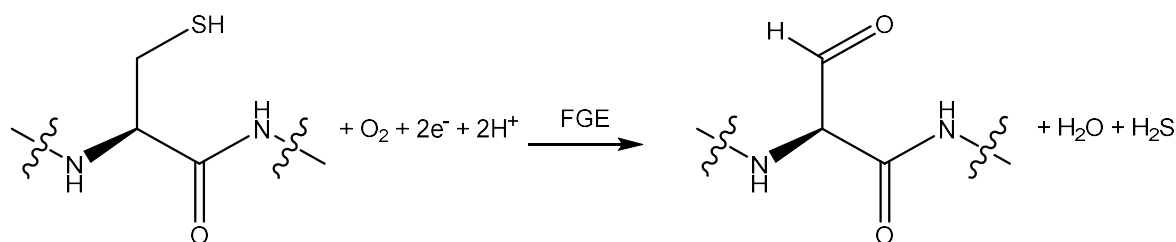
The binding of oligosaccharides has not shown to be a universal feature of LPMOs biochemistry. However, a number of examples are present in the literature of LPMOs showing the binding of soluble oligosaccharides demonstrated either by spectroscopic perturbations, crystallography, or indeed the presence of oxidised oligosaccharide products in the mass spectra of activity assays.<sup>65-67</sup> These particular enzymes are extremely valuable as they allow for analytical techniques that are not possible using their heterogenous, polysaccharide relatives. Protein crystal structures of the enzyme bound to oligosaccharide units have provided the field with valuable information pertaining the precise orientation in which the enzyme binds its substrate and which residues are key in facilitating this intermolecular interaction. These 3D structures also allow for the facile construction of computational models whereby spectroscopic properties and energetic barriers can be calculated for the LPMO reaction mechanism.

## 1.4 Copper Sites in Biology

Being one of the first-row transition metals, copper is a widespread and relatively abundant element for biology to recruit as a cofactor. Copper exists primarily in three main common oxidation states 0, +1, and +2. While metallic copper,  $\text{Cu}^0$ , provides has useful applications in manufacturing and electronics, it is not found in biological systems due to its total insolubility/heterogeneity. However, both  $\text{Cu}^I$  and  $\text{Cu}^{II}$  ions are regularly seen to occupy protein sites in biology. The  $\text{Cu}^I/\text{Cu}^{II}$  redox couple provides the platform for a range of roles in nature, stemming from  $\text{O}_2$  binding, activation, and corresponding substrate oxidations.<sup>68</sup> Though aqueous complexes of  $\text{Cu}^I$  are often found to be unstable with respect to disproportionation (Equation 1), examples of closed-shell,  $d^{10}$ , cuprous proteins exist in nature. Copper in the +1-oxidation state has a complete d-shell ( $d^{10}$ ) resulting in the absence of any d-d or LMCT features in the UV-Visible spectra. Moreover, the closed-shell  $\text{Cu}^I$  states of metalloenzymes make them inaccessible for study *via* useful techniques including electron paramagnetic resonance (EPR) spectroscopy, primarily limiting the study of their electronic structure to X-ray techniques.

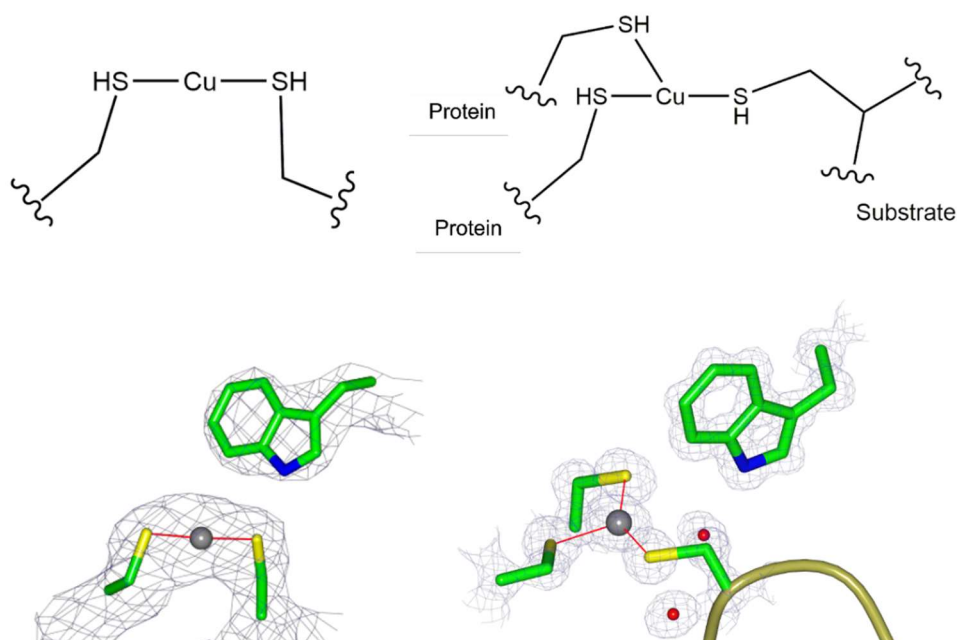


Traditionally, +1 is an unstable oxidation state for most copper complexes wrt. disproportionation to the +2 and 0 oxidation states (equation 1). However, enzymes exist which have shown to support a stable single copper ion in the +1 oxidation state. Formylglycine generating enzyme (FGE) is a newly discovered class of enzymes that incorporate a single  $\text{Cu}^I$  ion in its resting state.<sup>69</sup> The copper site in FGE catalyses the oxidation of cysteine residues to  $\alpha$ -formylglycine. Both crystallographic and EPR studies have confirmed the  $\text{Cu}^I$  oxidation state of the cofactor, which is unusual for that of most known oxidases.<sup>69</sup> Like the majority of copper oxidases, however, both protons and electrons are required from external donors to facilitate catalysis. A FGE reaction scheme can be seen below in **Figure 15**. FGEs ability to convert a natural amino acid into a residue containing an aldehyde group has led to its involvement in bioconjugation studies born out of the ease and specificity of functionalising the resulting aldehyde.<sup>70</sup>



**Figure 15. Reaction scheme of the transformation of a cysteine residue to  $\alpha$ -formylglycine facilitated by formylglycine generating enzyme.**

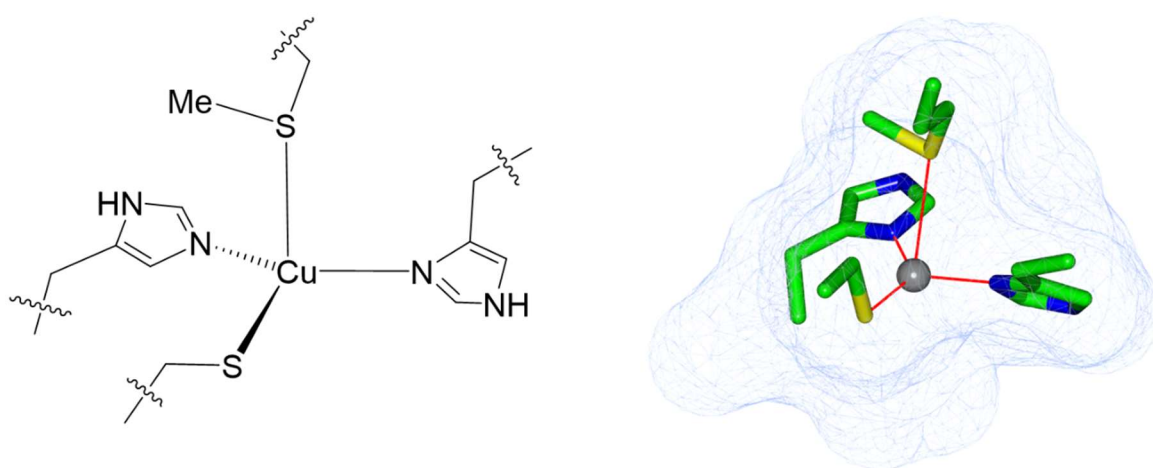
Structurally, FGEs contain a copper binding site that more closely resembles copper chaperone proteins than most copper oxidases. Its active site consists of two conserved cysteine residues that bind copper in a 2-coordinate, linear geometry (left, **Figure 16**).  $\text{Cu}^I$  is 'softer' than  $\text{Cu}^{II}$  attributed to its lower effective nuclear charge,  $Z_{eff}$ , and comparable electronic shielding. The valence orbitals in  $\text{Cu}^I$  are more diffuse and tend to form more favourable interactions with soft ligands. The sulfur atoms of the cysteine side chains coordinate the  $\text{Cu}^I$  favourably as they are softer ligands compared to the common N/O ligand analogues. As a result, a copper ion in this ligand field is well stabilised in the reduced,  $\text{Cu}^I$ , form of the enzyme. So much so in fact, that when the apo-enzyme is metalated with a  $\text{Cu}^{II}$  salt it auto-reduces to the  $\text{Cu}^I$  state evidenced by a time dependent quenching of the Cu EPR signal.<sup>69</sup> To stabilise the oxidised,  $\text{Cu}^{II}$ , state of the enzyme, recruitment of additional electron donating ligands is necessary. In this case, the binding of a cysteine sidechain from a suitable substrate (right, **Figure 16**) adds a crucial additional ligand now allowing for oxygen to bind and subsequently activate. FGEs are a fascinating example of a substrate-coupled  $\text{O}_2$  activation pathway whereby FGE can only produce reactive oxygen species once a suitable substrate has bound the enzyme. This helps to reduce any off pathway, deleterious chemistry from happening and potentially inactivating the enzyme.



**Figure 16. The skeletal (top) and crystal (bottom) structure of the active site of formylglycine generating enzyme from *Streptomyces coelicolor* and *Thermomonospora curvata* (PDB 6MUJ and 6S07).<sup>69, 71</sup> Figure adapted from Paradisi *et al.*<sup>72</sup>**

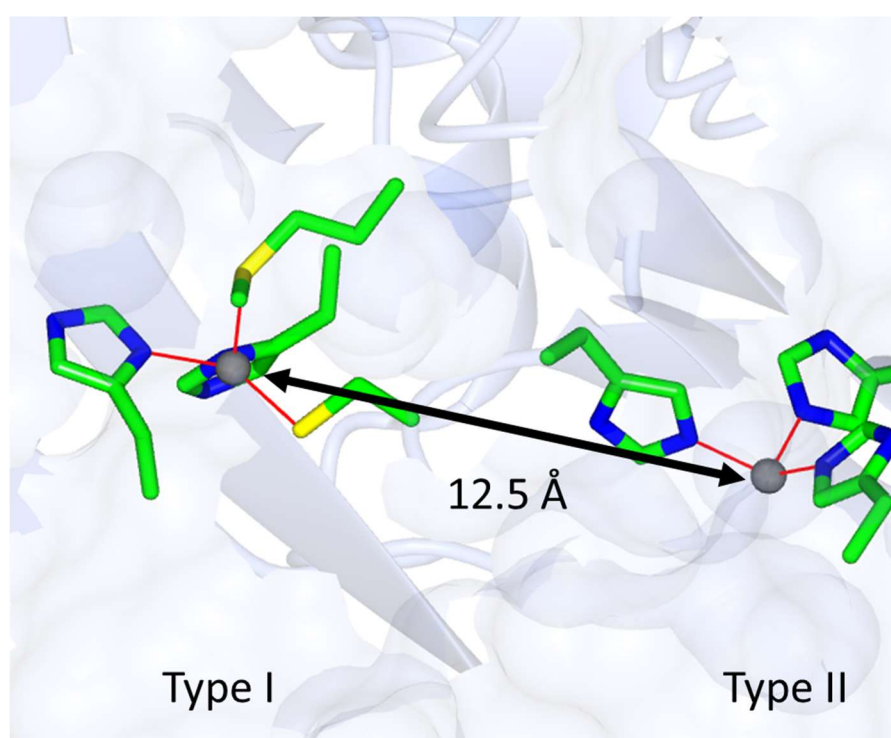
### 1.4.1 Type I 'Blue' Copper Proteins

The ubiquitous type I or 'blue' copper proteins contain a mono copper site in which a  $\text{Cu}^{\text{II}}$  ion coordinated by two histidine residues, a thiolate sulfur (from a cysteine), and thioether sulfur (from a methionine) in a distorted tetrahedral geometry.<sup>73</sup> They are commonly referred to as 'blue copper' sites since solutions of these proteins appear a striking blue colour. This arises from an intense ligand to metal charge transfer (LMCT) transition at around 600 nm occurring from the sulfur ligands to the vacant d orbital on the metal. Due to the two sulfur ligands, type I copper sites show a high degree of covalency which are reflected in their spectroscopic characteristics (small Cu hyperfine coupling values  $A_z < 200$  MHz).<sup>68</sup> The recruitment of histidines sidechains as well as the of softer, sulfur-containing ligands results in a copper site that is able to similarly stabilise both the +2 and +1 oxidation states. The similar energy of the oxidised and reduced form of type I copper sites allows for the facile transfer of an electron making them effective in electron transfer roles. Since type I copper sites take on a distorted tetrahedral geometry, they are not as susceptible to Jahn-Teller (JT) distortion that is usual for most  $\text{Cu}^{\text{II}}$  tetragonal complexes. As a result, the geometry around the copper ion does not significantly change upon one electron reduction to  $\text{Cu}^{\text{I}}$ . The lack of geometry change upon reduction is believed to be a major contributor in the ease of which type I copper centres can shuttle between the two oxidation states. This is a phenomenon referred to as the *entatic state*.<sup>74, 75</sup>



**Figure 17. Skeletal structure of a type I copper site (left) and crystallographic coordinates of a type I copper site in the azurin protein from *Pseudomonas aeruginosa* (PDB:5AZU).<sup>76</sup>**

Typical type I copper sites contain a relatively short Cu-S bond to the cysteine thiolate group of *ca.* 2.1 Å and an elongated Cu-S bond to the methionine thioether of *ca.* 2.9 Å. The distortion of the ligands away from tetrahedral results in the three shorter ligands forming a pseudo-plane incorporating the central copper ion. The resulting electronic structure of which has been determined to be a  $d(x^2-y^2)$  ground state defined by these strong in-plane ligands.<sup>68</sup> The proteins themselves have not been found to perform any particular catalytic transformations but remain widespread in nature as electron transfer partners. The most common of the type I copper proteins are plastocyanins and azurins which perform important electron transfers in photosynthesis and denitrification, respectively.<sup>77, 78</sup> Due to their role as electron transfer proteins, they are often small ( $\leq 10$  kDa) to maximise their mobility. Type I copper sites can, however, be found as part of more complex enzymes whereby they serve to deliver or remove electrons from another catalytic site.<sup>79</sup> An example of this can be seen in a nitrite reductase enzyme from *Hyphomicrobium denitrificans* (PDB:2DV6) where both type I and II copper sites are contained within the same enzyme separated by 12.5 Å (**Figure 18**).<sup>80</sup>

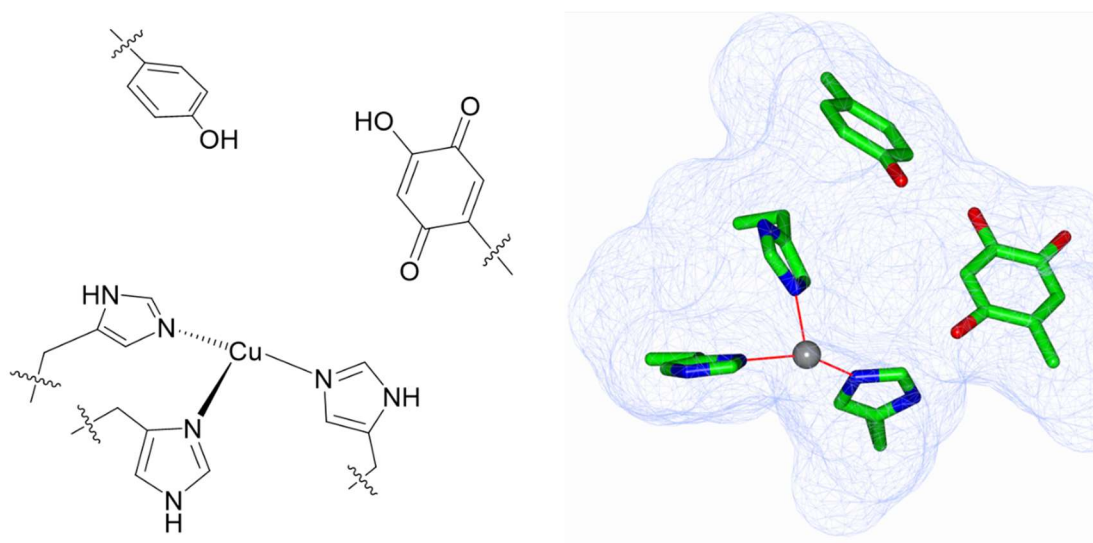


**Figure 18.** Nitrite reductase enzyme from *Hyphomicrobium denitrificans* (PDB:2DV6) highlighting both type I and type II copper sites.

### 1.4.2 Type II Copper Proteins

Type II copper proteins again incorporate a single copper ion and are found to perform a variety of applications in biology. They have a coordination geometry consisting of just N/O ligands

whose arrangement commonly resembles square-planar. In the absence of sulfur-containing ligands in the coordination sphere, the LMCT band at 600 nm that is characteristic in the UV-visible spectrum of type I copper proteins, is not present in type II. As a result, these enzymes do not display the same aforementioned blue colour. Typical type II copper proteins share a copper ion coordinated by 3 histidine sidechains and exogenous ligand(s) (e.g. water or chloride) completing a typical square planar coordination sphere. Despite their widespread similarities in geometric structure, the chemistry performed by type II copper sites has shown to be exceptionally diverse.<sup>81</sup> Type II copper sites are ubiquitous in nature and primarily have roles in oxidative biochemistry.<sup>68</sup> This type of copper site is present in many oxidase and oxygenase enzymes. The former of these enzymatic classes performs the simple oxidation of substrates without the incorporation of oxygen atoms, whereas the latter oxidises substrates *via* the incorporation of oxygen atom(s).<sup>72</sup> In addition to the three histidine residues, some type II copper sites have been seen incorporate post translationally modified (PTM) amino acids to execute nuanced and specific functions in the protein.<sup>82</sup> An example of this can be seen in copper-dependent amine oxidase enzymes where a modified tyrosine-derived quinone (**Figure 19**) participates in the catalytic conversion of primary amines into aldehydes, evolving both  $\text{NH}_3$  and  $\text{H}_2\text{O}_2$  in the process.



**Figure 19. Active site of a copper amine oxidase from *Arthrobacter globiformis* (PDB:1IU7) skeletal structure (left) and crystallographic coordinates (right).<sup>83</sup>**

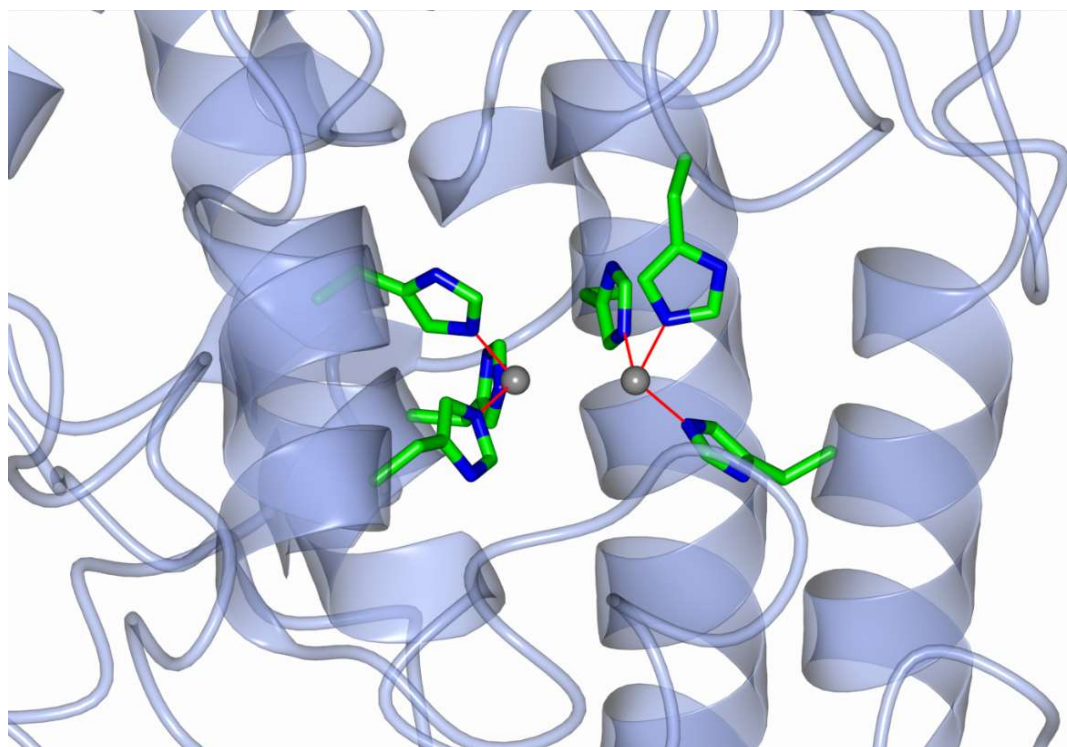
The valence orbitals of the histidine ligands are less diffuse than the sulfur containing analogues. This results in the ligands having poorer orbital overlap with the copper and therefore, less able to participate in covalent bonding. Moreover, the absence of softer, S-containing ligands results in the  $\text{Cu}^{\text{I}}$  oxidation state being less stabilised than in type I copper proteins. As a result, the type II copper sites in their resting state have a strong preference for the oxidised,  $\text{Cu}^{\text{II}}$ , form and require the



addition of external reducing agents and air-free techniques to support the reduced,  $\text{Cu}^{\text{I}}$ , form. The mechanism by which type II copper sites perform oxidations/oxygenations appears to show variation between the enzymatic classes with there still being significant questions remaining. The reactive intermediates expected to be employed by these enzymes are yet to be characterised. Much debate surrounds the identity of these intermediates, particularly in since they are prohibited from forming oxo complexes governed by the oxo wall. Oxygenase enzymes belong to the oxidoreductase family (Enzyme Classification 1, (EC 1.X.3)). These enzymes harness the oxidising power of molecular  $\text{O}_2$  to accept both electrons and protons to create reactive intermediates and leading to one or both oxygen atoms incorporated into the products. Whether one or both oxygen atoms appear in the oxidised substrate has led to the further sub-classifications of these enzymes as either *monooxygenase* or *dioxygenase*, respectively. Despite its oxidative potential, the reaction with  $\text{O}_2$  presents considerable kinetic challenges. Since  $\text{O}_2$  presents a triplet electronic ground state ( $S=1$ ), its reaction with singlet reactants ( $S=0$ ) to form singlet products is forbidden by the spin selection rule ( $\Delta S=0$ ).<sup>72, 84</sup> This difficulty in overcoming the spin state cross over (intersystem crossing, ISC) is essential in nature so that uncontrolled oxidations do not occur spontaneously. In the case of oxygenase enzymes, the copper cofactor facilitates this ISC process due to introducing a significant degree of spin-orbit coupling allowing for the facile activation of  $\text{O}_2$  and subsequent oxygenation.

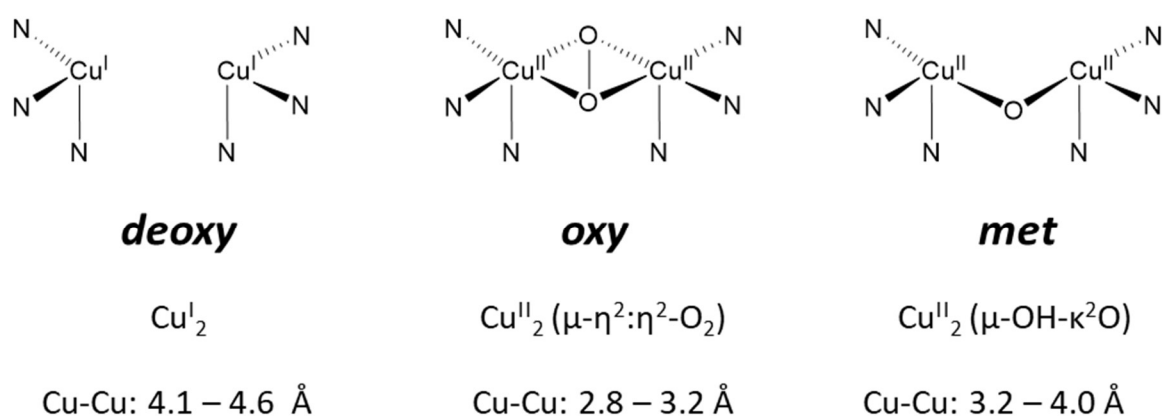
### 1.4.3 Type III Copper Proteins

Type III, or binuclear copper, sites contain two magnetically coupled  $\text{Cu}^{\text{II}}$  ions each coordinated by three histidine side chains. The two copper ions are held in close proximity to each other with a typical distance of less than 6 Å. These copper sites share some similarities to type II in that the copper ions are each ligated by three histidine sidechains. They again are typically found to perform oxidations of strong bonds across a variety of substrates. However, since they are binuclear, they undergo a separate mechanism utilising different reactive intermediates compared to the type II oxidase/oxygenase analogues. The proximity of the two copper ions allows for bridging ligands to coordinate both metal sites. By ligating two positively charged metal cations, the resulting reactive oxygen intermediates are thought to be even more electron deficient and thus more potent oxidisers. An example class of enzyme containing a type III copper site are tyrosinases which catalyse the oxygenation reaction of certain tyrosine residues to their corresponding ortho-quinones. Many of the solved crystal structures of tyrosinase enzymes show the complexation of chaperone proteins which are often coded for immediately upstream of the tyrosinase gene. These chaperones proteins have been shown to supply these oxygenase enzymes with copper so that they are able to perform catalysis.



**Figure 20. Active site of a tyrosinase enzyme from *Bacillus megaterium* (PDB:3NM8) highlighting the binuclear, type III copper site.<sup>85</sup>**

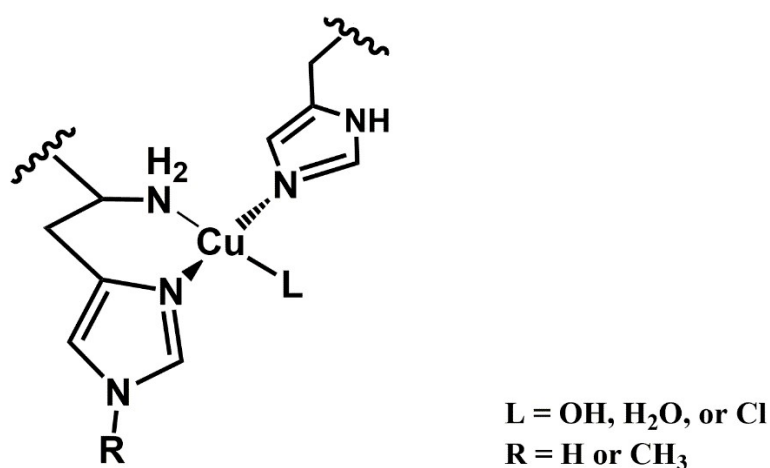
The two copper ions in type III copper sites can again exist in the +1 and +2 oxidation states. The active site of these enzymes adopts three common states during catalysis, the *deoxy*, *oxy*, and *met* state (**Figure 21**).<sup>86</sup> The *deoxy* state contains two Cu<sup>I</sup> ions and no exogeneous ligands. This oxidation state of the type III active sites sees the longest Cu-Cu distances (4.1 - 4.6 Å). This reduced state of the active site is primed for oxygen activation. The *deoxy* state can bind molecular oxygen across both metal centres in a  $\mu\text{-}\eta^2\text{:}\eta^2$  coordination mode. The binding of oxygen is accompanied with simultaneous activation as one-electron reductions occur from each copper ion yielding a dicopper(II) peroxy species.<sup>87</sup> This oxygen bound, peroxy, species is known as the *oxy* state of the enzyme. This state is an activated form of the enzyme within the catalytic cycles of type III copper sites.<sup>88</sup> Now the two copper ions share a bridging ligand, the displacement between the copper ions is much reduced wrt. the *deoxy* form with typical Cu-Cu distances between 2.8 and 3.2 Å. The *oxy* form can go on to oxidise substrates such as the conversion of phenols to catechols. The final form adopted by type III copper sites is the *met* state. In the *met* state, a single oxygen atom bridges the two Cu<sup>II</sup> ions, usually in the form of a hydroxide ligand.<sup>88</sup> This is the form that native type III copper sites will uptake the majority of the time. This state is considered to be the 'resting state' of the enzyme. The Cu-Cu displacement of this state is in-between that of both *oxy* and *deoxy* with typical distances between 3.2 and 4.0 Å. The *met* state of tyrosinase enzymes is unable to oxidise phenols but has shown to still catalyse the oxidation of catechols to their corresponding ortho-quinones. In the process, the *deoxy* state of the enzyme is reached where the copper site is primed again for reactivity with O<sub>2</sub>. As a result, tyrosinase enzymes in the presence of oxygen, will oxidise both phenols and catechols under separate catalytic cycles.



**Figure 21.** Different oxidation states type III copper sites may adopt during catalysis with corresponding typical Cu-Cu bond distances.

## 1.5 The 'Histidine Brace' – An Alternative Copper Site

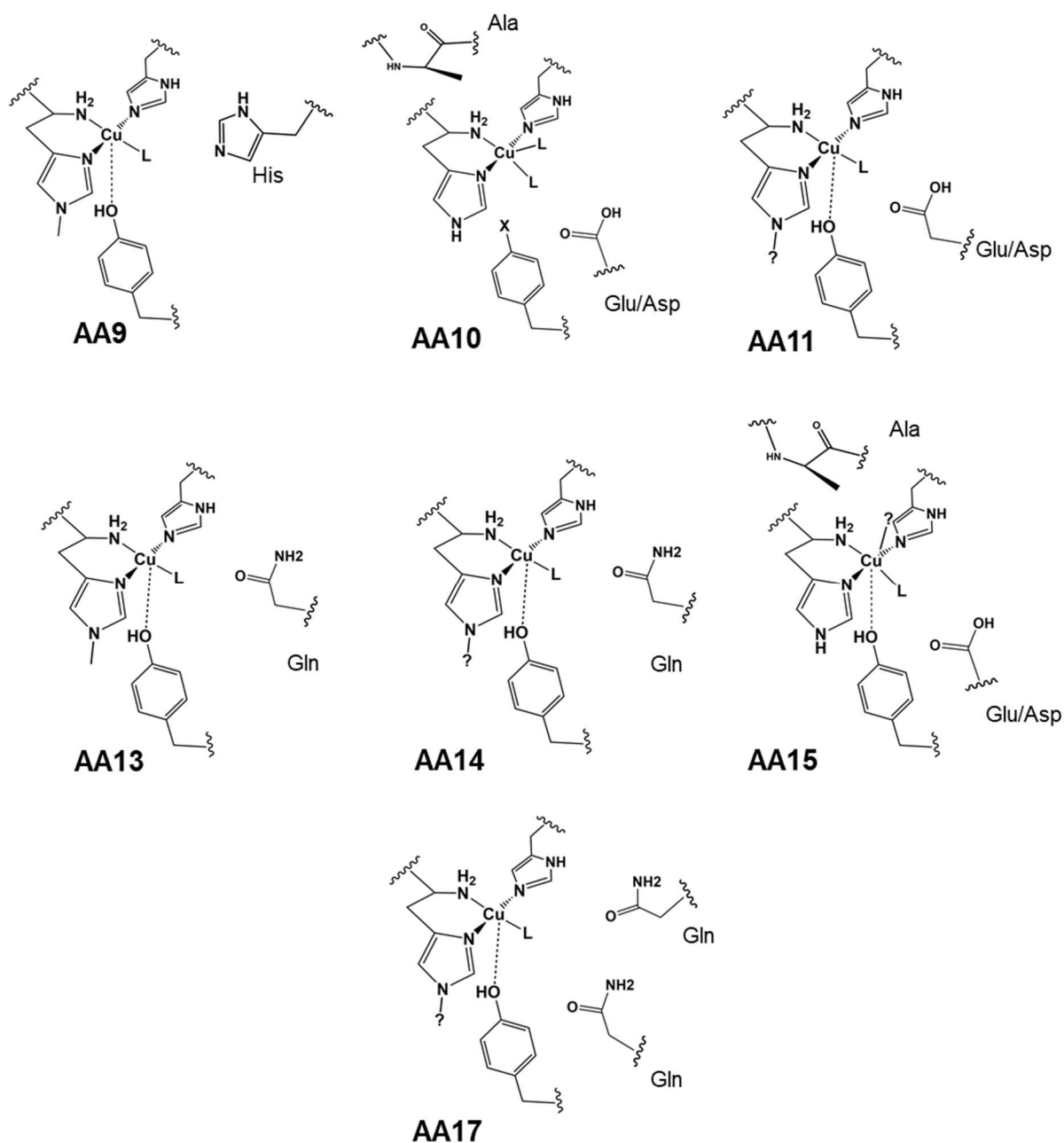
LPMOs have been shown to contain a single copper ion at the heart of their active site, which is coordinated by just two conserved histidine residues, completing the endogenous ligand field. One of these conserved histidine residues is the N-terminal residue which possesses a free amino group,  $-NH_2$ , uninvolved in the formation of a peptide bond. This allows the  $-NH_2$  group to serve as a ligand to the metal creating a bidentate chelating ligand from His1. The di-histidine moiety employed at the active site of LPMOs was coined as the 'histidine brace' in 2011 by Quinlan *et al.*<sup>46</sup> The *histidine brace* is a fitting name for the LPMO ligand field as it is comprised of a brace, or pair, of histidine residues that tightly coordinate to copper. As a mononuclear copper site tetragonally coordinated by histidine residues, they fall most closely into the type II classification for a copper site, though not without with some key differences. Unlike the type II copper sites, they do not possess a third histidine as part of their ligand field, with the amino terminus making up the otherwise vacant coordination site. The resulting T-shaped geometry of histidine brace is highly conserved and shows little variation between the LPMO families. The N-terminal histidine (His1) coordinates with the imidazole ring aligned with the N-Cu-N plane. Interestingly, the second histidine (His2) sidechain that coordinates the metal exhibits a twist with an angle of *ca.* 70 degrees between the best fit planes of the imidazole rings. The primary coordination sphere of an LPMO is depicted in **Figure 22**.



**Figure 22.** Histidine brace coordination geometry seen among LPMOs. Ligands, L, can be either water, hydroxide, or chloride. The substituent on the remote nitrogen of His1 can either be -H or  $CH_3$  depending on whether the organism expressing the enzyme has the enzymatic tools to perform this post-translational modification.

The  $Cu^I$  state of the enzyme sees just the 3N ligation of the histidine brace with a T-shaped geometry. In the  $Cu^{II}$  resting state, LPMOs recruit additional exogenous ligands to complete their primary coordination sphere. These are usually  $H_2O$  or  $OH^-$  ligands but there are examples where

chloride ions can be recruited. A study on an LPMO enzyme from *Lentinus similis*, LsAA9, highlighted that the binding of a cellulose oligosaccharide substrate to the LPMO saw the replacement of the water exogenous ligand with a chloride ion.<sup>65</sup> The final coordination geometry adopted by the LPMO shows variation among the families. This is usually mediated by conserved secondary coordination sphere residues that influence the final geometry by hydrogen bonding networks or steric crowding. The active site structures typical of the distinct families can be seen in **Figure 23** (AA16 omitted as no structures are available to date).



**Figure 23.** Typical active site structures for the different LPMO families. Exogeneous ligands are denoted by 'L' which can be H<sub>2</sub>O, OH<sup>-</sup>, or Cl<sup>-</sup>. For certain LPMO families, the substituent on the

**remote nitrogen of His1 is unknown since it is unclear whether the host organism has the capability of performing the PTM methylation denoted by '?'.**

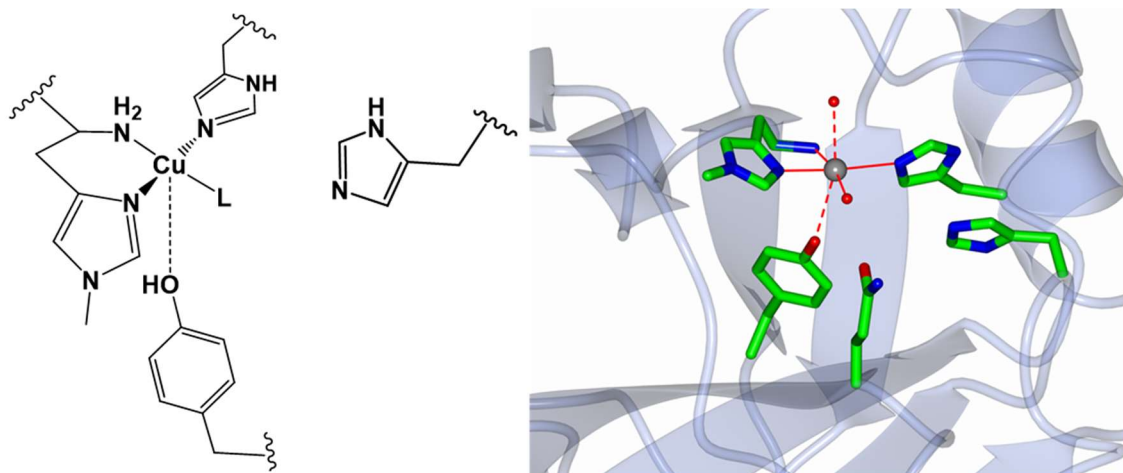
For most of the LPMO families (all but chitin-active AA10s), a tyrosine is located beneath the copper equatorial coordination plane and is weakly coordinated in the Jahn-Teller distorted axial position. For chitin-active AA10s, this residue is replaced by a phenylalanine. The phenylalanine residue provides similar steric bulk but misses the phenol functional group available for electron donation. Variation also exists between the families over the substituent on the remote nitrogen of His1. LPMOs originating from fungal organisms usually exhibit a post translational modification in the means of methylating the remote nitrogen on His1. The purpose of this methylation is still under debate, though a study from Petrović *et al.* indicated that the methylation of an LPMO from *Thermoascus aurantiacus*, TaAA9, led to greater resistance to oxidative damage when reacted with substrate and H<sub>2</sub>O<sub>2</sub>.<sup>89</sup> For the AA11, AA14, and AA17 families, it remains unclear whether the host organisms perform this PTM (as such, the substituent is denoted with '?' in **Figure 23**).

In addition to the histidine brace, certain secondary coordination sphere residues are also conserved as part of the active site structure of LPMOs depending on the family. The AA9 class for example contains a third conserved histidine residue which is located nearby (*ca.* 5 Å) to the copper but does not partake as a ligand. It is thought that residues such as these may participate in important hydrogen bonding roles in applications such as directing and stabilising intermediates of the catalytic cycle.<sup>90</sup> In the case of LPMOs from the AA10 class, a conserved acidic residue, either a glutamate or aspartate, is situated in close proximity to the central copper ion and strongly hydrogen bonds to one of the two water ligands present in its coordination sphere.<sup>63, 91</sup> This along with an sterically bulky alanine residue lead to a different geometry around the central copper ion in chitin-active AA10s when compared to the other classes.

### 1.5.1 The AA9s

The AA9 class of LPMOs are some of the most studied LPMOs in the literature so far.<sup>46, 54, 66</sup> They were first structurally characterised in 2011 by Quinlan *et al.* where an enzyme from the soil fungus *Thermoascus aurantiacus* was isolated and showed to perform oxidative degradation of cellulose in the presence of oxygen and a suitable reducing agent.<sup>46</sup> This class of LPMOs showed some significant differences to the AA10 class that was first discovered a year earlier. One of the primary differences from this enzyme from the previously discovered AA10 enzymes was the demonstration of activity on cellulose rather than chitin. The oxidative activity on cellulose drew in considerable interest due to their possible role in exploiting biomass deconstruction for biofuel production. Further studies have demonstrated that some members of the AA9 family show activity also on hemicelluloses such as xyloglucan.<sup>92</sup>

Quinlan and co-workers highlighted for the first time not only the correct metal cofactor, but also the unique methylation of the  $\tau$  nitrogen on His1. The AA9 LPMOs primarily stem from fungal organisms which possess the necessary enzymatic toolbox to perform such post translational modifications. Both methylated and non-methylated variants of these enzymes have shown activity on turning over oxidised oligosaccharides raising questions over the reason for these PTMs. A computational study by Kim *et al.* found that the PTM had no significant effect in reducing the theoretical energetic barrier height (*ca.* 1 kcal/mol) during catalysis.<sup>93</sup> To produce the AA9 enzymes containing the methylation on His1, the practitioner must use a suitable host organism to express the enzyme. Despite being the most common organism used to express proteins, *E.coli* are not capable of performing this PTM. A study in 2017 by Simmons *et al.* saw the production of CvAA9 in *Aspergillus oryzae* where the expression in the fungal host retained the methylation of the histidine nitrogen. However, the same enzyme expressed in *Pichia pastoris* and in bacteria saw the absence of this PTM.<sup>66</sup> An example of the AA9 active site structure can be seen in **Figure 24**.



**Figure 24. Typical active site structure of AA9 enzymes (left) and active site crystal structure of the first characterised AA9 LPMO, T $\alpha$ AA9 (PDB:2YET).<sup>46</sup>**

The AA9 coordination sphere is best described as axial, whereby the ligands are all oriented in a Jahn-Teller distorted geometry, typical for d<sup>9</sup> copper complexes. The histidine brace plane sees shorter metal ligand bond lengths (average 2-2.1 Å) than in the axial positions (*ca.* 2.6-2.9 Å). A conserved tyrosine is positioned in the nominal axial position beneath the plane with a typical Cu-O distance of around 2.8 Å. A water molecule is commonly seen in crystal structures to occupy the remaining axial coordination site. A third histidine and asparagine are the next nearest neighbours to the copper centre and are well positioned to hydrogen bond the equatorial ligand or intermediates generated as part of the LPMO catalytic cycle.

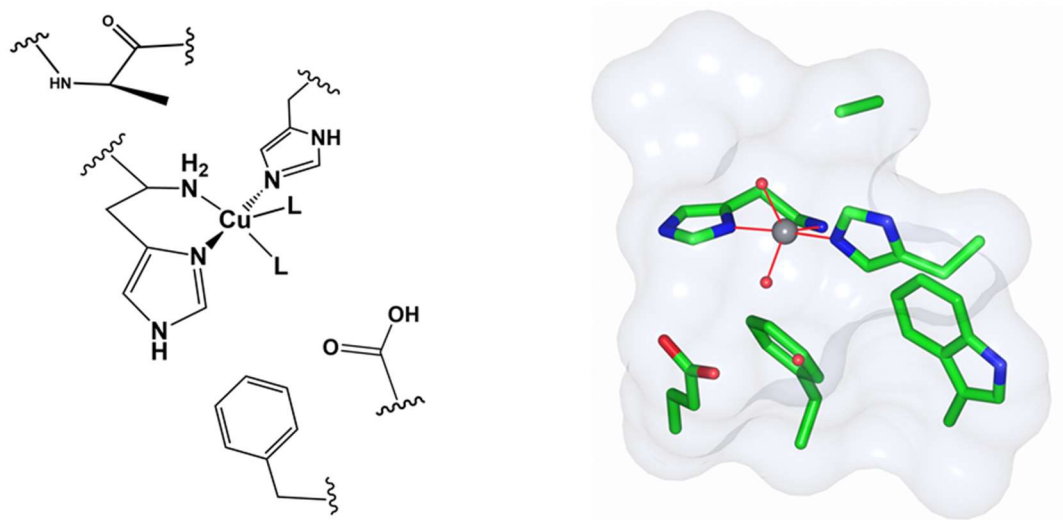


## 1.5.2 The AA10s

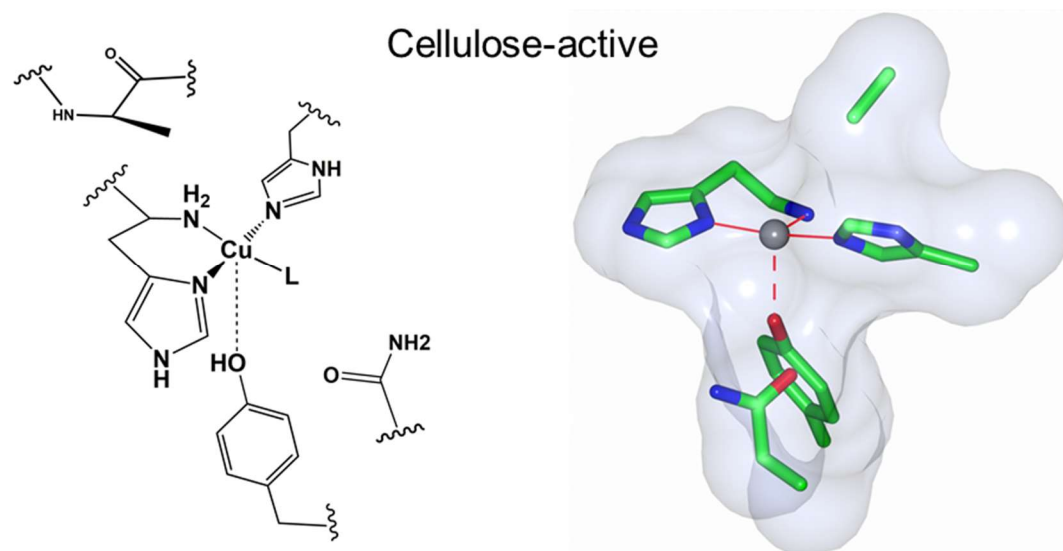
The first characterised LPMO, from *Serratia marcescens*, now belongs to the AA10 family.<sup>45</sup> These enzymes were formerly classified as chitin-binding modules (CBMs) and specifically belonged to the CBM33 family. Their first discovery in 2010 was a breakthrough in the field and sparked further studies. The crystal structure of *SmAA10* (PDB:2BM) was reported in this important paper, highlighting the di-histidine moiety near the proteins surface.<sup>45</sup> Their crystal structure, PDB: 2BEM, was modelled with a sodium ion occupying this metal binding site. In this initial paper, it was shown that the reactivity was thwarted on addition of EDTA. On this basis, it was postulated that the activity was instead dependent on a divalent cation occupying this site, such as  $Mg^{2+}$  or  $Zn^{2+}$ . The identity of the metal cofactor was correctly identified to be a copper ion a year later in a second landmark paper in two years.<sup>46</sup>

The AA10 family are primarily found in the genomes of bacterial organisms. Oftentimes, these are soil bacteria that possess a plethora of enzymes that enable them to degrade detritus. The majority of AA10s have shown activity on chitin, the major structural polysaccharide that makes up the exoskeletons of many insects and the shells of many crustaceans. However, some AA10s have shown activity on cellulose, akin to AA9s. A study by Forsberg *et al.* in 2014 compared the activity and EPR spectra of four separate AA10 LPMOs. They found that the cellulose active variants contained an active site tyrosine residue, in keeping with the AA9 family. The chitin active AA10s, however, contain a phenylalanine residue in the position. The resulting EPR signatures of these enzymes reflect this difference in coordination geometry.<sup>94</sup> The cellulose-active AA10s possess an EPR signature much more in keeping with typical tetragonal copper complexes. In comparison, the chitin active AA10s demonstrate considerable rhombicity indicating a totally different coordination geometry. The typical active site structure of a chitin active and cellulose active AA10s can be seen below in **Figure 25**.

## Chitin-active



## Cellulose-active



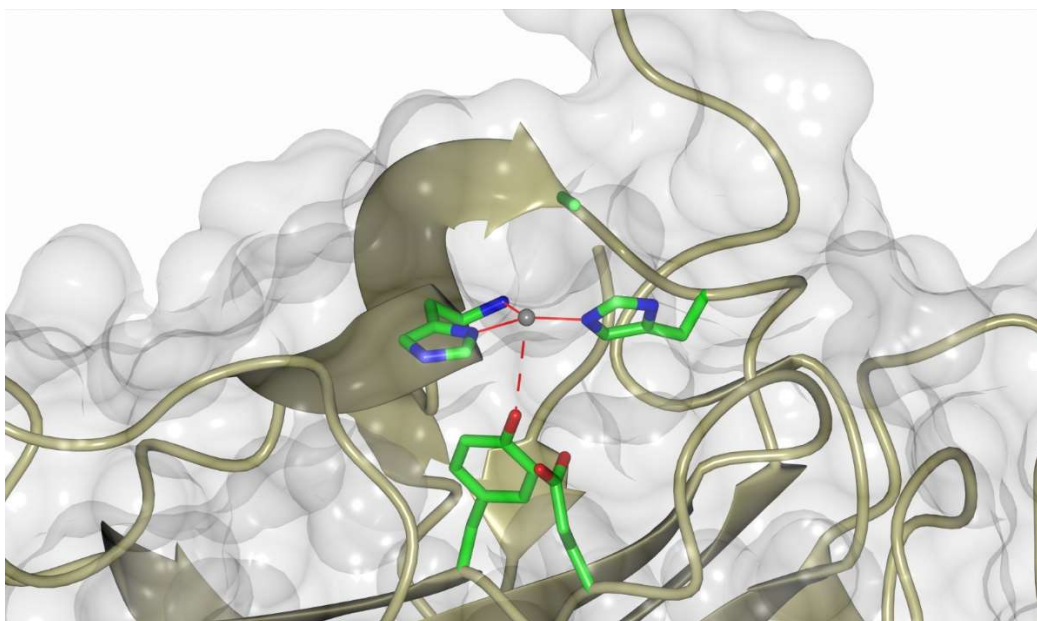
**Figure 25. Typical active site geometry of chitin-active AA10s and crystal structure coordinates of *BaAA10* enzyme (PDB:5IJU) (top), and typical active site geometry of cellulose-active AA10s and crystal structure coordinates from *MaAA10* (PDB:5OPF) (bottom).**<sup>63, 94</sup>

The coordination structure around the metal ion in chitin-active AA10s resembles trigonal bipyramidal or a distorted square pyramidal geometry, with two water molecules occupying positions above and below the histidine brace plane. This coordination geometry is unique among the LPMO families, which usually display an axial coordination geometry in keeping with traditional type II copper sites. Both the alanine and the glutamate sidechains could be responsible for the arrangement of the exogenous ligands. The glutamate is positioned below the histidine brace plane and appears to form a hydrogen bond with one of the coordinating water molecules. The conserved alanine also appears

to be positioned above the histidine brace plane, blocking the axial position *via* steric crowding. The resulting geometry is 5-coordinate and rhombic. The cellulose active AA10s in comparison, retain a more axial coordination geometry, likely stemming from the absence of the strongly H-bonded glutamate/aspartate residue and the restoration of the weakly bonding tyrosine residue in the axial position beneath the plane of the histidine brace. The difference in substrate specificity is believed to stem from differences in certain key surface residues that have been linked to substrate binding.<sup>95</sup> Particularly, the carbohydrate binding surface appears to be less polar than in the chitin-active AA10s, presumably due to the requirement of binding a more apolar substrate. Large, negatively charged regions are present on the binding surface of the chitin-active AA10s which are expected to play an important role in their observed substrate specificity.<sup>96</sup>

### 1.5.3 The AA11s

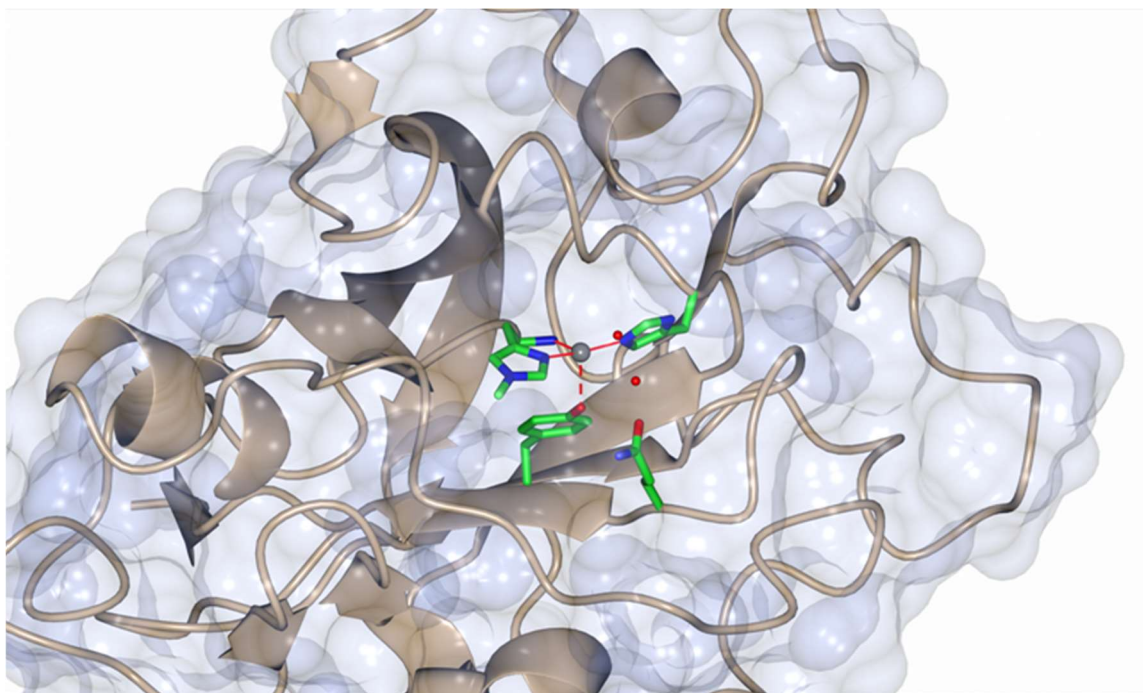
The discovery of the next family of LPMO, the AA11s, happened some three years later (2014) when Hemsworth *et al.* characterised an enzyme from *Aspergillus oryzae*, AoAA11, with a sequence distinctly unique from the previously discovered families.<sup>97</sup> The AA11 class was found to share structural similarities to both the AA9 and AA10 classes. The spectroscopic characteristics of the AA11 family also appears to share a likeness to both of the previously reported families, perhaps more so with the AA9s. The AA11s are primarily produced by fungal organisms to facilitate the degradation of biomass material. However, unlike the AA9s, these enzymes have been shown to be active on chitin.<sup>97</sup> The structure of AoAA11 appears to be in the Cu<sup>I</sup> state since it possesses no exogenous ligands. This is a common feature among the Cu-LPMO crystal structures as they are prone to photoreduction in the x-ray beam.<sup>98</sup> The active site of these enzymes display a long bond between the copper and an axial tyrosine, akin to the previously reported AA9 family. The presence of this tyrosine, therefore, can be excluded from the substrate specificity argument between chitin and cellulose. The second coordination sphere glutamic acid residue seen in chitin active AA10s also appears to be present in the AA11 family. The active site geometry of AoAA11 can be seen below in **Figure 26**.



**Figure 26. Crystal structure of the first discovered AA11 enzyme (PDB:4MAI), AoAA11, highlighting the primary and secondary coordination sphere residues shown as sticks.<sup>97</sup>**

#### 1.5.4 The AA13s

The next family to be discovered came in 2015 when Lo Leggio *et al.* discovered a new LPMO with a distinct genetic sequence that established the AA13 class in the CAZy database.<sup>99</sup> The first discovered member of the AA13 family was also found in the fungal organism, *Aspergillus oryzae*. The AA13s were shown to be *starch*-degrading LPMOs adding a new substrate to the scope of LPMO-mediated oxygenations. This was an interesting discovery as it identified a rather surprising new substrate for LPMOs given that starch is usually more tractable to enzymatic attack based on its increased solubility compared to cellulose. The discovery of the AA13 class raised the question as to the scope of what substrates are available to enzymatic breakdown by LPMOs. Although usually easier to degrade than cellulose, certain regions of starches can be heterogeneous and resistant degradation; necessitating the employment of LPMOs to increase their tractability. These AA13 LPMOs have shown to increase the breakdown rate of these starches by up to 100 times.<sup>99</sup>

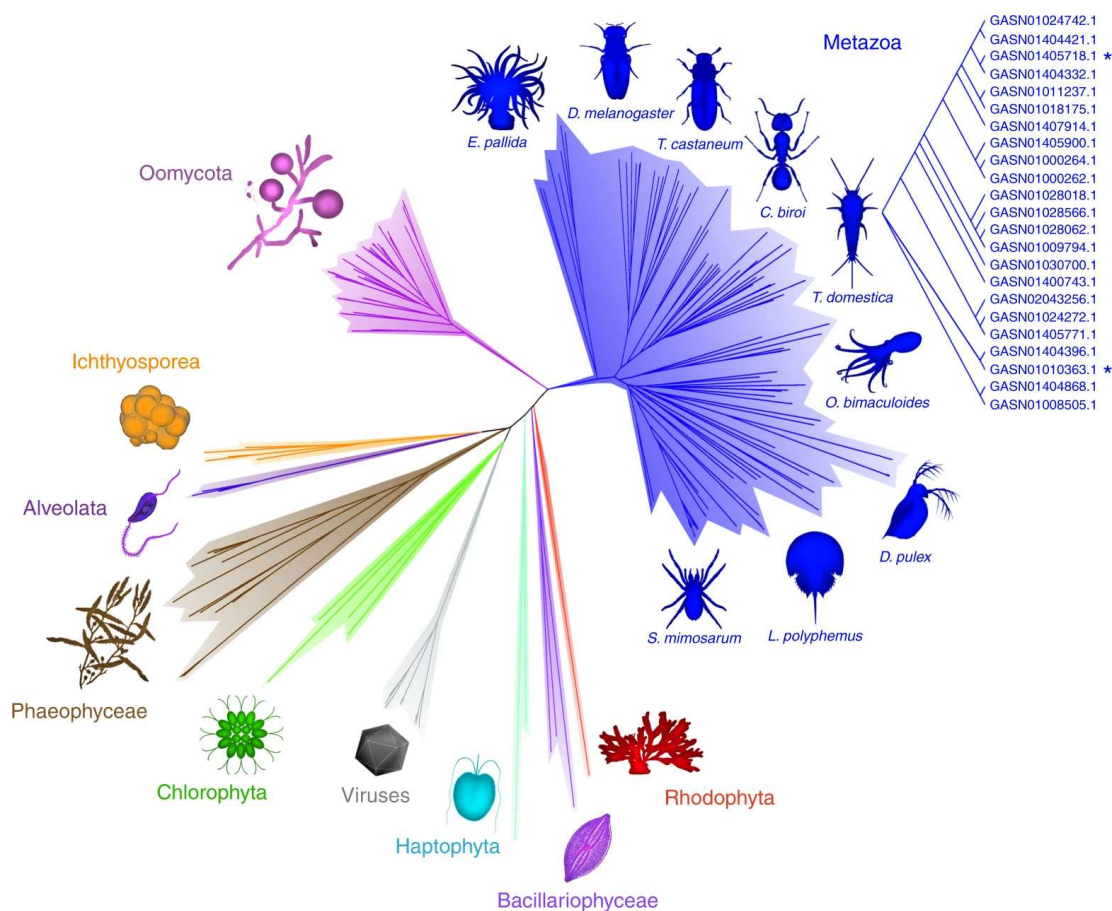


**Figure 27. Crystal structure of the first discovered AA13 enzyme (PDB:4OPB), AoAA13, highlighting the primary and secondary coordination sphere residues shown as sticks.<sup>99</sup>**

The active site geometry of this newly discovered family is in keeping with that of the AA9s. The coordination geometry represents a typical tetragonal geometry which is reflected in their EPR spectra. The crystal structure of AoAA13, however, shows two water molecules nearby the metal with incongruous bond lengths for a  $\text{Cu}^{\text{II}}\text{-OH}_2$  species (3.63 and 3.66 Å). Again, this is likely a consequence of photoreduction in the x-ray beam, this time showing a snapshot in time as the exogenous ligands dissociate from the metal following reduction as metal coordination becomes energetically unfavourable. The conserved residues that make up the AA13 active site appear to show great similarity with the AA9 class including the axially ligated tyrosine and the supporting glutamine residue.

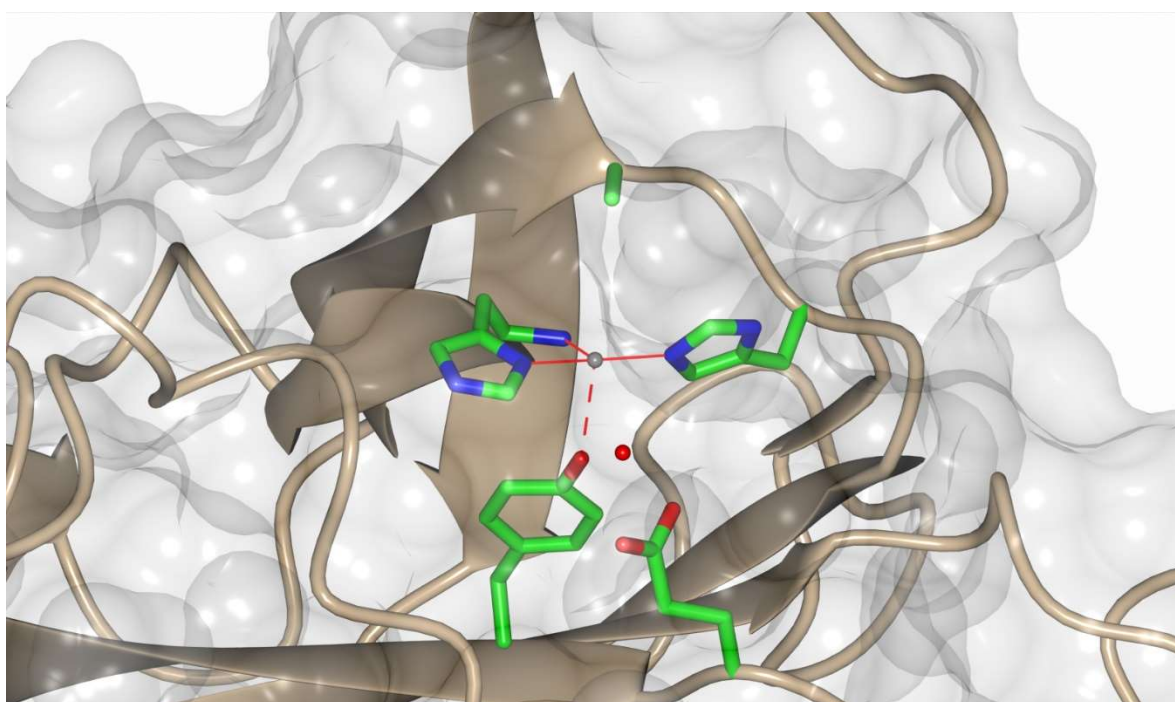
## 1.5.5 The AA15s

The next discovered family of LPMOs were the AA15s, discovered in 2018 by Sabbadin *et al.*<sup>50</sup> They discovered this new family of LPMO in *Thermobia domestica* which belongs to an ancient group of insects, with evidence of them existing some 420 million years ago.<sup>50, 100</sup> Blast searches of known databases based on the sequences found in *Thermobia domestica* revealed a vast and diverse list of organisms that hosted this newly discovered family of LPMO. These ranged from marine and invertebrates ranging from molluscs and crustaceans to insects and spiders. The discovery of this new AA15 family expanded the current understanding of how ubiquitous these enzymes are in nature. It is clear that a vast array of species depends on the recruitment of LPMOs facilitate the efficient breakdown of biomass. Interestingly, this widespread family of enzymes has shown to be active in the oxidation of *both* cellulose *and* chitin, the two most abundant biopolymers on earth. The biological diversity of this family of enzymes is visible by the phylogenetic tree (Figure 28).



**Figure 28. Phylogenetic diagram of the AA15 LPMO family across Taxa. Sequences of this family were found among animals (Metazoa), Oomycota, algae and many more. Figure sourced from Sabbadin *et al.*<sup>50</sup>**

From a structural perspective, the AA15 family possesses features in keeping with the previously reported families. They contain the axial, weakly coordinating tyrosine residue, akin to the AA9 enzymes while also containing an alanine residue whose sidechain imposes into the other remaining axial site with similar positioning as to that in AA10s. This particular active site arrangement appears closely related to the starch active AA11 family. The first characterised structure from *TdAA15* (PDB:5MSZ) shows a single water molecule in the active site vicinity. Again, the distance of this water molecule from the metal is very long (nearly 4 Å) consistent with the metal cofactor becoming photoreduced in the x-ray beam (**Figure 29**).



**Figure 29.** Crystal structure of the first discovered AA15 enzyme (PDB:5MSZ), *TdAA15*, highlighting the primary and secondary coordination sphere residues shown as sticks.<sup>50</sup>

## 1.5.6 The AA16s

The AA16 family of LPMO was discovered in 2019 by Chastel and co-workers from fungal *Aspergillus* species. Their work focussed on identifying LPMO candidates that could perform synergistically with *Trichoderma reesei* in the breakdown of cellulose. *Trichoderma reesei* is an organism rich with cellulolytic enzymes encoded into its genome and likely named after Elwyn Reese, a pioneer of biological cellulose degradation.<sup>44, 51</sup> This fungal organism is known to possess a host of enzymes that degrade cellulose and it has become a model system for making enzymatic cocktails.<sup>101</sup> The majority of this enzymatic arsenal is comprised of glycosidic hydrolases with only three AA9 LPMO genes coded for in its genome (most fungal saprotrophs contain many more AA9 LPMO genes than this, usually between 40-50).<sup>51</sup> Despite being studied for decades, Chastel *et al.* identified a unique family of LPMOs within several *Aspergillus* species including *Aspergillus niger* and *Aspergillus japonicus*. These AA16 LPMOs showed to boost the cellulolytic activity of the native *Trichoderma reesei* enzymatic cocktail and again built on the growing appreciation for how widespread LPMOs are in nature. As of yet, there are no crystal structures available for the AA16 family requiring further studies to enable better understanding of the nuances of the active site structure.

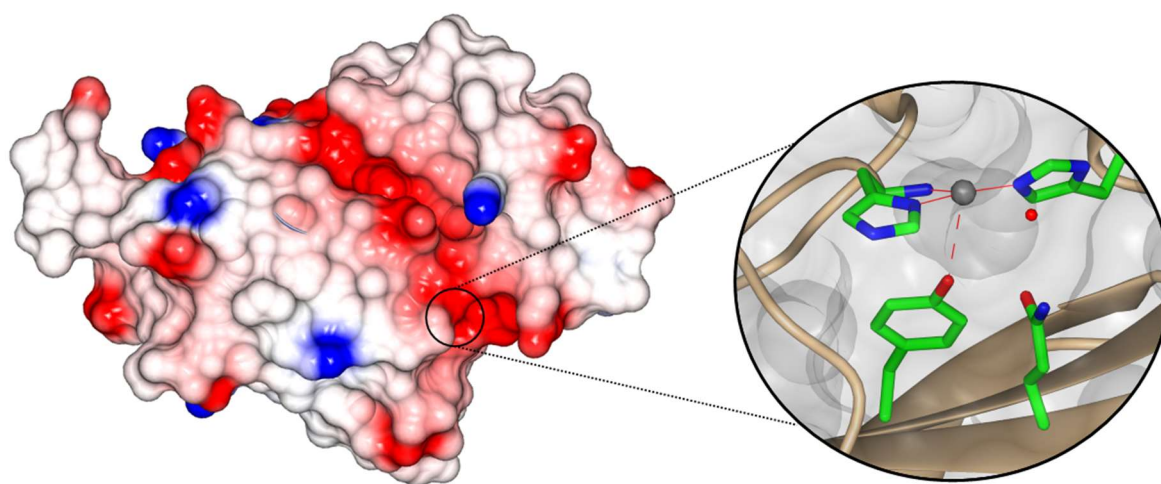
## 1.5.7 The AA17s

The most recently discovered family of LPMOs are the AA17s, which were discovered in 2021 by Sabbadin *et al.*<sup>52</sup> The AA17 family was first identified in the organism responsible for driving the potato famines in the 1800s, the oomycete *Phytophthora infestans*.<sup>102</sup> It is a pathogenic oomycete which causes late blight in species of potato and tomato (Solanaceae). The *PiAA17* enzyme itself serves as a virulence factor, facilitating the oomycetes entry into the host. In their experiments, they determined that this new family of LPMO performed the oxidative cleavage of polygalacturonic acid, the backbone of pectin. As discussed in section 1.2.4, polygalacturonic acid is a polysaccharide possessing negative charges due to the presence of carboxylate groups on the C6 carbon. Pectin is one of the largest components of plant cells walls, providing a primary layer of defence against pathogenic infection. Sabbadin and co-workers showed that *Phytophthora infestans* depends on the production of an AA17 LPMO to disrupt the plant cell wall structure allowing for the organisms' subsequent invasion of the plant. They demonstrated that the silencing of the LPMO gene rendered the organism



ineffective in plant pathogenesis. This was a significant finding as it opens up the avenue for new methods of crop protection, expanding the scope of LPMOs even further.

This discovery was also the first example of an LPMO that is active on a charged polysaccharide. Due to the hydrophilic backbone, pectin has unusual macro properties compared to other common polysaccharides, one of which being the ability to form gels.<sup>27</sup> To effectively bind pectin, the *PiAA17C* enzyme has evolved to contain a groove of charged residues to create substrate binding surface near the active site of the protein (negatively charged groove shown in red in **Figure 30**, left). This is a unique property of this family that provides some insight in how LPMOs show substrate specificity. As to the active site geometry, *PiAA17C* shows much similarity with the AA9s incorporating the axial tyrosine making up a canonical type II-like copper site (right, **Figure 30**).



**Figure 30.** Electrostatic potential of *PiAA17C* (PDB:6Z5Y) with blue showing positively charged and red as negatively charged (left) and exploded view of copper active site residues (right).<sup>52</sup>

## 1.6 LPMO Protein Preparation

Effective preparation and isolation of LPMO enzymes is a crucial part of their study. Different research groups have employed a range of approaches to produce and isolate LPMOs using several host organisms.<sup>45, 46, 63, 97, 103</sup> *Escherichia coli* is the most commonly chosen host organism used to express LPMO enzymes.<sup>104</sup> Competent *E. coli* cells (e.g., BL21 DE3) are regularly exploited in research laboratories by altering their genetic sequence with an engineered plasmid so that they possess the necessary genes to express the desired protein.<sup>105</sup> The ease of genetic modification and simple

cultivation make *E.coli* a highly attractive host organism for LPMO production, particularly those enzymes already originating from bacterial organisms (e.g., AA10s).

The majority of LPMOs characterised to date, however, are found in fungal hosts rather than bacteria.<sup>104</sup> This makes LPMO production using *fungal* expression systems highly attractive for these enzymes. There are clear advantages to producing LPMOs in this way including: the ability to perform post-translational modifications (including the  $\tau$ -N-methylation of the His1 sidechain) as well as glycosylations and ensuring correct protein folding.<sup>46</sup> The purpose of the methylation to the remote nitrogen on His1 remains ambiguous, though it is speculated to play a role in stabilising reactive oxygen species. Glycosylations of enzymes have been shown to have positive effects on both protein stability and secretion.<sup>106</sup> Most bacteria do not possess the machinery necessary to perform these PTMs -this being an obvious major drawback of expressing fungal LPMOs in *E.coli*. Though examples exist, LPMO protein expression using fungal hosts is far less established in most laboratories, with bacteria still proving to be the major workhorse of protein production.<sup>104</sup>

Competent *E. coli* cells are routinely transformed by incorporating engineered plasmids that encode the necessary genes to express the target protein. These engineered bacteria are typically grown in suspensions of Luria Bertani (LB) or Terrific Broth (TB) media which contain the necessary nutrients for healthy bacterial growth.<sup>107, 108</sup> In order to decrease the likelihood of contamination/infection of the growing bacterial cultures, an antibiotic resistance gene is often inserted into the bacterial genome. Typical antibiotics used for this purpose are kanamycin or ampicillin.<sup>109</sup> Successful insertion of an antibiotic resistance gene into the *E.coli* plasmid allow the practitioner to grow the desired culture in the presence of the antibiotic. By doing so, the risk of a rival organism infecting the culture is greatly reduced. The production of the desired protein is carefully controlled by an operon sequence immediately upstream of the LPMO gene. Typically, these are *lac*-operon style promoter/operator sequences, whereby the LPMO gene is not able to be transcribed until an inducer (usually a sugar such as IPTG or lactose) is added. The addition of the inducer releases the *repressor* protein from prohibiting transcription, allowing the desired LPMO gene to be transcribed and later, translated.

For most *E.coli* strains, initial cultures are optimally grown at 37 °C until they reach an optical density 600 nm (OD<sub>600</sub>) between 0.6-0.8.<sup>104</sup> Between these values bacterial cultures are understood to be at a suitable concentration to maximise protein production. Significantly less than 0.6 and the colony is typically too small for maximum protein production and much higher than 0.8 and the colony can become overcrowded and stressed for nutrients which can lead to apoptosis. Significant levels of cell lysis are highly undesirable when producing protein as the cell contents are released into the soluble medium and can make downstream purification significantly more challenging. Once this optimal

colony size is reached, the relevant inducer is added to trigger the overexpression of the LPMO protein. At this point the temperature is often decreased to better suit protein synthesis rather than cell reproduction (*ca.* 20 °C).

### 1.6.1 Periplasmic Expression

In the literature, recombinant LPMO expression has commonly been performed by secreting the overexpressed protein into the *periplasm* of the bacteria.<sup>94, 97, 103</sup> This is achieved by placing a signal gene immediately upstream of the target LPMO gene. The resulting translated protein contains a signal peptide that is recognised by the cell and is transported to the cell membrane (**Figure 31**). Once transported to the cell membrane, the polypeptide is translocated into the periplasmic space. The signal peptide is simultaneously cleaved without the requirement of any restriction enzymes. Enzymes secreted into the periplasmic space can be isolated through methods such as cold osmotic shock, whereby the permeability of the bacterial cell wall is modified using osmosis to retrieve the desired crude protein. This method of protein expression has been used to good effect to obtain untagged LPMOs.<sup>94, 97, 103</sup> Once the overexpressed protein has been extracted from the periplasm, it can be separated from the cell debris *via* centrifugation. Here, the cell material is pelleted, and the soluble protein remains in the supernatant. LPMOs at this stage are very crude and require additional purification steps precluding their detailed study. With bacterial LPMOs serving as extracellular enzymes (with their mode of action taking place outside of the cell) they are usually expressed in the periplasm in their parent organism. As such, achieving correct protein folding is unlikely to be an issue when producing LPMOs in this way.

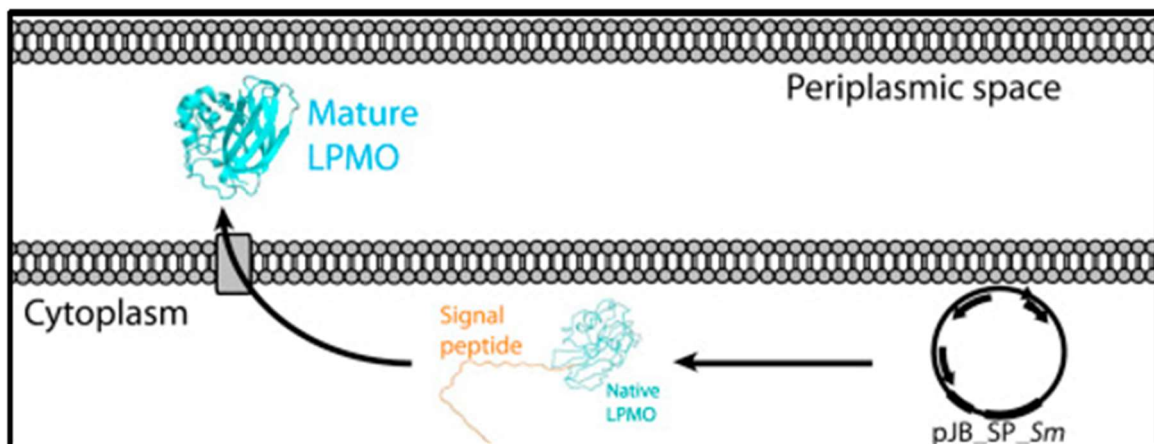


Figure 31. Periplasmic LPMO expression using a signal peptide. Figure adapted from Courtade *et al.*<sup>103</sup>

## 1.6.2 Cytoplasmic Expression

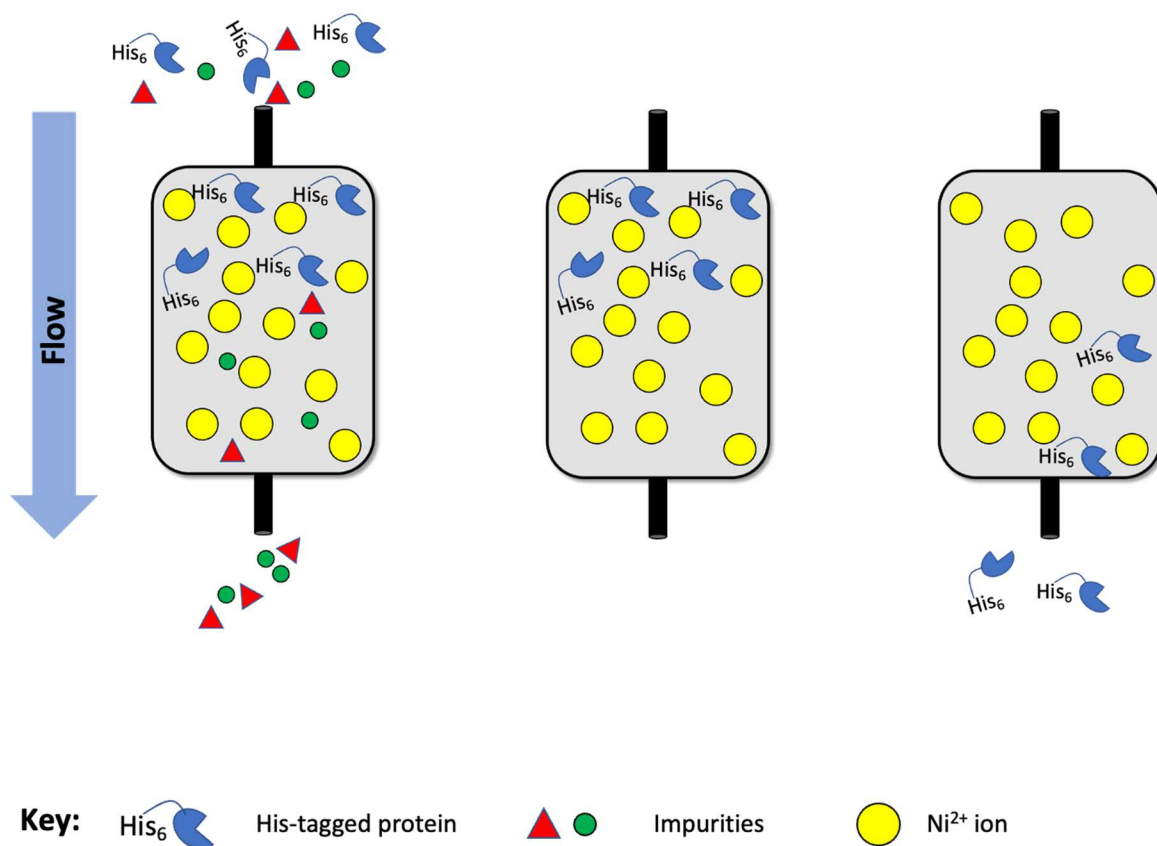
It is also possible, however, to express and obtain LPMO enzymes from the cytoplasm of the cultivated cells, negating the requirement of a signal gene/peptide. Proteins produced in the cytoplasm of *E. coli* can give good yields but can provide added complications.<sup>63,104</sup> Here, *E. coli* cultures are grown up and protein is expressed in much the same way as previously described (**Section 16**), but now missing the signal peptide. As a result, the protein is not translocated to the periplasmic space and remains in the cytoplasm. To obtain the overexpressed protein from the cytoplasm, the cells must be lysed. There are several ways to lyse cells, by far the most common of these are sonication or the addition of lysozyme enzyme (or sometimes both). The major drawback of cytoplasmic expression is that the protein must be extractable from the entire of the soluble contents of the host cell as lysis expels all cell contents into the medium. To combat this, experimentalists often employ the use of a peptide tag in order to provide a means of separating the desired protein from all other cell contents. These can take various forms, with the most common being a 'His-tag' or a 'Strep-tag'.

## 1.6.3 Protein Purification

To purify untagged LPMOs (e.g., the LPMOs often expressed in the periplasm), practitioners often employ the use of anion exchange chromatography. Proteins contain a net surface charge that is dictated by the pH of the buffer solution. At a particular pH known as the isoelectric point (IP), LPMOs contain no net surface charge. Careful buffer selection with a pH either above or below this IP value gives rise to a protein with either a negative or positive surface charge, respectively. This surface charge can be exploited to aid in protein purification using anion exchange chromatography. An anion exchange column contains a charged resin that binds to the charged surface of the proteins. Since different proteins vary in primary amino acid sequence their total net surface charge usually differs at a given pH. As a result, different proteins have a variable affinity for the columns stationary phase. This allows the practitioner to apply a salt gradient to elute and separate the proteins as the ionic strength of the mobile phase increases.

For LPMOs expressed in the cytoplasm, His-tags are usually added to the protein sequence to provide a more rigorous means of isolating the desired protein from the cell lysate contents. These amino acid tags are short peptides that can either be added to the N-terminus or the C-terminus side of the protein sequence. In the case of LPMOs, the N-terminus serves as a crucial active site ligand making it

essential for the tags removal if positioned there. A ‘His-tag’ typically consists of six consecutive histidine residues that are added onto the protein sequence. The histidine’s imidazole side chains act as electron donating groups that have a strong attraction toward electrophiles. As such, proteins with a His-tag can be trivially isolated from crude reaction mixtures by affinity chromatography employing a positively charged column, most commonly a Nickel column, or ‘His-trap’.



**Figure 32. Schematic of protein isolation using a nickel-affinity column of a His-tagged protein. Figure shows crude mixture entering the His-trap with impurities flowing through (left), His-tagged protein retained on the column (middle) and purified His-tagged protein eluting from the column with imidazole gradient.**

Affinity chromatography is performed by loading the filtered supernatant (containing crude protein) onto a His-trap column. The impurities are washed off the column using a buffer containing a low concentration of imidazole (typically  $\leq 30$  mM). The imidazole helps to mobilise compounds through the column so only strongly bound species are still retained (middle, **Figure 32**). The purified His-LPMO protein is then retrieved from the column by applying a buffer gradient to increase the imidazole concentration (right, **Figure 32**). Eluted fractions are analysed by gel electrophoresis and His-LPMO-containing fractions are pooled and up-concentrated. It is often desirable to remove His-tags from LPMO enzymes as they are known to interfere with copper binding and EPR studies.<sup>104</sup> To achieve this some research groups have utilised a cleavable SUMO-tag or TEV-tag to anchor the His-tag to.<sup>63, 110</sup>

This allows the practitioner to easily isolate the LPMO enzyme from crude mixtures using affinity chromatography, but then fully remove the tags using a SUMO or TEV protease enzyme, yielding the desired protein sequence.

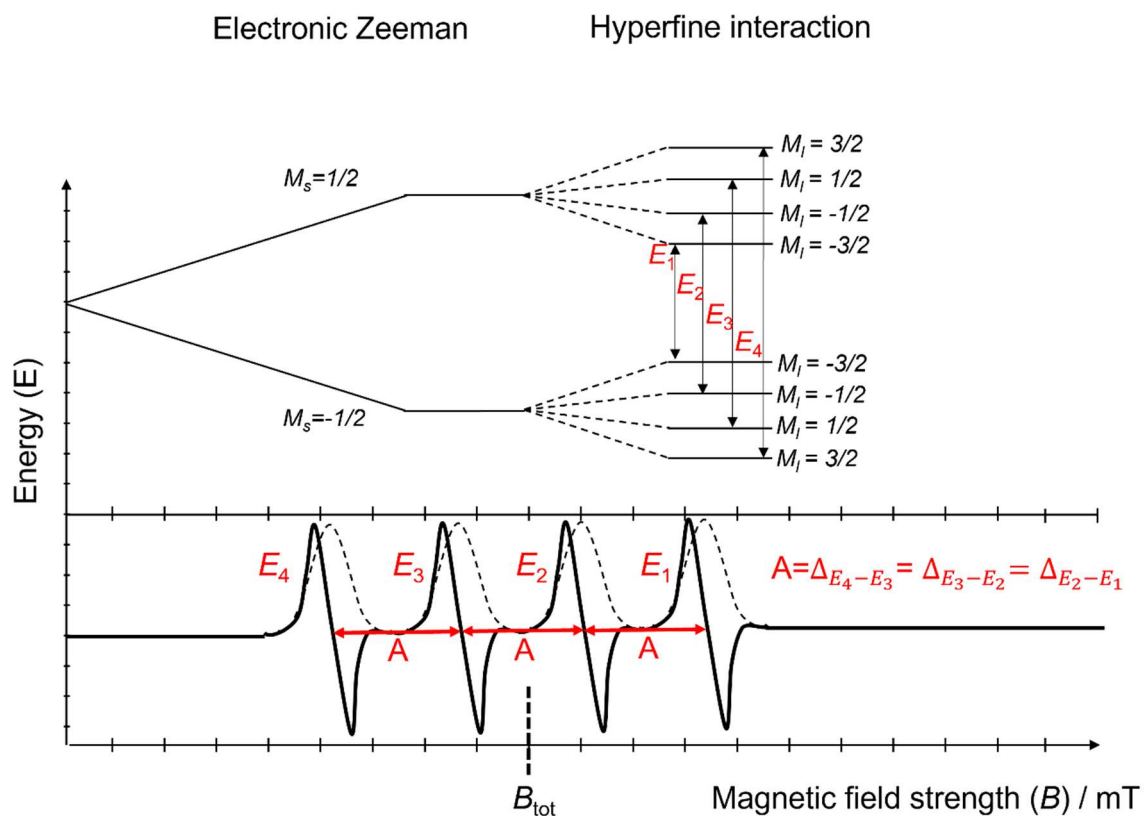
Following either anion exchange or affinity chromatography methods, it is common to perform a further purification step to ensure good purity of the final enzyme solution. -This is particularly important if a SUMO tag must be cleaved to enable separation of the LPMO from the SUMO tag and the SUMO protease. Size exclusion chromatography (SEC) is often the method used for this additional purification step. In SEC compounds are separated on a large column packed with a gel matrix. Compounds are separated based on size (molecular weight) with large compounds eluting fastest and small compounds eluting slowest. Again, fractions are identified by gel electrophoresis and purified LPMO-containing fractions are pooled and up-concentrated. Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is a common strategy to assess the purity (and identity) of protein throughout the purification process. It again separates species by size and can be semi-quantitative by introducing standard protein ladders of known molecular weight. Purified protein samples can then be used for crystallographic, kinetic or electron paramagnetic resonance studies.

## 1.7 Electron Paramagnetic Resonance (EPR) Spectroscopy

### 1.7.1 EPR Theory

While crystallographic studies have helped to provide considerable understanding of the variation in secondary and tertiary structures among the LPMO families, certain techniques such as electron paramagnetic resonance (EPR) spectroscopy can be more informative regarding the precise *electronic structure* of the LPMO active site. Since the resting state of the enzymes contains a single copper ion in the +2 oxidation state, they possess a  $d^9$  electron configuration with a single unpaired electron ( $S = \frac{1}{2}$  Kramer's doublet) in its d subshell.<sup>111</sup> Only species containing unpaired electrons are suitable for investigation using this method; eliminating enzymes either without metal cofactors or which contain metal cofactors with singlet ( $S = 0$ ) electronic ground states such as Zn(II) and Mg(II). The unpaired electron in the copper d orbital manifold has two spin states ( $\alpha$  and  $\beta$ ) denoted by the magnetic spin quantum number,  $m_s$  ( $m_s = \pm \frac{1}{2}$ ). In the absence of an external magnetic field these two spin states are degenerate in energy. However, when an external magnetic field is applied, the degeneracy of the two spin states is lifted: known as the Zeeman effect. Under this regime, the paramagnetic species will uptake a Boltzmann distribution favouring the population of the lower energy spin state.<sup>112</sup> The EPR method relies on exploiting this electronic Zeeman effect by irradiating

the sample with microwave energy and sweeping the magnetic field strength until the energy separation of the two spin states is in resonance with the microwave energy (left, **Figure 33**).



**Figure 33.** Electronic Zeeman and hyperfine interaction effect on the degeneracy of spin energy levels of a  $\text{Cu}^{\text{II}}$  ( $I = 3/2$ ) unpaired electron in an external magnetic field (top) and resonant energy absorption when microwave energy matches the energy separation between states ( $E_1$ - $E_4$ ) (dashed, bottom) depicted commonly as a first derivative (solid, bottom). In a  $\text{Cu}^{\text{II}}$  ion, the nuclear Zeeman interaction is much smaller than the hyperfine and excluded from the figure for clarity.

### 1.7.1.1 The Electronic $g$ -factor

EPR spectroscopy can be exceptionally useful in obtaining a wealth of knowledge including metal identity, oxidation state, coordination geometry/symmetry, degree of covalency, and the electronic ground states of paramagnetic complexes.<sup>113</sup> As well as applied magnetic field, an unpaired electron possesses its own local magnetic field caused by its own intrinsic angular momentum. The magnetic field caused by the electron ( $B_e$ ) adds to the applied field ( $B_0$ ) to give a total magnetic field ( $B_{tot}$ ) strength experienced by the unpaired electron.

$$B_0 + B_e = B_{tot}$$

EPR practitioners treat the disparity in the applied magnetic field *versus* the total magnetic field with a correction factor,  $g$ , or  $g$ -factor. An electron free in space, unperturbed by a nucleus, creates its own magnetic field that gives rise to an experimental  $g$ -factor of 2.0023. This value is known as the free-electron  $g$ -factor, or  $g_e$ . The perturbation to the total magnetic field varies depending on both the nucleus on which it resides and the immediate environment. For free-radicals localised on relatively light atoms (such as C, N & O) the perturbation to the magnetic field is very similar to that of the free electron resulting in  $g$ -factors very similar to the free-electron  $g$ -factor,  $g_e$ . However, for unpaired electrons situated on heavier elements (such as transition metals) there can be considerable deviations to the magnetic field experienced by the electron due to the significant introduction of *spin-orbit coupling* (SOC). This deviation can either add or take away from the applied magnetic field giving rise to the possibility of  $g$ -factors both greater and smaller than  $g_e$ . The values of these  $g$ -factors are an invaluable diagnostic as to the identity the radical (eg. a typical value for Cu(II) complex would be 2.2 and a typical value for V(IV) complex being 1.95). The  $g$ -factor determines at which field position the spin states come into resonance and in turn where we observe peaks in our EPR spectra. The energy separation between the  $m_s=+\frac{1}{2}$  and  $m_s=-\frac{1}{2}$  spin states is given by the following equation:

$$E = h\nu = \mu_B \times g \times B$$

Where  $h$  is Planck's constant ( $6.626 \times 10^{-34} \text{ J Hz}^{-1}$ ),  $\nu$  is the frequency of light,  $\mu_B$  is the Bohr magneton ( $9.3 \times 10^{-24} \text{ J T}^{-1}$ ),  $g$  is the  $g$ -factor, and  $B$  is the magnetic field strength (T).

For a gaseous transition metal ion unperturbed by ligating atoms, the  $g$ -factor is intrinsically related to the angular momentum quantum numbers, described by the Landé formula.<sup>114</sup>

$$g = 1 + \frac{J(J + 1) + S(S + 1) - L(L + 1)}{2J(J + 1)}$$

Where  $g$  is the  $g$ -factor,  $J$  is the total angular momentum,  $L$  is the orbital angular momentum, and  $S$  is the spin angular momentum.

In the presence of a ligand field, the coordinating atoms have the effect of quenching the orbital angular momentum,  $L$ , so that the total angular momentum,  $J$ , become equal to the spin angular momentum,  $S$ .

$$L = 0$$

$$\therefore J = S$$

The quenching effect that the ligands have on the orbital angular momentum means that transition metal ions *should* have a  $g$ -factor close to that of the free electron,  $g_e$  (2.0023). However,



spin-orbit coupling reintroduces some angular momentum which perturbs this value away from  $g_e$ . Spin-orbit coupling becomes significant for larger atoms ( $\geq 3$  row elements) due to a fourth power dependency on effective nuclear charge ( $\lambda \propto Z_{\text{eff}}^4$ ).<sup>115</sup> Therefore, deviations from  $g_e$  become more significant for the heavier elements appearing later in the periodic table. For example, the spin orbit coupling constant,  $\lambda$ , of a free carbon atom is  $29 \text{ cm}^{-1}$  compared to  $-830 \text{ cm}^{-1}$  for a free Cu(II) ion.<sup>68, 116</sup> This disparity in spin orbit coupling constants explains the differing  $g$ -factors for C and Cu(II) based radicals.

For radical species that interact with neighbouring atoms or ligands, the  $g$ -factor becomes anisotropic. The anisotropy of these systems means that the electronic Zeeman interaction is dependent on the orientation of the paramagnetic species with respect to the external magnetic field. This is a particularly important factor when the unpaired electrons are localised in non-symmetric atomic orbitals, such as p and d, or in molecular orbitals which contain a significant percentage of these atomic orbitals. In these circumstances, the  $g$ -factor is better described a matrix, or tensor, which can be diagonalised to provide three principal values accounting for the 3 principal directions ( $g_x$ ,  $g_y$ , and  $g_z$ , or  $g_1$ ,  $g_2$ , and  $g_3$ ).<sup>116</sup> The geometry of the ligands around a metal complex affects the electronic structure of the metal ion by perturbing the energies of the d-orbitals. As such, the number, orientation, and type of ligands determine the ordering of the d-orbital energies and hence, the overall electronic ground state of the metal complex.

For  $d^9$  transition metal ions like Cu(II), the common electronic ground states are  $d(x^2-y^2)$  or  $d(z^2)$ . Copper (II) complexes are commonly tetra-, penta-, or hexa-coordinate and adopt a number of coordination geometries including elongated octahedron, square pyramidal, square planar, tetrahedral, or trigonal bipyramidal.<sup>114</sup> In the square planar/pyramidal or elongated octahedral cases, the ground states are usually  $d(x^2-y^2)$  due to the electrostatic repulsion of the equatorial ligands elevating the energy of the orbital making it energetically unfavourable to be doubly occupied. However, in the trigonal bipyramidal case, the electronic ground state is typically  $d(z^2)$  due to a combined depression in  $d(x^2-y^2)$  energy and a corresponding elevation in  $d(z^2)$  energy. The values of the  $g$  matrix are informative of the energy separations of the d-orbital manifold. As such, the evaluation of the  $g$ -values from EPR spectroscopy can be used to decipher the electronic ground states and hence the geometry of transition metal complexes. The values of the  $g$ -matrix report on the energy separation of the d orbitals and are reflective of the overall strength of the ligand field.

The perturbation to the  $g$ -factors depends on covalency, the degree of SOC, and the energy separation between the electronic ground state and suitable, magnetically coupled excited states shown by the general equation:

$$g = g_e - \frac{\alpha^2 n \lambda}{\Delta E}$$

Where  $g$  is the observed  $g$ -factor,  $g_e$  is the free-electron  $g$ -factor,  $\alpha^2$  is a covalency parameter ( $\leq 1$ ),  $\lambda$  is the SOC constant,  $\Delta E$  is the energy separation between the electronic ground state and the suitable excited state, and  $n$  is a quantum mechanical coefficient determining the degree of mixing with the excited state.<sup>116</sup> For transition metal ions showing a high degree of covalency with its ligands, the value of  $\alpha^2$  decreases leading to smaller deviations from  $g_e$ .

The magnitude of  $n$  in this equation denotes the degree of mixing between the ground state with suitable excited states and is determined by group theory. This coefficient is defined by the overlap of the orbital angular momentum operators ( $\widehat{L}_x, \widehat{L}_y, \widehat{L}_z$ ) in the three principal cartesian directions,  $x, y, z$ , described by the following equation:

$$\frac{n_{ij}}{\Delta E} = \sum_{n \neq 0} \frac{\langle 0 | \widehat{L}_i | n \rangle \langle n | \widehat{L}_i | 0 \rangle}{E_n - E_0}$$

Where  $|0\rangle$  corresponds to the ground state wave function and  $|n\rangle$  is the wave function of the  $n$ -excited states. In the case of a Cu(II) ion with a  $|x^2 - y^2\rangle$  ground state ( $|0\rangle$ ), the  $g_z$  parameter therefore solves as:

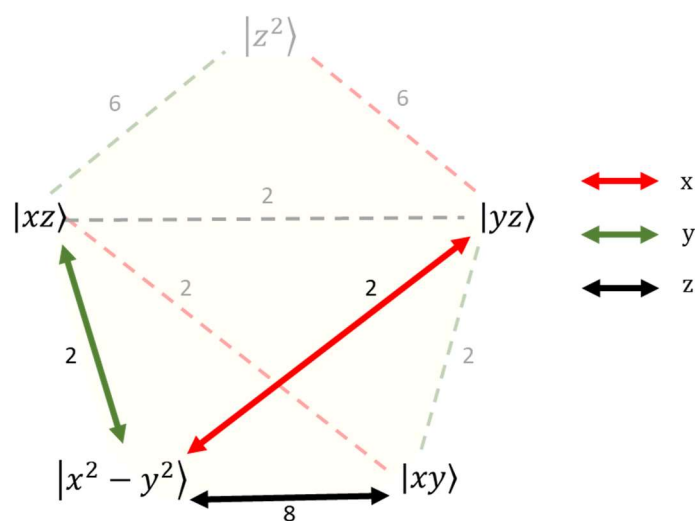
$$g_z = g_e - \lambda \sum_{n \neq 0} \frac{\langle 0 | \widehat{L}_z | n \rangle \langle n | \widehat{L}_z | 0 \rangle}{E_{|xy\rangle} - E_{|x^2-y^2\rangle}} = 2.0023 - \frac{8\lambda}{E_{|xy\rangle} - E_{|x^2-y^2\rangle}}$$

and assuming tetragonal symmetry,  $g_x$  and  $g_y$  are equivalent and solve for:

$$g_{x,y} = g_e - \lambda \sum_{n \neq 0} \frac{\langle 0 | \widehat{L}_{x,y} | n \rangle \langle n | \widehat{L}_{x,y} | 0 \rangle}{E_{|xz\rangle/|yz\rangle} - E_{|x^2-y^2\rangle}} = 2.0023 - \frac{2\lambda}{E_{|xz\rangle/|yz\rangle} - E_{|x^2-y^2\rangle}}$$

In the case of a  $d^9$  Cu(II) complex in an axial coordination geometry, the ground state of the metal ion is essentially described by a function which resembles  $|x^2 - y^2\rangle$ . The  $|x^2 - y^2\rangle$  ground-state can interact with excited states of suitable symmetry ( $|xz\rangle, |yz\rangle$ , and  $|xy\rangle$ ) when placed in an external magnetic field.<sup>116, 117</sup> In this example,  $E_0$  is the energy of the  $|x^2 - y^2\rangle$  ground state and  $E_n$  in the  $g_z$  equation is the energy of the  $|xy\rangle$  excited state and  $E_n$  in the  $g_{x,y}$  equations is the energy of the  $|xz\rangle$  and  $|yz\rangle$  excited states. Therefore, the corresponding values of  $g_x, g_y$  and  $g_z$  values can be used to ascertain the energy separation between the electronic ground state and suitable excited states, providing valuable information regarding the electronic structure of the paramagnet. A

simplified depiction of how the electronic ground state of a  $d^9$  transition metal ion can interact with d orbitals of suitable excited states is shown by the ‘magic pentagon’ where the values of ‘ $n$ ’ are shown on the connecting vectors (**Figure 34**).<sup>116, 117</sup>



**Figure 34.** The ‘magic pentagon’ showing how spin orbit coupling connects suitable excited states to the different possible electronic ground states of a  $d^9$  transition metal complex. Red, green, and black vectors show excited states connected by  $g_x$ ,  $g_y$ , and  $g_z$ , respectively. Values of ‘ $n$ ’ are shown on connecting vectors.

By accounting for the values of ‘ $n$ ’ shown in **Figure 34**, the general formula for  $g$  can be rewritten for each of the anisotropic  $g$ -values.  $g_x$ ,  $g_y$ , and  $g_z$ .<sup>118</sup>

$$g_x = g_e - \frac{2\lambda}{\Delta_{yz}}; \quad g_y = g_e - \frac{2\lambda}{\Delta_{xz}}; \quad g_z = g_e - \frac{8\lambda}{\Delta_{xy}};$$

The directionality of the  $g$ -factor can be understood by visualising the orientation of the d orbitals involved in transition. For example, the  $g_z$  parameter for a radical with a  $|x^2 - y^2\rangle$  ground state depends on the energy separation between the  $|xy\rangle$  excited state. By considering the orientation of both  $d(xy)$  and  $d(x^2 - y^2)$  orbitals, they can be superimposed by applying a rotational symmetry operation about the  $z$  axis ( $R_z$ ). The formulae, in effect, describe around which axis the electronic ground state can be transformed into the coupled excited state by applying a simple symmetry operation. The circulation of the electron around the axis generates a magnetic field (in this case the  $z$ -direction) which contributes to the observed total magnetic field in the experiment, thus perturbing the value of  $g$ .

For a Cu(II) complex with a perfect square-planar/pyramidal coordination geometry, the energies of the  $|xz\rangle$  and  $|yz\rangle$  are degenerate. This results in two of the principal values of  $g$  matrix

being equivalent  $g_x = g_y \neq g_z$ , known as an *axial* system. Complexes with lower symmetry have totally non-degenerate d orbital manifolds and hence 3-nondegenerate g-values, known as a rhombic system.

### 1.7.1.2 Hyperfine Coupling – the Effect of *Nuclear Spin*

As well as probing the electronic spin states and the related g-factor, the EPR experiment is also sensitive to magnetically coupled nuclei with a non-zero nuclear spins ( $I \neq 0$ ). The coupling of the nuclear spin to the electron spin resonance gives rise to fine structure, known as hyperfine coupling (HFC). Nuclei with a non-zero spin quantum number ( $I \neq 0$ ) possess intrinsic orbital angular momentum which produce a magnetic moment,  $\vec{\mu}_I$ , which interacts with the magnetic moment arising from the spin angular momentum  $\vec{\mu}_S$ . Using copper as an example, the two common isotopes for copper ( $^{63}\text{Cu}$  and  $^{65}\text{Cu}$ ) both possess a non-zero nuclear spins ( $I = 3/2$ ) and as such, the Cu(II) EPR spectra display hyperfine coupling. The number of peaks arising from the copper hyperfine coupling obeys the  $2nI + 1$  rule (where n is the number of coupled nuclei and I is the nuclear spin) resulting in 4 separate peaks separated by a hyperfine coupling constant, A (usually given in units of MHz). The magnitude of this coupling constant varies on the gyromagnetic ratio of the nucleus on which the unpaired electron is located and the degree of covalency to which the nucleus is bonded to neighbouring atoms. A radical that resides on an atom which participates in negligible covalent bonding resides primarily on a single nucleus, giving rise to larger hyperfine coupling constants. Conversely, unpaired electrons that experience significant covalency are more spread across neighbouring nuclei resulting in reduced hyperfine coupling constants arising from the parent nucleus.

There are two principal mechanisms which contribute to the overall hyperfine interaction, the dipole-dipole (or ‘spin dipolar’) interaction, denoted by T, and the Fermi contact interaction, denoted by  $a_{\text{iso}}$ . The overall hyperfine interaction, therefore, is the total of these two principal interactions:  $A = A_{\text{iso}} + T$ . The spin dipolar term (T) arises from the direct interaction between the dipoles produced by the non-zero electron and nuclear spins. The electron-spin dipole moment is averaged over the shape of the singly occupied orbital, e.g.  $d(x^2-y^2)$ , and interacts with the nuclear-spin dipole. The dipolar contribution to the hyperfine coupling depends on the distance, r, and the angle,  $\theta$ , between the magnetic moments.<sup>119</sup> The magnitude of the dipolar contribution to the hyperfine coupling has an inverse, cubic relationship with the distance between the interacting dipoles ( $A_{\text{dip}} \propto \frac{1}{r^3}$ ).<sup>120</sup> Since the dipole-dipole hyperfine parameter is orientation dependent, it is anisotropic and is best described by a tensor with individual values for x, y, and z ( $A_x$ ,  $A_y$ , and  $A_z$ ). In an axial system, the diagonalised components of the anisotropic hyperfine matrix are as follows:  $T = [-T, -T, 2T]$

yielding the hyperfine terms:  $A_x = A_y = a_{iso} - T$ , and  $A_z = a_{iso} + 2T$ .<sup>117</sup> As a result, a paramagnetic Cu(II) complex with an axial  $d(x^2-y^2)$  ground state is expected to have a large  $A_z$  parameter due to the large  $2T$  spin dipolar interaction. This results in a dominant hyperfine interaction in the z-direction of the spectrum. As an anisotropic contribution, the sum of the dipolar hyperfine coupling constants always equals zero.<sup>68</sup> The dipole-dipole interaction dominates for paramagnetic species where the unpaired electron resides in an orbital with a node at the nucleus (p, d, or f). For paramagnetic species where the unpaired electron resides in an s-orbital, or molecular orbitals containing some s-character, the electron now possesses non-zero electron density at the nucleus. Non-zero electron density at the nucleus allows for the direct contact between the electronic and nuclear spins, known as the Fermi contact interaction.

The Fermi contact interaction provides a purely isotropic part of the overall hyperfine interaction ( $a_{iso}$ ), contributing equally to the individual terms of the hyperfine tensor,  $A_x$ ,  $A_y$ , and  $A_z$ .<sup>68, 121</sup> It arises from finite the electron density at the nucleus that is present when there is some s-orbital character in the SOMO.<sup>116</sup> The overall magnitude of the Fermi contact term to the overall hyperfine interaction thus depends on the degree of s-orbital character in the SOMO. For example, a SOMO with  $sp^3$  hybridisation contains reduced s-orbital character than a SOMO with  $sp^2$  hybridisation (25% vs 33%). Therefore, a carbon-based radical with a  $sp^2$  SOMO will display larger hyperfine constants compared to that of the  $sp^3$  analogue due to the larger s-orbital character and hence, larger Fermi contact contribution. A final (and often minor) mechanism exists which can contribute to the overall hyperfine coupling arising from *spin polarisation* of orbitals.<sup>122</sup>

Additional fine structure can sometimes arise from coupling to adjacent nuclei which are magnetically coupled to the unpaired electron. Such fine structure is known as *superhyperfine* coupling (SHFC). For example, in biological systems a histidine sidechain can ligate a copper ion leading to superhyperfine coupling, whereby fine structure can be resolved due to the unpaired electron on the copper nucleus being magnetically coupled to the  $I = 1$  nitrogen ligand. Radicals that experience significant covalent bonding (such as Type I copper sites) will be better shared across multiple nuclei, reducing the hyperfine coupling constant arising from the parent nucleus.

The relationship between the covalency of the ligands and the resulting hyperfine/ $g$ -values of copper complexes has led to the development of “Peisach Blumberg plots” or “PB plots” whereby the quick evaluation of the  $g_z$  and  $A_z$  parameters can be informative of the coordinating ligands.<sup>123</sup> A highly covalent ligand field such as in Type I copper proteins (2N2S) results in reduced  $A_z$  hyperfine couplings and  $g_z$  values when compared to the Type II. The dependency of these values on the ligands allows an EPR practitioner to gauge the identity of the ligating atoms by placing them on a PB plot (**Figure 35**).

In more covalent systems, the better sharing the electron across multiple nuclei oftentimes leads to the resolution of greater *super*hyperfine couplings (providing the coupled nuclear is  $I \neq 0$ ). The consequence of all these features leads to highly complex and multi-featured spectra containing vast amounts of information about a paramagnetic chemical species.

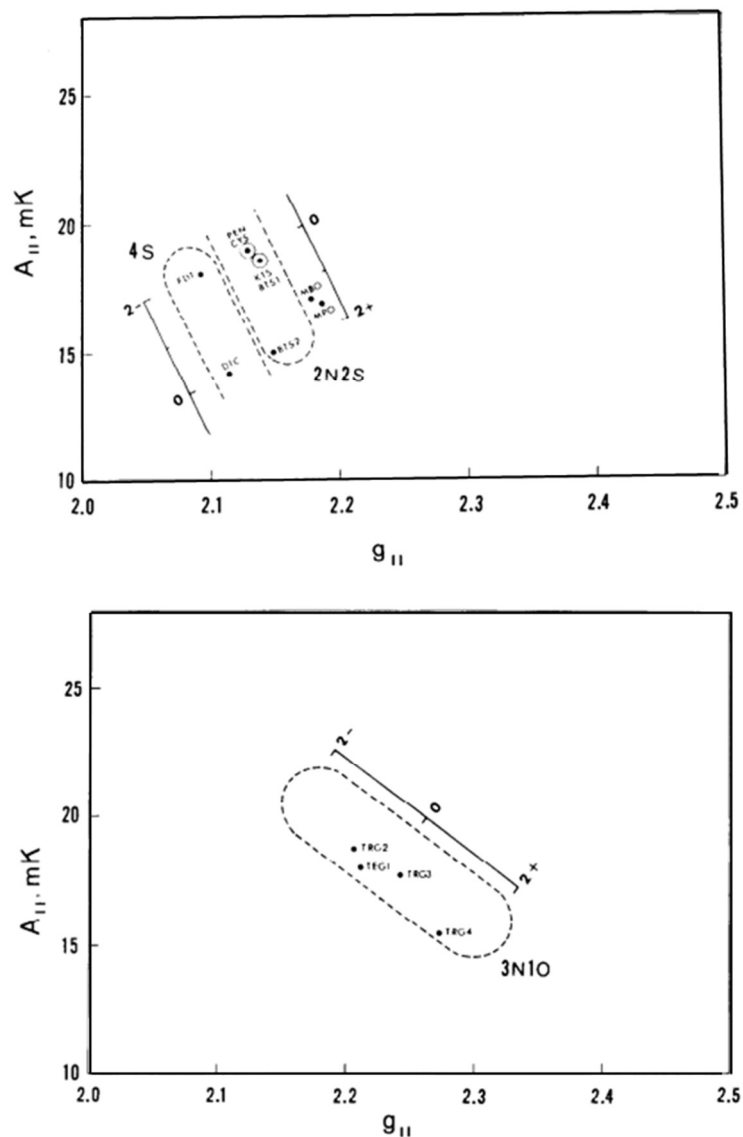


Figure 35. Peisach Blumberg plot of Type 1, 2N2S, copper complexes (top) and Type 2, 3N1O, copper complexes (bottom). Figure adapted from Peisach and Blumberg, 1974.<sup>123</sup>

### 1.7.1.3 The Effective and Spin Hamiltonian

The electronic behaviour of radical systems can be well described using an *effective Hamiltonian*. An effective Hamiltonian describes only part of the eigenvalue spectrum of a full

Hamiltonian and can have varying degrees of appropriateness. A spin Hamiltonian, regularly used to describe the characteristics of radical species, is usually a very appropriate effective Hamiltonian; able to fully describe the electronic behaviour with a few parameters and operators. A spin Hamiltonian comprises of spin electron and nuclear operator terms that are dependent on the magnetic field (B) and two magnetic field independent terms: the zero-field splitting and hyperfine operators. These operators describe how the unpaired electron spin angular momenta interact with one another and how they interact with nuclear spin angular momenta, respectively, shown by the following equations:

$$\hat{H}_0 = \hat{H}_{EZ} + \hat{H}_{NZ} + \hat{H}_{HFI} + \hat{H}_{ZFS}$$

$$\hat{H}_0 = \underbrace{\beta_e S_g B}_{\text{Spin Hamiltonian}} + \underbrace{\beta_N I_g B}_{\text{Electronic Zeeman}} + \underbrace{SAI}_{\text{Nuclear Zeeman}} + \underbrace{SDS}_{\text{Hyperfine interaction}} + \underbrace{\quad}_{\text{Zero field splitting}}$$

The values incorporated in the spin Hamiltonian include the anisotropic *g*-factors, hyperfine coupling constants (A) and zero field splitting parameters. As discussed previously, these values are informative as to the identity of the radical, the coordinating atoms, as well as the overall geometric structure. In order to obtain all such values describing the spin Hamiltonian, an EPR spectrum must be simulated. Once obtained, vast amounts of information can be learned regarding the precise electronic structure of a given species.

## 1.7.2 The EPR Experiment

EPR is an exceptionally sensitive technique due to the large gyromagnetic ratio of an electron. A proton in comparison has a gyromagnetic ratio three orders of magnitude lower making an EPR experiment *ca.* 1000 times more sensitive than the analogous NMR study (**Table 1**). This allows for the detailed investigation of paramagnetic systems even in relatively dilute systems. This makes EPR a fantastic spectroscopic method for studying paramagnetic metalloenzymes as protein solutions are rarely stable much beyond 1 mM in concentration. Typically, a protein solution of concentrations  $\geq$  0.1 mM with a total volume of 0.1 mL is suitable for study using EPR spectroscopy. These enzyme solutions are typically placed into high-quality, quartz EPR tubes which lack any paramagnetic metals or radicals that may perturb the signal.

**Table 1. Gyromagnetic ratios of sub-atomic particles**

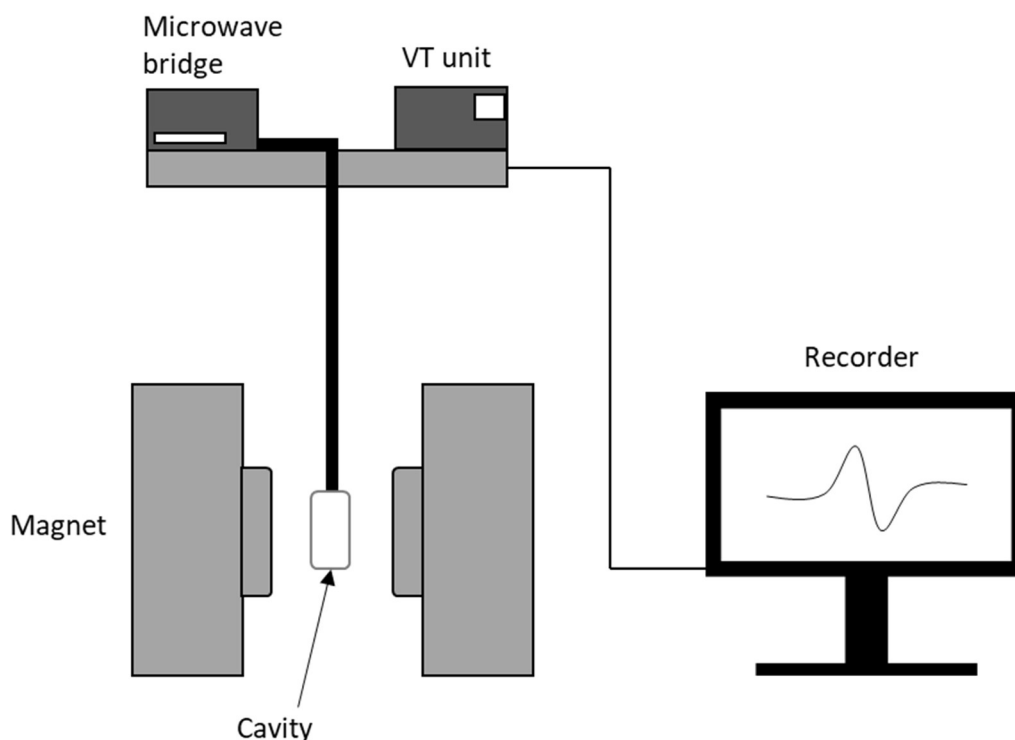
Particle	Gyromagnetic ratio / $\text{rad}\cdot\text{s}^{-1}\cdot\text{T}^{-1}$
Proton	$2.68 \times 10^8$
Electron	$1.76 \times 10^{11}$

### 1.7.2.1 Continuous Wave (cw) EPR

The EPR technique can be informative of both the *solution* or *frozen solution* phase of LPMO enzymes. Since enzymes almost always operate within an aqueous medium, the direct study of the active site electronic structure in (frozen) solution is important for our understanding of their reactivity. Studies of LPMO *in crystallo* often require specific conditions in order to develop protein crystals that may lead to a copper site containing a ligand field inconsistent with that in the solution phase. The typical EPR experiment utilises a continuous-wave (cw) of microwave radiation where the sample is continuously irradiated with a low intensity, fixed frequency microwave radiation whilst the magnetic field strength is swept.<sup>113</sup> As the *g*-values are specific to the identity of the radical, the resonant field position can vary significantly between different systems (particularly with different transition metals). As such, the EPR practitioner will have to scan a wide field range for an unknown sample to ensure they are obtaining all the available information.

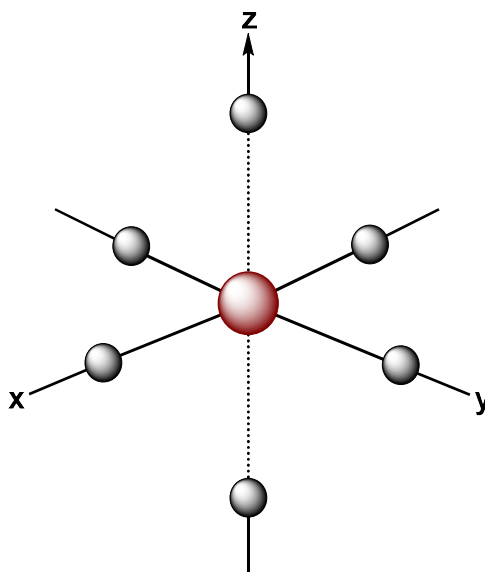
EPR spectra of copper complexes are often collected at low temperatures (*ca.* 150 K) mediated by a variable temperature (VT) control unit which controls the flow rate of liquid nitrogen to the sample cavity. These temperatures create frozen solutions of the enzyme samples in which each enzyme in the frozen solution possesses a random orientation with respect to the applied magnetic field. The EPR spectra of frozen solution samples are as such, anisotropic, giving the practitioner additional information on the paramagnetic species. Unlike organic radicals, the anisotropic *g*-values of copper complexes require large magnetic field ranges to fully separate due to the introduction of spin-orbit coupling (SOC) from the transition metal into the spin Hamiltonian. Typical microwave frequencies most often used for EPR experiments are S-band (2-4 GHz), X-band (*ca.* 9.3 GHz), and Q band (*ca.* 35 GHz) with X-band being by far the most common. Usually, a combination of X and Q-band is suitable to appropriately ascertain the desired *g*-factors and HFC constants of most copper complexes. An example cw-EPR experimental setup can be seen in **Figure 36**.





**Figure 36. Schematic diagram of the essential hardware of an EPR spectrometer highlighting the microwave source (bridge), electromagnet, variable temperature (VT) unit, sample cavity, and recording device.**

The anisotropic spectra from frozen solution EPR experiments consists of three individual  $g$  and  $A$  values (corresponding to orthogonal  $xyz$  directions), commonly denoted as either:  $g_x, g_y, g_z$  and  $A_x, A_y, A_z$ , or  $g_x, g_y, g_z$  and  $A_x, A_y, A_z$ . A typical  $d^9$  copper complex possesses tetragonal symmetry after undergoing Jahn-Teller distortion away from octahedral (**Figure 37**). This elongation of a ligand axis results in a coordination geometry that has two similar or identical axes (usually  $x/y$  or  $1/2$ ) and one unique axis (usually  $z$ , or  $3$ ). These complexes are referred to as axial, whereby they contain two similar parameters for  $x/y$  and one unique for  $z$  ( $g_x \approx g_y \neq g_z$  and  $A_x \approx A_y \neq A_z$ ).<sup>114</sup> Unlike free complexes in solution, enzymes can adopt a multitude of different ligand field arrangements in their metal active sites dictated by the positioning specific amino acid sidechains. As a result, metal cofactors in the active sites of proteins can contain ligand field arrangements with surprising symmetries for their  $d$ -electron counts, or which adopt highly asymmetric conformations altogether.



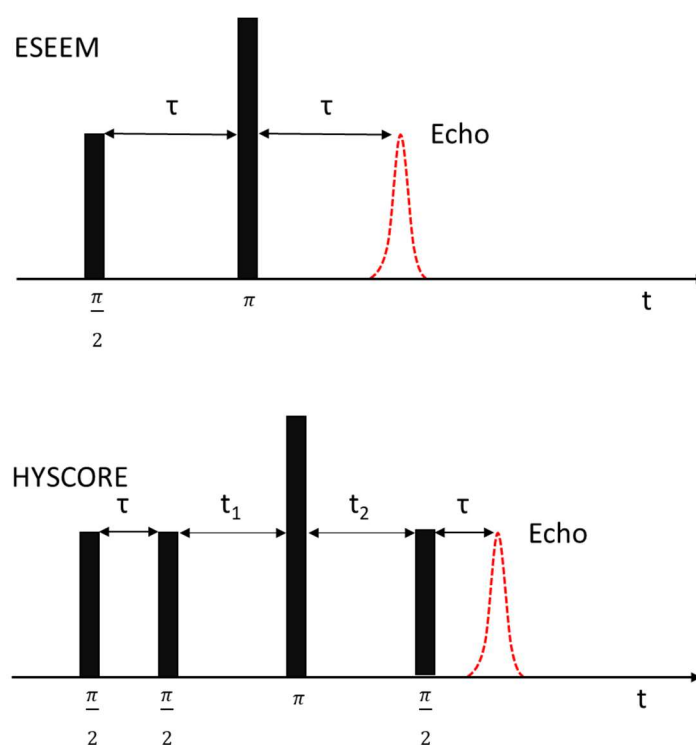
**Figure 37. Example coordination geometry of a  $d^9$  copper complex displaying a Jahn-teller distorted axial geometry. Copper ion shown as copper sphere and ligands shown as grey spheres.**

### 1.7.2.2 Advanced EPR Techniques (Pulsed and HYSCORE)

Despite the wealth of information possible to obtain using the cw-EPR experiment ( $g$ -values, hyperfine and superhyperfine couplings), smaller SHFCs can prove challenging to resolve due to the magnitude of the coupling often being smaller than the experimental line width. In such cases, the development of pulsed techniques have shown to be highly valuable in the field of EPR spectroscopy.<sup>117</sup> Pulsed microwave sequences offer greater spectral resolution of these small SHFCs and are often employed as complementary techniques to cw-EPR spectroscopy. Used in tandem with cw-EPR, pulsed EPR techniques allow a practitioner to extract as much information about the paramagnetic species as possible. There are a multitude of pulsed techniques, with the most popular ones including: pulsed Electron Nuclear Double Resonance (ENDOR) spectroscopy and HYperfine Sublevel COrRELation (HYSCORE) spectroscopy. For a detailed introduction into these techniques, I direct the reader to a comprehensive text by Schweiger and Jeschke.<sup>124</sup> Briefly, the ENDOR experiment utilises a combination of both microwave and radio irradiation frequencies to simultaneously probe the EPR and NMR transitions (hence, a *double* resonance technique). The pulsed ENDOR technique is particularly useful in determining SHFCs of strongly coupled nuclei (i.e. primary coordination sphere nuclei).<sup>113</sup> In the case of the LPMO Cu(II) paramagnet, the ENDOR technique could be useful in the extraction of the SHFC values of the coordinating nitrogen atoms.

For the detection of more weakly coupled nuclei (e.g., the remote imidazole nitrogen atoms of the histidine brace), a technique known as Electron Spin Echo Envelope Modulation (ESEEM) is

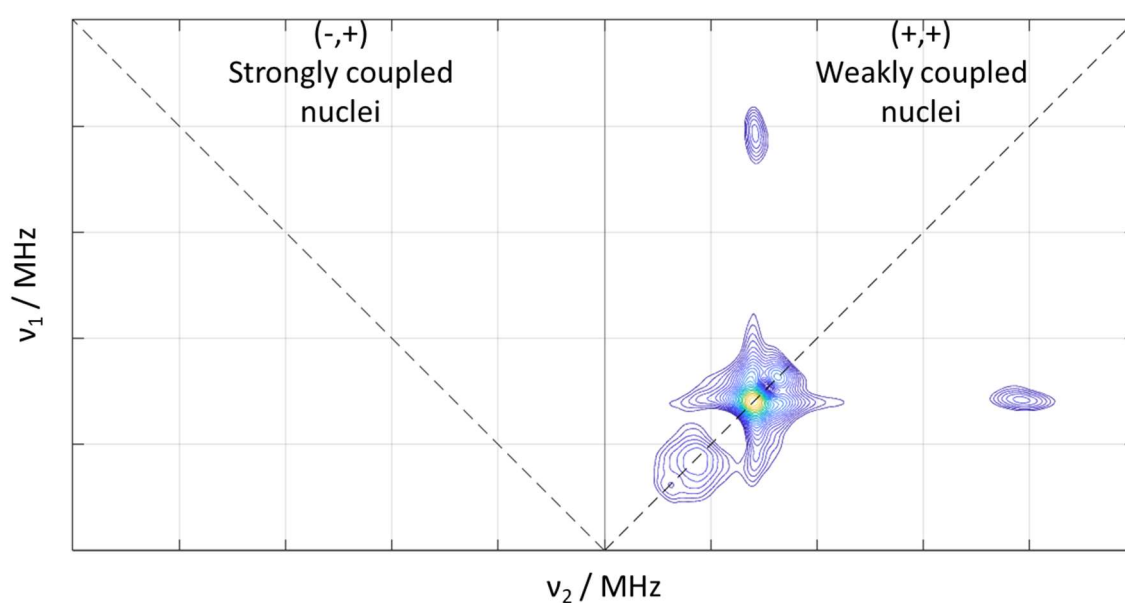
typically used. This method requires a pulse sequence where pulses of fixed microwave frequency to irradiate the sample separated by fixed and variable time intervals (i.e., where  $\tau$  is varied (top, **Figure 38**)). The pulse sequence results in a measurable spin echo (dashed red peak in **Figure 38**). The amplitude of this spin echo is modulated as a result of weakly coupled nuclei and returns a modulated time domain signal.<sup>113, 124</sup> Fourier transformation of this signal indirectly provides access to the both the NMR transitions and the SHFCs of weakly coupled nuclei.<sup>125</sup> In practise, the ESEEM experiment is rarely performed and is instead replaced with the more useful 2-dimension ESEEM technique, HYSCORE.<sup>124</sup> The HYSCORE technique is particularly sensitive to low frequency signals (<5 MHz) which is exceptionally useful given these couplings usually evade detection in the cw-EPR experiment.



**Figure 38. Visual representation of a two-pulse ESEEM sequence (top) and a 4-pulse HYSCORE sequence (bottom). In the ESEEM experiment,  $\tau$  is varied to return a modulated time domain signal. In the HYSCORE experiment,  $t_1$  and  $t_2$  are varied independently in order to obtain a two-dimensional data set.**

In the HYSCORE experiment, the inter-pulse delays,  $t_1$  and  $t_2$  (bottom, **Figure 38**), are varied independently to yield a 2-dimensional dataset. This technique has advantages over ESEEM when the studied paramagnet has more than one coupled nucleus. A Fourier transform of the modulated time decay data in both dimensions ( $t_1$  and  $t_2$ ) returns a 2-dimensional frequency domain spectrum with  $\nu_1$  on the y-axis and  $\nu_2$  on the x-axis (**Figure 39**).<sup>117, 126</sup> Coupled nuclei appear as cross-peaks, with weakly coupled nuclei ( $|a_{iso}| < 2|\nu_L|$ ) appearing in the (+,+) quadrant (right, **Figure 39**) and strongly coupled nuclei ( $|a_{iso}| > 2|\nu_L|$ ) appearing in the (-,+) quadrant (left, **Figure 39**) where  $\nu_L$  is the Larmor

frequency of the nucleus. The clear separation of these couplings allows for the accurate characterisation of complex spin systems with multiple coupled nuclei. The HYSCORE experiment also allows for the determination of the anisotropy of the dipolar hyperfine interaction ( $T$ ) in addition to the isotropic hyperfine interaction ( $a_{\text{iso}}$ ). Anisotropy in the dipolar hyperfine interaction leads to the correlation peaks to broaden, forming ridges (as seen by the contour lines in **Figure 39**). The maximum curvature of these ridges (with respect to the antidiagonal) relates to the magnitude of  $T$  and the value of  $a_{\text{iso}}$  can be ascertained from the ridge end points. Careful simulation of these data by a skilled EPR practitioner allows for the retrieval of accurate SHFC data, including the specific isotropic (Fermi) and anisotropic (dipolar) contributions to the spin Hamiltonian. For a more exhaustive description of HYSCORE spectroscopy, I direct the reader to Jeschke *et al.* and for a more general introduction, I suggest helpful texts by Roessler *et al.* and Murphy *et al.*<sup>113, 117, 124</sup>

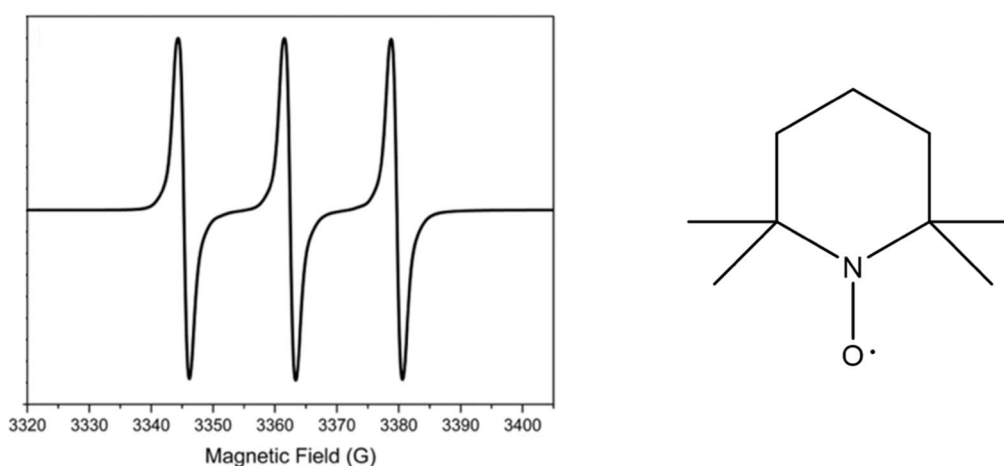


**Figure 39.** HYSCORE spectrum of an LPMO enzyme (*B/AA10*) collected at 3385 Gauss and  $\tau=136$  ns. Weakly coupled nuclei appear present in the  $(+,+)$  quadrant (right) and no strongly coupled nuclei appear in the  $(-,+)$  quadrant (left). Dataset included in a paper by Courtade *et al.*<sup>126</sup>

### 1.7.2.3 Simulation of EPR Spectra

In order to accurately retrieve the spin-Hamiltonian (SH) parameters from an EPR experiment, its spectra must be simulated. Simulations are performed by inputting a set of SH parameters and generating a simulated spectrum using a toolbox such as EasySpin, implemented by MATLAB.<sup>127</sup> This is an iterative process whereby the practitioner sequentially alters the parameters until the agreement

between the experimental and simulated spectrum is sufficiently good. For simple systems (such as a TEMPO, **Figure 40**), this iterative simulation process can be performed computationally. This is advantageous as it can be fast and removes human error/bias. However, for more complex systems involving heavier elements and multiple coupled nuclei (such as copper complexes), the computational expense is typically too large to obtain a suitable simulation using this automated method. As such, simulations of Cu(II)-LPMO EPR spectra are best by a skilled practitioner to manually adjust the SH parameters to best fit the experimental spectrum.<sup>104</sup> For example, if the simulated set of peaks appear too far downfield, the  $g$ -values should be increased to better match with the experiment. The  $g$ -factors have an inversely proportional relationship to the field position (i.e., if the resonant peak position is at a lower magnetic field than that of the simulation, the  $g$ -factor must be increased and *vice versa*).

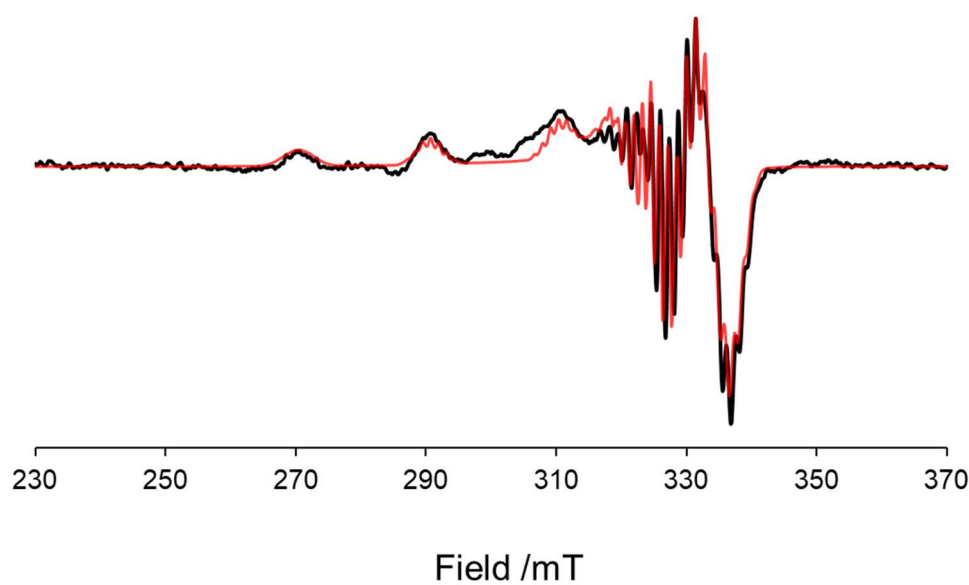


**Figure 40. EPR spectrum of TEMPO (left) and skeletal structure (right). Figure adapted from Mendoza *et al.*<sup>128</sup>**

The TEMPO radical is a relatively straightforward case for an EPR practitioner to simulate. The unpaired electron resides on a light atom (O), resulting in a  $g$ -value that is close to that of the free electron,  $g_e$  (2.0023). The TEMPO EPR spectrum displays hyperfine coupling arising from coupling of the unpaired electron with the  $I = 1$   $^{14}\text{N}$  nucleus; giving rise to three peaks separated by a hyperfine coupling constant,  $A$ . When simulating this HFC, the practitioner will need to input the identity of the coupled nucleus into script (so that the correct number of peaks are calculated) as well as the magnitude of the hyperfine coupling tensor. If the separation between the peaks does not match the experimental spectrum, these values should be increased and decreased accordingly.

In LPMO EPR spectroscopy, the number of spin-Hamiltonian parameters required to adequately describe the spectra can be large owing to the  $I = 3/2$   $^{63/65}\text{Cu}$  nuclei and the three

magnetically coupled  $I = 1$   $^{14}\text{N}$  nuclei. This is further complicated in anisotropic experiments, whereby three individual parameters are required to describe both the  $g$  and  $A$  values to account for the individual  $x$ ,  $y$ , and  $z$  components. In the case of anisotropic EPR experiments of Cu(II)-LPMOs, the multiline EPR spectra demand the careful curation of many spin Hamiltonian parameters in order to properly describe the system (**Figure 41**).



**Figure 41. X-band EPR spectrum (black) and simulation (red) of BaAA10 (0.3 mM) with  $\beta$ -chitin substrate in MES buffer (20 mM pH 6.0). Data collected at 150 K.**

The vast majority of EPR experiments on Cu(II)-LPMO complexes have been performed using X-band spectrometers (*ca.* 9.3 MHz). However, meaningful additional detail can be ascertained using lower (S-band, *ca.* 3-4 MHz) and greater (Q-band, *ca.* 34-35 MHz) microwave frequencies. The former can allow for better quantitation of the number of coordinating ligand atoms and better determination of the SHFCs and the latter allows for greater separation of the anisotropic  $g$  and  $A$  tensors.<sup>49, 111</sup> Most families LPMOs show approximately axial EPR spectra shown by the individual  $x$ ,  $y$ , and  $z$  values of the  $g$ -matrix and hyperfine tensor ( $g_z > g_y \approx g_x$  and  $A_z \gg A_y \approx A_x$ ) (**Table 2**). The spectra are indicative of a well separated  $|x^2 - y^2\rangle$  ground state with typical  $g_z$  and  $A_z$  values of 2.26 and 500 MHz, respectively<sup>49, 66</sup>

However, *chitin* active members of the AA10 family have shown to have a different electronic structure compared to the resting states of the other families of LPMOs. The chitin active AA10s still show an overall  $|x^2 - y^2\rangle$  electronic ground state but with some notable differences. Their EPR spectra display a reduced  $A_z$  copper hyperfine coupling and an overall rhombic  $g$ -matrix and hyperfine tensor ( $g_z > g_y > g_x$  and  $A_z \approx A_x > A_y$ ).<sup>63, 126, 129</sup> With the aid of crystallographic and computational

studies, this alteration in electronic structure has been shown to arise from two coordinating water ligands which sit outside of the basal plane of the histidine brace.<sup>63, 126, 129</sup> The geometry resembles a distorted square pyramid and leads to a suppression in  $d(x^2-y^2)$  character in the SOMO and a corresponding increase in  $dz^2$  mixing. The electronic ground state of the Cu(II) ion in chitin active AA10s therefore contains a reduced fraction of  $d(x^2-y^2)$  and contains a greater mixture of the other d-orbitals, notably  $d(z^2)$ . The resulting SH parameters from the chitin active AA10s are very different from the other families (highlighted in **Table 2**).

**Table 2. EPR spin Hamiltonian parameters of LPMO enzymes characterised in the literature from the different auxiliary activity (AA) families.**<sup>49, 50, 52, 97, 99, 126, 130, 131</sup>

Enzyme <sup>†</sup>	<i>g</i> -matrix			Cu Hyperfine tensor / MHz		
	<i>g<sub>z</sub></i>	<i>g<sub>y</sub></i>	<i>g<sub>x</sub></i>	<i>A<sub>z</sub></i>	<i>A<sub>y</sub></i>	<i>A<sub>x</sub></i>
<i>LsAA9</i> <sup>130</sup>	2.23	2.06	2.04	460	75	75
<i>TtAA10A</i> <sup>131</sup>	2.27	2.09	2.05	420	125	100
<i>BIAA10A</i> <sup>126</sup>	<b>2.26</b>	<b>2.10</b>	<b>2.03</b>	<b>336</b>	<b>110</b>	<b>255</b>
<i>AoAA11</i> <sup>97</sup>	2.28	2.10	2.03	440	48	34
<i>AnAA13</i> <sup>99</sup>	2.26	2.08	2.05	513	98	78
<i>PcAA14A</i> <sup>49</sup>	2.27	2.07	2.06	508	55	28
<i>TdAA15A</i> <sup>50</sup>	2.25	2.07	2.04	525	60	47
<i>PiAA17</i> <sup>52</sup>	2.27	2.07	2.04	488	68	40

<sup>†</sup>The AA16 family have not yet been satisfactorily characterised with EPR spectroscopy

The combination of spectroscopic, crystallographic and computational studies have led to the current understanding that LPMOs coordinate exogenous ligands (either H<sub>2</sub>O/OH<sup>-</sup>/Cl<sup>-</sup>) adopting a 4/5-coordinate geometry in the Cu(II) resting state (of which the precise ligand field is specific to the individual protein families) and a 3-coordinate Cu(I) geometry in the activated state; lacking the recruitment of additional ligands. The Cu(I) state of the enzyme is primed to react with either O<sub>2</sub> or H<sub>2</sub>O<sub>2</sub> leading to the reactive oxygen species (ROS) necessary to oxygenate the aliphatic C-H bonds in a variety of recalcitrant polysaccharides. Multiple studies have shown certain LPMO Cu(II) ligand fields to be influenced by the presence or absence of polysaccharide substrate.<sup>66, 126, 129, 132, 133</sup> These studies have gathered significant interest in recent years as they show that the binding of substrate to enzyme can lead to marked alterations in the electronic structure which may be important in oxygen activation and subsequent catalysis.<sup>104</sup> Typically, these EPR studies have shown the substrate-bound case to have an elevated *A<sub>z</sub>* copper hyperfine coupling and a reduced *g<sub>z</sub>* value.<sup>129, 133</sup> These changes to the SH parameters are indicative of a greater energetic gap between the  $|x^2 - y^2\rangle$  ground state and  $|xy\rangle$  excited state, suggesting a stronger ligand field primarily in the x/y plane. In some cases, it has been

shown that this can occur by the displacement of (pseudo-)axial water molecules after binding of the polysaccharide substrate.<sup>126, 129</sup>

Whilst the understanding of both structural and electronic features of the LPMO active site have grown following the various EPR and crystallographic studies, additional methods are sometimes needed to better relate spectroscopic properties with geometrical structure. As such, many research groups apply computational approaches to augment these studies.

## 1.8 Computational and Theoretical Studies

Many research groups perform computational/theoretical studies to augment their spectroscopic and crystallographic experiments to provide additional information and insight.<sup>126, 129, 134</sup> There are several major quantum mechanical software packages which are routinely used for this purpose, such as Gaussian, ORCA, and Turbomole.<sup>135-138</sup> Using these packages, a range of computational methods can be employed to obtain the optimized coordinates of a molecule/molecular cluster, total energies, a wealth of spectroscopic properties, and the energetic barriers of chemical reactions. For small molecules, calculations of this type can be routinely performed on desktop computer without too much computational expense (in both equipment and time). However, for larger molecules (particularly with those with heavier elements or open shell electronic configurations) the use of a dedicated computational cluster is often required.

There are a variety of theoretical methods that computational chemists use to examine chemical systems such as Hartree-Fock (HF) or Density Functional Theory (DFT), the latter usually providing the best balance of accuracy and computational cost.<sup>139</sup> Occasionally, more advanced, multiconfigurational approaches such as Complete Active Space Self-Consistent Field theory (CASSCF) are necessary to properly describe open shell molecules for which the previous two methods can sometimes fall short.<sup>140</sup> The appropriateness of both the method and the level of theory will vary on the studied chemical system and needs to be decided and explored by an experienced computational chemist. If little is known about the suitability of computational methods for the studied system, a wide range of approaches should be performed to see how they compare, and which methods can best reproduce the observable characteristics.



## 1.8.1 Level of Theory - Functionals and basis sets

Once decided on a computational approach, a suitable functional (i.e., a function of another function) must be chosen to execute the desired calculation. In the case of density functional theory, a particular functional is selected which is used to calculate the energy of the system by evaluating the spatial distribution of electron density.<sup>141</sup> The first widely recognised type of DFT functional were the “local density approximation (LDA)” functionals which were proposed by Hohenberg and Kohn in their ground-breaking DFT paper published in 1964.<sup>142</sup> These functionals approximate the energy of the true electron density using a local constant density based on that of uniform electron gas.<sup>143</sup> These are the oldest and most simplistic DFT functionals and were developed to describe the ground states of homogenous gases. Their applicability to molecules and more complex systems is limited, though they have shown some success in describing the physical properties of some solid-state materials.<sup>144</sup> These functionals tend to overestimate binding energies and are generally poor at evaluating properties such as hydrogen bonding.<sup>144</sup> As such, the local density approximations were improved upon by including density gradients leading to an improved set of functionals, known as the generalised gradient approximation (GGA) functionals.<sup>143, 145</sup> These approximations were first developed by Perdew, Burke, and Ernzerhof in 1998 and showed considerably improved calculation results of molecules.<sup>143</sup> Functionals of this type (PBE) are still widely used today in applications including describing the spectroscopic properties of LPMOs.<sup>134</sup> GGA functionals such as ‘BP86’ tend to be among the fastest options to perform optimisations of coordinates with reliable final geometries.<sup>135</sup> The final major type of DFT functional are the hybrid functionals, the major development of which came about in 1988 by Lee, Yang, and Parr, and later in 1993 by Becke.<sup>146, 147</sup> The exchange and correlation energy approximations developed by these groups has led to the most commonly used DFT functional which is named after their initials, B3LYP. These DFT functionals incorporate a percentage of exact exchange derived from Hartree-Fock (HF) theory into the density functional, forming a *hybrid* functional. The degree of HF exchange included in the approximation depends on the functional, with common ones being B3LYP (20%) and PBE0 (25%).<sup>145-147</sup> These hybrid functionals have often shown to give even more accurate calculations, especially when predicting properties such as exact total energies, spectroscopic properties and vibrational frequencies.

Along with the functional, the level of theory in calculation is also dictated by the *basis sets*. Basis sets are an additional collection of functions that are applied to describe to the orbitals of the individual atoms in a studied system. These basis sets are assigned by the practitioner and can be specifically tailored to add a greater depth of theory to specific atoms which benefit from an increased

level of accuracy. A basis set which contains twice as many basis functions as the minimal basis set is referred to as a “double zeta basis set” and basis sets which contains triple or quadruple the number of basis functions are known as “triple” or “quadruple”-zeta basis sets, respectively. Heavier basis sets (those which apply a greater number of functions to describe the molecular orbitals) provide better accuracy in a given calculation but come with additional computational cost. For the majority of atoms in a given LPMO active site model, a double zeta or split valence basis set (e.g. def2-SVP) is usually sufficient to obtain the desired accuracy with reasonable computational expense. However, it is necessary to use a larger basis set (such as a triple-zeta, def2-TZVP) to in order to correctly describe the central copper ion and its coordinating atoms. Using a too small basis set on the copper ion and coordinating atoms could result in inaccurate bond lengths and lead to unreliable downstream property calculations. In such instances, these calculations are referred to as ‘basis set limited’, whereby increased basis set sizes lead to greatly improved accuracy of the calculated properties. In addition to functional and basis sets, leading computational chemistry experts have developed approaches to approximate the effects of both dispersion and solvation; sometimes playing a pivotal role in final geometries, energies, and spectroscopic properties.<sup>148, 149</sup>

## 1.8.2 Geometry Optimizations

In computational chemistry, a practitioner requires a set of internal coordinates denoting the 3D atomic positions of their studied system to calculate desired energies and properties. One of the best methods of acquiring atomic coordinates for the basis of computational calculations is to use those from crystal structures. This is a desirable means of acquiring initial coordinates as they are based off experimental evidence. Protein crystals are obtained by screening purified protein solutions with an array of conditions to encourage crystal growth. For suitably large and ordered crystals, X-Ray Diffraction (XRD) can be performed to obtain a scatter pattern which can later be solved for crystallographic atomic coordinates.<sup>150</sup> Whilst the resolution for protein x-ray crystallography can be reasonably good (ca. 1-2 Å) and getting ever better, the experimental coordinates are often not suitably accurate for the meaningful calculations of energies and spectroscopic properties. As such, the initial atomic coordinates are usually “geometry optimized”. Geometry optimizations are a means of adjusting the given geometry of a molecule or molecular cluster to find atomic coordinates that minimize the total energy. The resulting optimized coordinates can be valuable for the practitioner, providing the theoretically favourable bond lengths and angles of their desired system which otherwise may have been unreliably represented by crystallography. The reliability of DFT optimized

geometries vs crystallographic coordinates will depend on the quality of the computations performed vs the resolution of the x-ray crystal structures.

Several different theoretical approaches can be employed to achieve these optimized coordinates such as Hartree-Fock (HF), Molecular Dynamics (MD) and Density Functional Theory (DFT); with the latter being by far the most popular. To perform a geometry optimization, an input file script must be written which specifies the level of theory, the desired calculation, and the initial geometric coordinates to be optimized. The level of theory is principally specified by selecting the desired functional. For LPMOs, the full size of the protein is too large to calculate using a purely quantum mechanical approach. As such, cluster models are often used to describe the key parts of the studied structure.<sup>129, 151, 152</sup> In the case of LPMOs, this would usually include the central copper ion surrounded by just the immediate ligands and key supporting residues. In order to account for the structural constraints that are usually enforced by the rest of the protein, it is possible to keep certain atoms frozen throughout the optimization of cluster models. If all atomic coordinates of the protein are desired in the calculation, a combined quantum mechanics-molecular mechanics (QM-MM) approach is necessary. QM-MM treats the majority of atoms using molecular mechanics (MM) and only a small selection are treated with the quantum mechanical (QM) approach.<sup>90, 151</sup> An example of a DFT geometry optimisation script for a superoxide molecule can be seen below in **Figure 42**.

Selecting the DFT functional

Selecting the basis set

Command for geometry optimization

```
! BP86 def2-SVP Opt

%geom

*xyz -1 2
0 -0.526979000 0.826667000 0.000000000
0 0.833646000 0.826667000 0.000000000

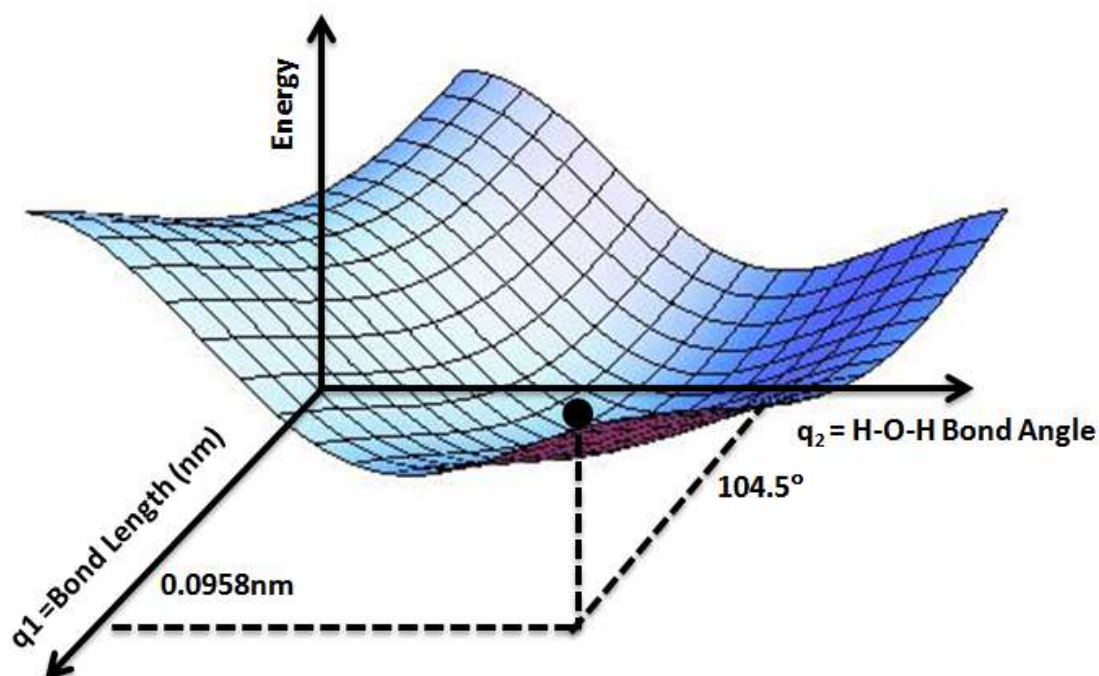
end
end
```

Charge and spin multiplicity

xyz coordinates of molecule

**Figure 42.** Example script for a DFT geometry optimization of a superoxide ion ( $O_2^-$ ) using ORCA.

When a geometry is successfully optimized, it should lie at a minimum on the potential energy surface (PES); whereby any geometric adjustments lead to an increase in the total energy of the system. To check that a minimum has been reached on the PES; a computational chemistry practitioner will often perform a *frequency* calculation on the final coordinates. A frequency calculation calculates the thermodynamic parameters of the system (S, H, and G) as well as the IR vibrational spectra. A geometry that is not optimized to an energy minimum will have IR stretching frequencies with significant negative values. A negative value means moving the geometry of the 'optimized' coordinates leads to an exothermic release of energy. In the case that large, negative stretching frequencies are obtained the geometry should be re-optimized to find the true minimum. For cluster models, performing such frequency calculations are more challenging to interpret (and often trivial) due to the stretching between atoms that were fixed throughout the calculation returning large negative values. Stretching modes between such atoms would be impossible when accounting for the full size of the protein rendering this step unnecessary. A potential energy surface can be generated by plotting the displacement of atoms against the total energy of the system. For instance, a simple case of plotting the O-H bond length and  $\angle$  H-O-H bond angle in water against the energy of the system leads to the PES shown in **Figure 43**. In this case, the global minimum of the PES is found at an O-H bond length of 0.958 Å and a  $\angle$  H-O-H bond angle of 104.5°. From here, any adjustment to the geometric coordinates leads to an increase in the total energy.



**Figure 43.** Potential energy surface of water showing the total energy of the system as a function of O-H distance and the  $\angle$  H-O-H bond angle. Image free for public use from [www.wikipedia.org](http://www.wikipedia.org)

### 1.8.3 EPR Property Calculations

Once found, the optimized coordinates of computational models can be used for property calculations, whereby many spectroscopic properties can be calculated. One of the most useful of these in the LPMO field are EPR calculations. An EPR property calculation can be performed for a given set of coordinates to return valuable information including the  $g$  and  $A$  values. Moreover, DFT calculations can return the specific contributions to  $g$ -values and hyperfine couplings including the dipolar and Fermi interactions, providing additional insight. Again, a suitable functional and basis set regime must be specified (hybrid DFT functionals such as B3LYP are popular for the prediction of copper EPR spectra) to return meaningful results. A single point calculation is usually performed whereby the total energy as well as all of the atomic orbitals are calculated using the desired functional/basis set regime. From here, a command at the end of the script such as 'eprnmr' in ORCA, will trigger the calculation of the EPR properties of the system. It is important to evaluate the effect of increasing the basis set depth as well as different functionals to make sure the data obtained from such calculations are reliable.

```
! UKS B3LYP Def2-TZVP

* xyz 0 2 _____ Charge and spin multiplicity
[ ] _____ Space for xyz coordinates

%eprnmr _____ Requesting EPR property calculation
gtensor 1
ori Centerofelcharge } Calculate  $g$  from the centre of electronic charge
nuclei = all N {aiso, adip, fgrad, rho} } Calculate hyperfine constants for N and Cu atoms
nuclei = all Cu {aiso, adip, aorb}
end
```

Figure 44. Example script for a DFT EPR calculation for a neutral, doublet radical using ORCA.

There have been some comprehensive studies assessing the reliability of various functionals and basis set regimes for assessing the EPR properties of metal complexes using DFT. I direct the reader to a paper by Frank Neese, the developer of one of the major quantum chemistry software packages, ORCA, in which he details some of the successes and shortcomings of calculating EPR properties of metalloproteins and complexes using DFT.<sup>153</sup> In this paper, he highlights the challenges in calculating the EPR properties of transition metals, with the results often tending to be qualitative guides than quantitative. They did, however, note that the hybrid functionals such as B3LYP tend to perform better at reproducing experimental EPR parameters than those from the GGA functionals (e.g. BP86).

A detailed paper by the Solomon group identified the tendency for DFT functionals to overestimate metal-ligand covalency, leading to unreliable  $g$ -factors and hyperfine couplings.<sup>154</sup> In their paper, they used a systematic approach to assess the viability of a range of functional/basis set regimes to describe the well-studied  $\text{CuCl}_4^{2-}$  complex as a model system to benchmark the calculations.<sup>154</sup> They demonstrated that two of the most common DFT functionals, BP86 and B3LYP, grossly overestimated the covalency of the metal ion with the chloride ligands, leading to a reduced spin density on copper of around 40% (compared to the experimental value of 62%).<sup>154</sup> To combat this over-covalency issue, they trialled increasing the degree Hartree-Fock exchange (HFX) ratio in the functional to increase the ionicity of the bonding behaviour. In this study, they found 38% HFX to be the optimised value to introduce into the density functional in order to reproduce the metal-ligand behaviour of a model copper (II) complex. The corrected balance of covalency/ionicity of the bonding led to a more accurate depiction of Cu-L bond lengths and spin densities and thus more accurate downstream calculations of  $g$ -values and hyperfine couplings. Such adjustments to the HFX in DFT functionals have been used to good effect in describing the spin Hamiltonian parameters of LPMOs.<sup>126,</sup>

130

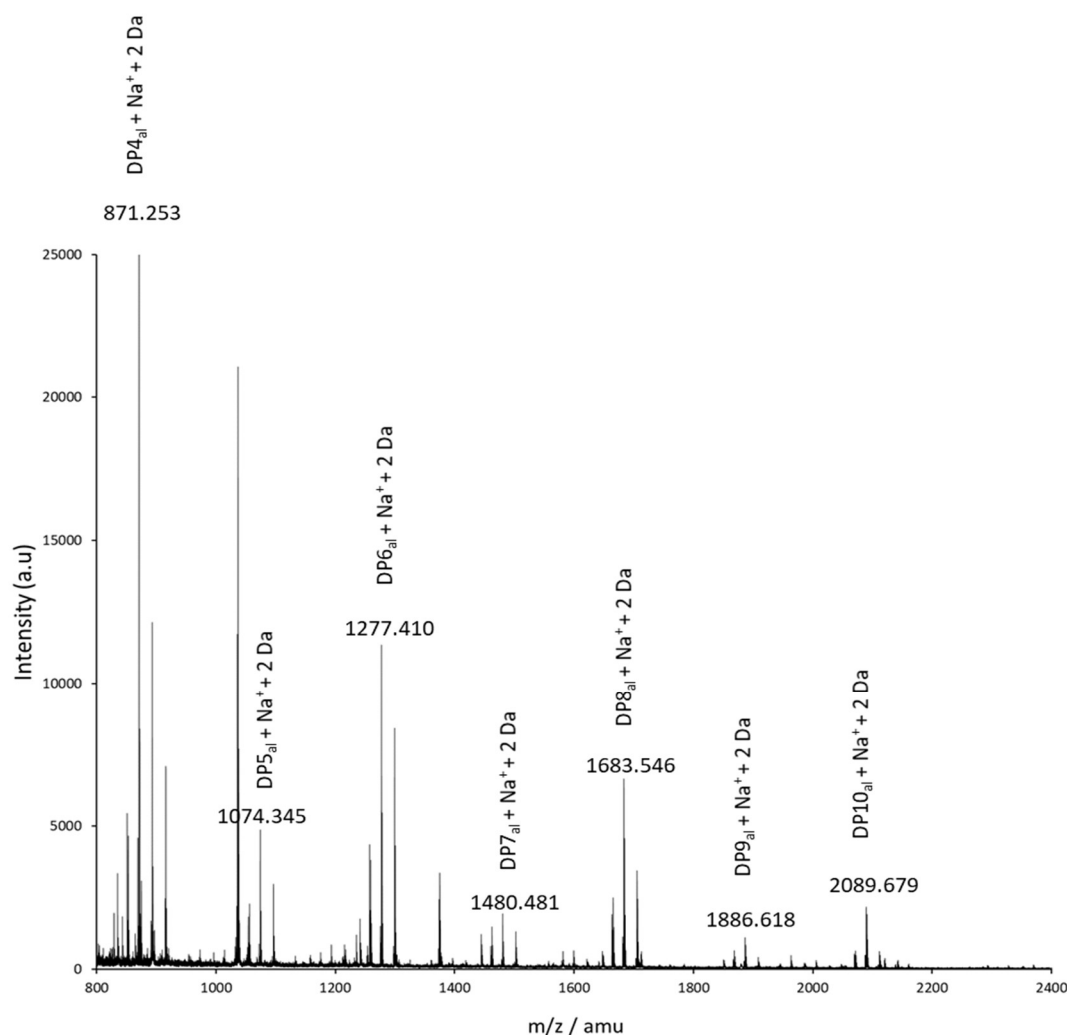
In relation to LPMOs, there have been some dedicated studies trialling functionals and basis set regimes in order to best reproduce the experimental spin Hamiltonian parameters of the copper active site.<sup>134, 155</sup> Theibich *et al.* performed a thorough investigation into the effect of the computational set up on the final copper and ligand hyperfine couplings.<sup>134</sup> They found the DFT functionals to overestimate the SHFCs and that the hyperfine couplings in general were a parameter that is highly sensitive to the functional and basis set choice. They concluded that meaningful hyperfine couplings for LPMOs could be obtained using DFT, though they noted the importance of using an expanded 'core-properties' basis set to describe the copper ion or the dipolar and Fermi contact interactions will carry significant error.<sup>134</sup>

There are some good examples in the literature of DFT calculations making meaningful contributions to the study of LPMO EPR parameters. A study by Bissaro *et al.* in 2018 saw electronic perturbations to the active site of a chitin-active AA10, *SmAA10A*, following the addition of chitin. Both the substrate bound state and 'resting state' of the enzyme were qualitatively described using DFT calculations with suitable agreement.<sup>129</sup> Another study by Ciano *et al.* in 2020 saw the use of DFT EPR calculations to describe the *g*-matrix and copper hyperfine tensor of a fungal active AA9 enzyme, *LsAA9*. This is a fascinating study investigating the orientation dependence of the EPR spectra of an LPMO upon binding a highly ordered, crystalline substrate (cellulose fibres from celery).<sup>130</sup> These computations showed reasonable agreement between the calculated spin Hamiltonian parameters the experimentally derived values, though not without a significant degree of error, a known issue for Cu(II) property calculations.<sup>156</sup> In this study, however, the major contribution of the DFT calculations was to aid in the determination of the orientation of the *g*-matrix with respect to the copper active site coordination geometry, which found the x-direction of the *g* matrix to be aligned to N(His)-Cu-N(His) direction, the y-direction to be aligned with the N(NH<sub>2</sub>)-Cu-X direction, with the z-direction being orthogonal to the other two.<sup>130</sup> This is an example of the additional insight that can be gained by performing DFT EPR calculations that would be challenging to determine experimentally.

## 1.9 On the LPMO Catalytic Mechanism

Since their initial discovery in 2010, much has been learned about the spectroscopic properties, structural characteristics, and product profiles of LPMO reactions which have helped towards understanding their overall catalytic mechanism.<sup>45, 52, 54, 65, 66</sup> It is generally accepted that in order to perform oxygen activation, the Cu(II) resting state of the active site must accept an electron from an external electron donor to form a reduced, Cu(I) species. This reduced form of the enzyme has been shown by both crystallography and DFT calculations to lose its exogenous ligands (OH<sup>-</sup>/H<sub>2</sub>O/Cl<sup>-</sup>). The 3-coordinate Cu(I) complex ligated by just the histidine brace possesses a vacant equatorial binding site that is electronically primed for the binding and activation of an oxidative co-substrate. The original 2010 study by Vaaje-Kolstad *et al.* provided the foundation of understanding for the LPMO mechanism as they showed that molecular oxygen can be the source of the oxygen in the hydroxylated substrate.<sup>45</sup> They were able to confirm this by performing activity assays with a chitin active AA10, using either isotopically labelled <sup>18</sup>O<sub>2</sub> or H<sub>2</sub><sup>18</sup>O in the experiments.<sup>45</sup> Similar studies were performed a few years later confirming monooxygenase activity for the AA9 family of LPMOs.<sup>54, 56</sup> Isotopic labelling studies of this type have shown the identification of heavy oxidised oligomers using

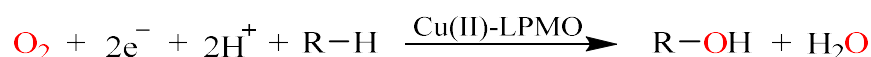
Matrix Assisted Laser Desorption Ionisation Time Of Flight (MALDI-TOF) mass spectrometry. We have performed experiments of this type on an LPMO from *Aspergillus Oryzae* (AoAA11), demonstrating monooxygenase activity also for the AA11 family. An AoAA11 protein sample (1  $\mu$ M, generously supplied by Dr. Alessandro Paradisi) was used for reactions with cellulose (1 mg), ascorbic acid (1 mM) and  $^{18}\text{O}_2$  (1 atm). These reactions led to a series of heavy (+2 Da), oxidised, breakdown products characterised using MALDI-TOF mass spectrometry (**Figure 45**).



**Figure 45.** MALDI-TOF spectrum of  $\beta$ -chitin breakdown products performed by AoAA11 (1  $\mu$ M), Na ascorbate (1 mM) and  $^{18}\text{O}_2$  (1 atm). The highlighted masses correspond to sodiated aldonic acid (al) chitooligomers with varying degrees of polymerisation (DP<sub>n</sub>). The '+ 2 Da' is due to the presence of a single  $^{18}\text{O}$  atom in the products.

The findings of these studies allowed for the proposal of a general equation of the LPMO oxygenation reaction as follows:





Following binding and activation of O<sub>2</sub> at the LPMO active site, the logical product is a Cu(II)-superoxide species. However, it has been hypothesised by some that this intermediate may not be a potent enough oxidiser to break the polysaccharide C-H bond in cellulose or chitin, with the energetic barrier for hydrogen atom abstraction (HAA) predicted to be endothermic by DFT.<sup>93, 157</sup> The nature of the LPMO catalytically active species remains speculative due to the lack of experimental characterisation. As such, their investigation has relied largely on computational modelling to determine the theoretical barrier heights for the corresponding reactions. It has been suggested by some groups that copper complexes are capable of forming reactive [Cu(III)-OH]<sup>2+</sup> intermediates which can perform the difficult HAA step.<sup>158</sup> Much more commonly considered to the relevant catalytic species is the Cu(II)-oxyl (Cu-O<sup>•</sup>).<sup>90, 93, 159</sup> Both the [Cu(III)-OH]<sup>2+</sup> and [Cu(II)-O]<sup>•</sup> intermediates are predicted to lead to exothermic HAA from polysaccharide substrates, indicating they are more potent oxidisers than the Cu(II)-superoxide intermediate.<sup>157</sup>

### 1.9.1 Oxygenase vs Peroxygenase Mechanism

As well as the identifying the relevant reactive oxygen species (ROS), another key question that remains is the identity of the ‘correct’ oxidising co-substrate, with there being substantial evidence for both O<sub>2</sub> and H<sub>2</sub>O<sub>2</sub>.<sup>160-162</sup> Activity studies have shown that both oxidising co-substrates can turnover oxidised oligomers from a polysaccharide substrate.<sup>63, 163</sup> Molecular oxygen, O<sub>2</sub>, is the obvious consideration as the natural co-substrate for LPMO activity due to its omnipresence in the earth’s atmosphere. O<sub>2</sub> has been relied upon for the vast majority of LPMO activity assays as the oxidising equivalent and was first performed in 2010 by Vaaje-Kolstad *et al.*<sup>45</sup> As discussed above, it have been shown through isotopic labelling studies that the oxygen atom from O<sub>2</sub> can be the source of oxygen in the oxidised products. In the same study, they also showed that in the absence of polysaccharide substrate, uncoupled turnover from the reaction of Cu(I)-LPMO with O<sub>2</sub> led to the in-situ generation of H<sub>2</sub>O<sub>2</sub>. DFT studies have also shown that O<sub>2</sub> can be readily converted into H<sub>2</sub>O<sub>2</sub> at the active site of LPMOs *via* two proton-coupled electron transfer (PCET) steps, rendering this debate slightly mute.<sup>45, 164</sup>

Despite swathes of evidence showing oxygen’s efficacy as an oxidative partner, numerous studies have suggested that H<sub>2</sub>O<sub>2</sub> is *preferred* as an oxidative co-substrate in LPMO catalysis.<sup>161, 163</sup> In many of these reports, they claim that the production rate of oxidised oligosaccharides is much improved (100 times) compared to the corresponding reactions with O<sub>2</sub>.<sup>163</sup> Whilst these studies have

shown boosted reaction rates by supplying the reaction mixture with H<sub>2</sub>O<sub>2</sub>, it is still unclear whether it is correct or indeed preferred as a co-substrate for the LPMO reaction. As extracellular enzymes, a reliance on a supply of H<sub>2</sub>O<sub>2</sub> seems impractical, whereas in almost all circumstances, O<sub>2</sub> will be available to the enzyme. There are many examples in enzymology of monooxygenase enzymes also possessing peroxygenase activity, however, it is scarcely seen for peroxygenases to also show monooxygenase activity.<sup>165, 166</sup>

Despite being suitably oxidising, O<sub>2</sub> reduction typically experiences considerable kinetic challenges due to the of breaking the spin selection rule:  $\Delta S = 0$  (by changing from the triplet state in O<sub>2</sub> to a singlet state in the products).<sup>167</sup> Transition metal complexes are usually recruited to reduce the kinetic barriers of O<sub>2</sub> reduction (*via* SOC) in order for these reactions to become feasible. In the case of hydrogen peroxide, these major kinetic barriers have already been overcome as the molecule possesses both the protons and electrons necessary for facile substrate oxidation. The elimination of the slow O<sub>2</sub> activation steps is likely the principal cause for the boosted rates of reaction observed when using hydrogen peroxide in activity assays. Along with this boosted activity, H<sub>2</sub>O<sub>2</sub> been shown to increase the rate of enzyme inactivation due to self-oxidation.<sup>168</sup> When H<sub>2</sub>O<sub>2</sub> concentrations are not scrupulously controlled, mass spectrometry analyses have reveal significant oxidative damage to the enzyme and in some cases the irreversible formation of inactive species such as the ‘purple species’ seen in *LsAA9*.<sup>152</sup> The outcome of these findings are that peroxide can be used to accelerate the rate of reaction within LPMOs, but the system loses an element of control leading to more deleterious off-pathway chemistry.

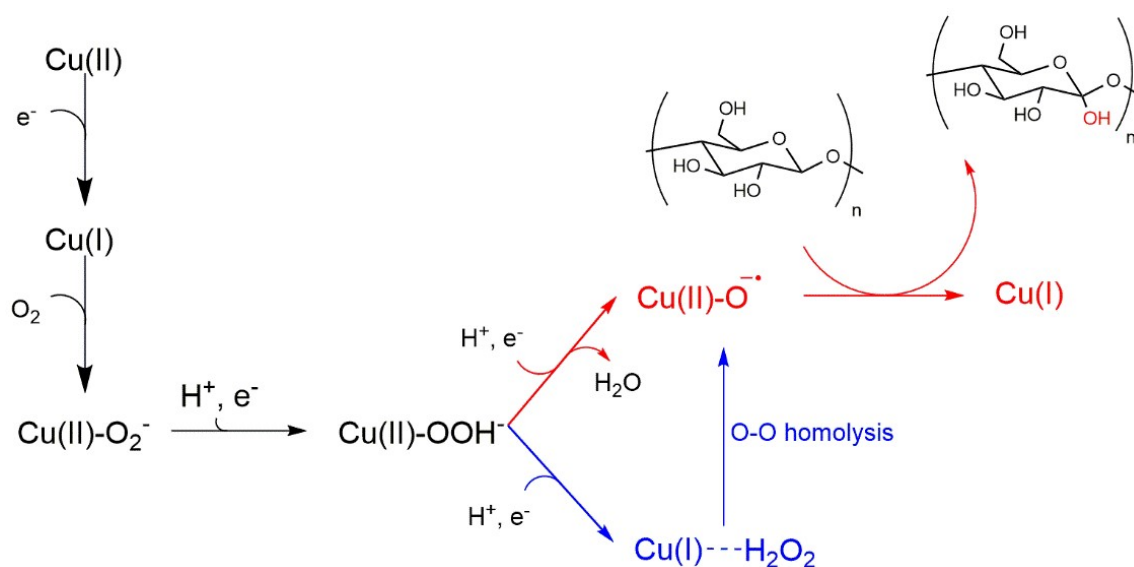
What emerges though, is the possibility of using H<sub>2</sub>O<sub>2</sub> as a shunt in the catalytic cycle; by removing the kinetically challenging O<sub>2</sub> activation steps leading to more rapid generation of reactive oxygen species. Such a shunt in the catalytic cycle has been reported for the generation of compound 1 (CMP1) in cytochrome P450 chemistry.<sup>169</sup> In the absence of characterised intermediates, computational modelling is necessary to better gauge the thermodynamics and feasibility of the proposed LPMO mechanistic pathways.

## 1.9.2 Computational Insights

As well as generating optimised coordinates and calculating EPR properties, computational studies have made significant contributions to understanding the LPMO catalytic mechanism.<sup>54, 90, 151</sup> Computational studies have primarily considered a reactive copper oxygen species (ROS) such as

Cu(II)-superoxo,  $[\text{Cu(II)-O}_2]^+$ , or Cu(II)-oxyl,  $[\text{Cu(II)-O}]^+$ , to be the likely candidates to perform the difficult hydrogen atom abstraction (HAA) the BDE of which is estimated to be around  $100 \text{ kcal mol}^{-1}$ . It is thought that the reduced copper ion can bind to and activate  $\text{O}_2$  leading to the formation of a  $[\text{LPMO-Cu(II)-O}_2]^+$  species. This copper(II)-superoxo species was originally considered to be the key reactive intermediate as the logical product of the reaction of Cu(I)-LPMOs with oxygen.<sup>54</sup>

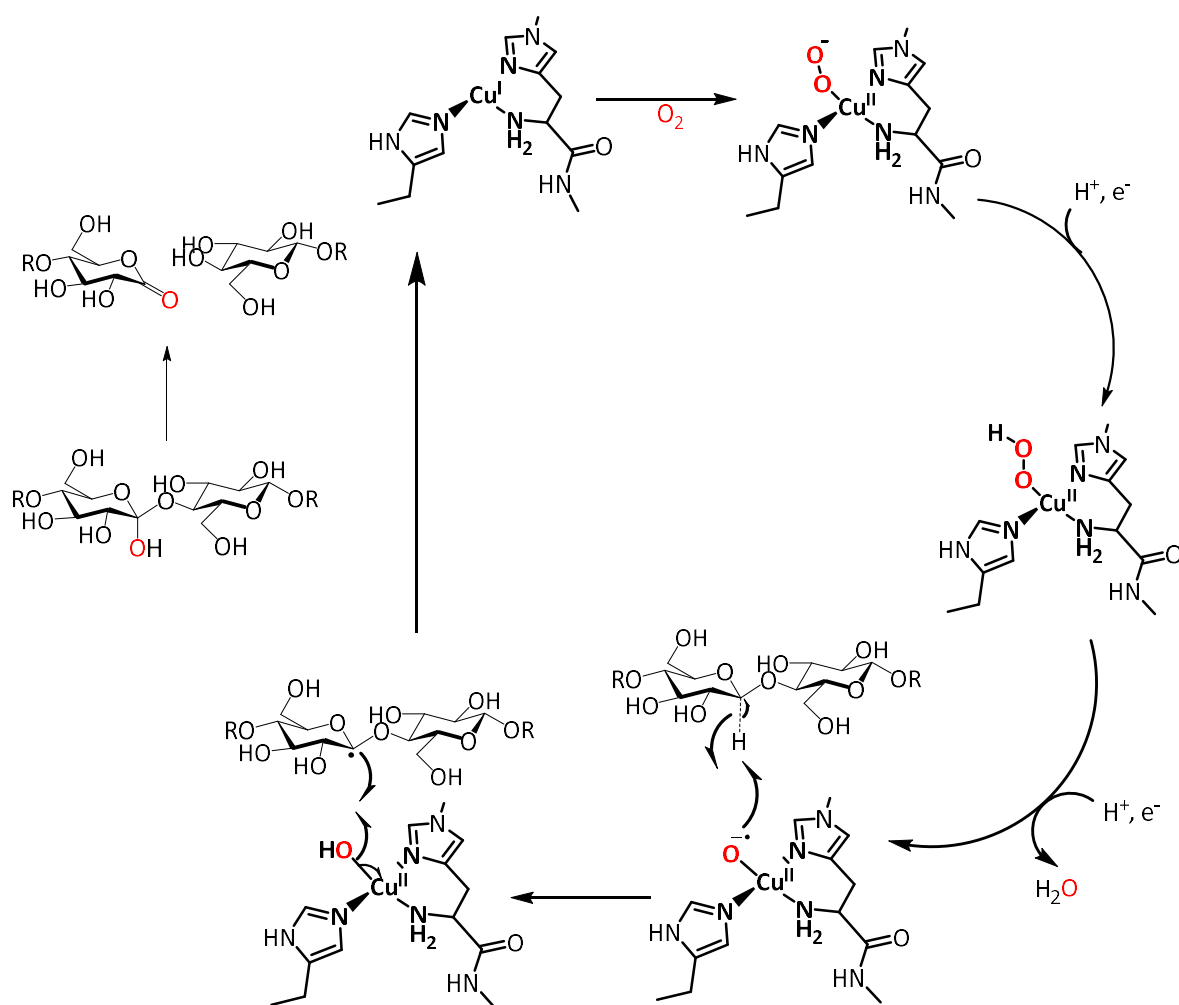
In 2014, Kim *et al.*, suggested that this oxidation will likely require a more potent ROS to perform overcome the HAA reaction barrier. They suggested a more potent Cu(II)-oxyl species is a more likely candidate involved in a proposed radical rebound mechanism.<sup>93</sup> In their study, they used DFT to calculate the geometries and energies for a number of proposed transition states and intermediates so that an overall reaction free-energy landscape could be drawn. They investigated the efficacy of a reactive copper-oxyl species positioned in the nominal *axial* position (opposite the tyrosine) oxidising a polysaccharide substrate. The computed intermediate showed sufficiently low energetic barriers to oxidise the strong C-H bond in cellulose. However, the formation of such an intermediate in the axial position is severely thermodynamically disfavoured with respect to the vacant equatorial position and is a severe oversight in these early studies. Subsequent studies have also since suggested the likely ROS to be a copper-oxyl species, though all of which positioned instead in the more thermodynamically accessible, equatorial position.<sup>90, 132, 159</sup>



**Figure 46. Proposed catalytic pathways of substrate hydroxylation for the  $\text{O}_2$ -dependent reaction of LPMOs. Figure based on mechanism proposed by Wang *et al.*<sup>164</sup>**

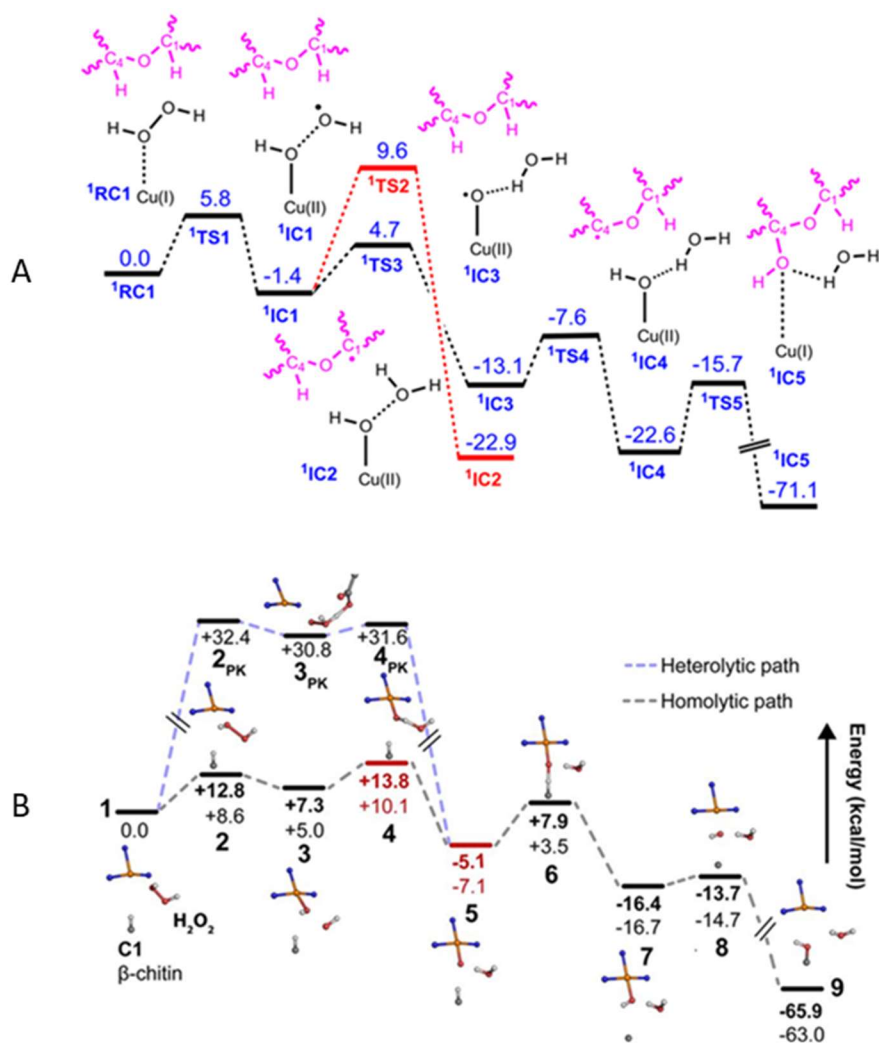
In the  $\text{O}_2$ -dependent catalytic cycle, it has been proposed the  $[\text{Cu(II)-OO}]^+$  intermediate first converts to the Cu(II)-hydroperoxo  $[\text{Cu(II)-OOH}]^+$  *via* a proton-coupled electron transfer (PCET) from a reducing co-substrate such as ascorbic acid.<sup>164</sup> The Cu(II)-hydroperoxo intermediate can then undergo

a second PCET step to form the copper(II)-oxyl and a molecule of water. Protonation of the distal atom (red in **Figure 46**) is expected to facilitate homolysis of the O-O bond and generate the copper(II)-oxyl along with a molecule of water.<sup>170</sup> More recently, it has been proposed that the PCET to the proximal oxygen atom (blue in **Figure 46**) has also been shown by DFT calculations to lead to formation of the copper(II)-oxyl via a Cu(I)-H<sub>2</sub>O<sub>2</sub> intermediate. It is thought this proceeds to form a [Cu(II)-OH]<sup>+</sup> species and a 'caged hydroxyl' radical which reclaims the hydrogen atom to yield the same proposed reactive [Cu(II)-O]<sup>+</sup> intermediate.<sup>90</sup> The O<sub>2</sub>-dependent pathway presented in **Figure 46** can be incorporated into a full reaction mechanism as shown in **Figure 47**. The key steps involve: O<sub>2</sub> binding and activation from the Cu(I)-LPMO (top), two PCET steps leading to formation of the Cu(II)-oxyl intermediate (right), radical rebound mechanism between the ROS and substrate (bottom), and hydroxylation of the substrate and regenerating the Cu(I)-LPMO (left).



**Figure 47.** Proposed LPMO catalytic mechanism of the oxygen-dependent pathway for the hydroxylation of cellulose.

There have been two major computational Quantum Mechanics/Molecular Mechanics (QM/MM) studies investigating the reaction of Cu(I)-LPMO active site with H<sub>2</sub>O<sub>2</sub> whereby a 'caged' or 'precision guided' hydroxyl radical is a key intermediate leading to the copper(II)-oxyl.<sup>90, 132</sup> The experiments of Wang *et al.* constructed their study based on crystal structure coordinates of a fungal LPMO, *LsAA9*, already bound to a celotriose oligosaccharide (PDB:5ACF)<sup>160</sup>; whereas, the experiments of Bissaro *et al.* constructed their models based on crystal structural coordinates from a bacterial LPMO, *SmAA10* (PDB:2BEM) and introduced the chitin substrate using a molecular mechanics (MM) approach.<sup>132</sup> Despite differing substrates and slightly different coordination geometries, both sets of calculations resulted in very similar energetic landscapes (**Figure 48**). This indicates that this particular catalytic pathway could be available across the various families of LPMO. What is notable in both reaction pathways shown below is that the steps are predominantly exothermic with only small energetic barriers to overcome (**Figure 48**). The largely exothermic free energy landscape of the reaction of a Cu(I)-LPMO with H<sub>2</sub>O<sub>2</sub> indicates that this reaction will occur readily producing oxidised products and regenerating the Cu(I)-LPMO. The lack of any significant activation barriers in this mechanism means that generation of ROS species such as the copper(II)-oxyl are very fast and could be the reason for the indiscriminate enzyme oxidation and inactivation seen in many experiments of LPMOs with H<sub>2</sub>O<sub>2</sub>.

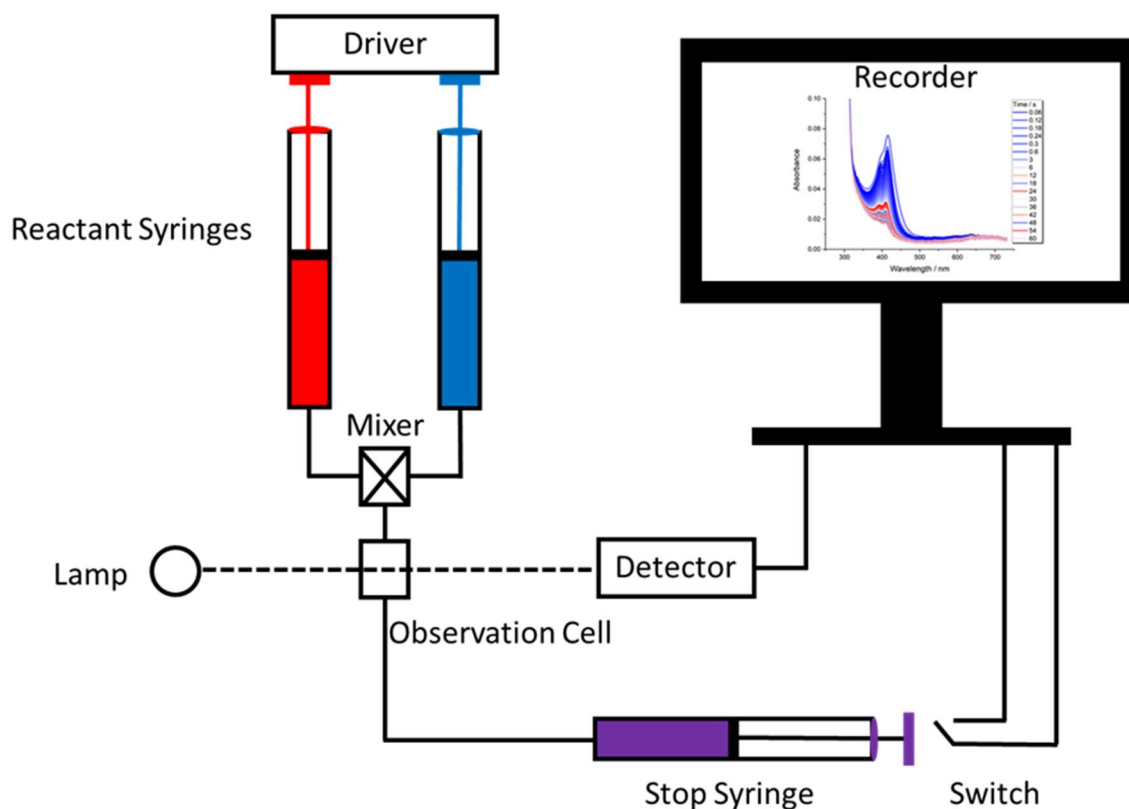


**Figure 48.** Calculated relative energies for the reaction of a Cu(I)-LsAA9-celotriose complex with H<sub>2</sub>O<sub>2</sub> (adapted from Wang *et al.* (A))<sup>90</sup> and a Cu(I)-SmAA10-chitin complex H<sub>2</sub>O<sub>2</sub> (adapted from Bisarro *et al.* (B)).<sup>132</sup>

These landmark computational papers have become pivotal in helping to build understanding as to how these cuproenzymes are able to oxidise such strong C-H bonds. Over time, the quality of the computational methods is rapidly improving with larger model sizes and higher accuracy methods becoming possible. This allows for an experimentalist to glean more information from their experiments by augmenting their experimental data with computational predictions. The predicted spectroscopic properties and energies from such calculations are becoming increasingly reliable and helping to unravel the remaining mysteries in the LPMO catalytic cycle.

### 1.9.3 Stopped-flow Spectrophotometry

While useful in building understanding of a chemical system, computational studies alone are not sufficient in unequivocally solving the mechanism of chemical reactions. As such, it is important to validate or debunk these studies with experimental evidence of intermediates and observables. Due to the oxidative potency of the predicted intermediates (copper(II)-oxyl and copper(II)-superoxide), experimental evidence has shown to be challenging to obtain. The lifetimes of these intermediates are likely to be incredibly short and capture of such species will necessitate particular conditions – most likely outside those routinely used for biological samples. One of the major approaches utilised to monitor reactions is the use of stopped-flow spectrophotometry. This technique is performed by storing two (or more) reactants in separate syringes (**Figure 49**). The contents of the syringes are then pushed into a mixing chamber upon command, facilitated by a driving piston. The driving of the contents into the mixing chamber and subsequently in the observation cell begins the reaction of the syringe contents. Furthermore, the increase in pressure causes a stop syringe to immediately trigger the switch which begins data collection. The data obtained from these experiments are incremental UV-Visible spectra of the newly reacting species across a user-specified time frame. By recording the UV-Visible spectrum over time, it is possible to monitor the progress of a reaction providing either the reactants, intermediates, or products (ideally all) contain a chromophore. For the studying LPMO mechanism, it is possible to use gas-tight syringes and to purge the system with nitrogen gas prior to data collection. This helps ensure that the reduced Cu(I)-LPMO does not oxidise prior to data collection.



**Figure 49. Schematic of a Stopped-Flow spectrophotometer apparatus based on the work of Bagshaw.<sup>171</sup>**

Several stopped flow studies of LPMOs have been performed to date in the means of reacting the reduced enzyme (Cu(I)-LPMO) with an oxidant, usually  $O_2$ ,  $H_2O_2$  or a peroxybenzoic acid such as meta-chloroperoxybenzoic acid (*mCPBA*).<sup>54, 152, 172</sup> So far, only oxidised amino acid residues have been detected, with Cu(II)- $O_x$  intermediates evading detection.

The oxidised amino acid residues detected so far have been characterised as tyrosyl or tryptophanyl radicals.<sup>172</sup> Experiments from Jones *et al.* showed two optical species arising from the reaction of the fungal enzyme *Hypocrea jecorina* LPMO9A (CuI-HjLPMO9A) with  $H_2O_2$ .<sup>172</sup> The first species formed was assigned to be a tryptophanyl radical determined by a characteristic absorption band centred at  $\lambda_{max} = 330$  &  $520$  nm as well as the EPR signature. The location of the oxidised tryptophan residue was found to be approximately  $6 \text{ \AA}$  from the copper centre. This radical subsequently decayed leading to the formation of a second species which was determined to be a tyrosyl radical. The tyrosyl radical was assigned to be the active site tyrosine that is present in the coordination environment of most LPMOs; the resulting radical of which is antiferromagnetically coupled to the Cu(II) centre. This assignment was in keeping with their experimental data showing optical spectra consistent with a tyrosyl radical but with a silent EPR signature.



These oxidised amino acid residues are believed to make up a 'hole-hopping pathway' somewhat protecting the enzyme from autoinactivation caused by uncoupled turnover.<sup>152, 172</sup> Despite these protective mechanisms, however, a study by Paradisi *et al.* identified the formation of a deactivated state of an enzyme *LsAA9* during a stopped-flow kinetics experiment with a unique, stable chromophore.<sup>152</sup> This chromophore possesses a unique absorption profile with intense feature at both 490 and 850 nm and was assigned to a Cu(II)-tyrosyl radical complex arising from the oxidation of the active site tyrosine residue at the active site of AA9 LPMOs. This oxidised species was found to be an irreversible and catalytically inactive form of the enzyme. Such a finding is significant in the argument of O<sub>2</sub> vs H<sub>2</sub>O<sub>2</sub> as the correct oxidising co-substrate.

Despite numerous attempts at characterising intermediates from the reaction of Cu(I)-LPMOs with an oxidant, little direct evidence exists that aid in the determination of the exact reaction pathway. As such, the chemical mechanism by which LPMO enzymes oxidise strong polysaccharide C-H bonds with pin-point accuracy remains a debated topic. Moreover, it is unclear whether the different LPMO families proceed through a universal chemical mechanism, or whether nuances exist between them depending on both the substrate and differences in primary and secondary coordination sphere arrangements. To properly unravel the LPMO catalytic mechanism it is necessary to shed more light on the intricacies and differences between the LPMO families by performing thoughtful crystallographic, spectroscopic, and theoretical studies.

# References

1. D. Klemm, B. Heublein, H.-P. Fink and A. Bohn, *Angewandte Chemie International Edition*, 2005, **44**, 3358-3393.
2. M. M. Martin, C. G. Jones and E. A. Bernays, *Philosophical Transactions: Biological Sciences*, 1991, **333**, 281-288.
3. P. Gupta, K. Samant and A. Sahu, *Int J Microbiol*, 2012, **2012**, 578925.
4. P. Baldrian and V. Valášková, *FEMS Microbiology Reviews*, 2008, **32**, 501-521.
5. S. Prasad and A. P. Ingle, in *Sustainable Bioenergy*, eds. M. Rai and A. P. Ingle, Elsevier, 2019, DOI: <https://doi.org/10.1016/B978-0-12-817654-2.00012-5>, pp. 327-346.
6. D. J. Tenenbaum, *Environ Health Perspect*, 2008, **116**, A254-A257.
7. R. A. Lee and J.-M. Lavoie, *Animal Frontiers*, 2013, **3**, 6-11.
8. R. M. Boddey, L. H. d. B. Soares, B. J. R. Alves and S. Urquiaga, in *Biofuels, Solar and Wind as Renewable Energy Systems: Benefits and Risks*, ed. D. Pimentel, Springer Netherlands, Dordrecht, 2008, DOI: 10.1007/978-1-4020-8654-0\_13, pp. 321-356.
9. L. Augusto Horta Nogueira and R. Silva Capaz, *Global Food Security*, 2013, **2**, 117-125.
10. E.-M. Aro, *Ambio*, 2016, **45**, 24-31.
11. H. C. J. Godfray, *The Journal of Agricultural Science*, 2014, **152**, 2-8.
12. M. H. L. Silveira, B. A. Vanelli and A. K. Chandel, in *Advances in Sugarcane Biorefinery*, eds. A. K. Chandel and M. H. Luciano Silveira, Elsevier, 2018, DOI: <https://doi.org/10.1016/B978-0-12-804534-3.00006-9>, pp. 135-152.
13. K. Robak and M. Balcerek, *Food Technol Biotechnol*, 2018, **56**, 174-187.
14. G. J. Davies, A. Planas and C. Rovira, *Acc Chem Res*, 2012, **45**, 308-316.
15. J. Hanske, J. Schulze, J. Aretz, R. McBride, B. Loll, H. Schmidt, Y. Knirel, W. Rabsch, M. C. Wahl, J. C. Paulson and C. Rademacher, *J Biol Chem*, 2017, **292**, 862-871.
16. J. Ø. Duus, C. H. Gotfredsen and K. Bock, *Chemical Reviews*, 2000, **100**, 4589-4614.
17. A. Udayan, M. Arumugam and A. Pandey, in *Algal Green Chemistry*, eds. R. P. Rastogi, D. Madamwar and A. Pandey, Elsevier, Amsterdam, 2017, DOI: <https://doi.org/10.1016/B978-0-444-63784-0.00004-7>, pp. 65-89.
18. S. Ummartyotin and H. Manuspiya, *Renewable and Sustainable Energy Reviews*, 2015, **41**, 402-412.
19. I. Miwa, K. Maeda and J. Okuda, *Experientia*, 1978, **34**, 167-169.
20. K. B. Wiberg, W. F. Bailey, K. M. Lambert and Z. D. Stempel, *The Journal of Organic Chemistry*, 2018, **83**, 5242-5255.
21. R. J. Moon, A. Martini, J. Nairn, J. Simonsen and J. Youngblood, *Chemical Society Reviews*, 2011, **40**, 3941-3994.
22. L. J. Gibson, *J R Soc Interface*, 2012, **9**, 2749-2766.
23. H. V. Scheller and P. Ulvskov, *Annual Review of Plant Biology*, 2010, **61**, 263-289.
24. V. K. Ponnusamy, D. D. Nguyen, J. Dharmaraja, S. Shobana, J. R. Banu, R. G. Saratale, S. W. Chang and G. Kumar, *Bioresource Technology*, 2019, **271**, 462-472.
25. H. Seddiqi, E. Oliaei, H. Honarkar, J. Jin, L. C. Geonzon, R. G. Bacabac and J. Klein-Nulend, *Cellulose*, 2021, **28**, 1893-1931.
26. R. Kohn and P. Kovac, *Chem. Zvesti*, 1978, **32**, 478-485.
27. B. R. Thakur, R. K. Singh and A. K. Handa, *Crit Rev Food Sci Nutr*, 1997, **37**, 47-73.
28. M. Kaya, E. Lelešius, R. Nagrockaitė, I. Sargin, G. Arslan, A. Mol, T. Baran, E. Can and B. Bitim, *PLoS One*, 2015, **10**, e0115531-e0115531.
29. J. L. Shamshina, P. Berton and R. D. Rogers, *ACS Sustainable Chemistry & Engineering*, 2019, **7**, 6444-6457.

30. M. Kaya, M. Mujtaba, H. Ehrlich, A. M. Salaberria, T. Baran, C. T. Amemiya, R. Galli, L. Akyuz, I. Sargin and J. Labidi, *Carbohydrate Polymers*, 2017, **176**, 177-186.
31. M.-K. Jang, B.-G. Kong, Y.-I. Jeong, C. H. Lee and J.-W. Nah, *Journal of Polymer Science Part A: Polymer Chemistry*, 2004, **42**, 3423-3432.
32. M. Nurunnabi, V. Revuri, K. M. Huh and Y.-k. Lee, in *Nanostructures for Oral Medicine*, eds. E. Andronesco and A. M. Grumezescu, Elsevier, 2017, DOI: <https://doi.org/10.1016/B978-0-323-47720-8.00015-8>, pp. 409-433.
33. D. Elieh-Ali-Komi and M. R. Hamblin, *Int J Adv Res (Indore)*, 2016, **4**, 411-427.
34. T. Dai, G. P. Tegos, M. Burkatovskaya, A. P. Castano and M. R. Hamblin, *Antimicrob Agents Chemother*, 2009, **53**, 393-400.
35. R. Jayakumar, M. Prabakaran, P. T. Sudheesh Kumar, S. V. Nair and H. Tamura, *Biotechnol Adv*, 2011, **29**, 322-337.
36. A. Amore, P. N. Ciesielski, C.-Y. Lin, D. Salvachúa and V. Sánchez i Nogué, *Australian Journal of Chemistry*, 2016, **69**, 1201-1218.
37. A. Ardèvol and C. Rovira, *Journal of the American Chemical Society*, 2015, **137**, 7528-7547.
38. V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho and B. Henrissat, *Nucleic Acids Research*, 2013, **42**, D490-D495.
39. H. Chen and L. Wang, in *Technologies for Biochemical Conversion of Biomass*, eds. H. Chen and L. Wang, Academic Press, Oxford, 2017, DOI: <https://doi.org/10.1016/B978-0-12-802417-1.00004-1>, pp. 65-99.
40. F. Calderaro, L. E. Bevers and M. A. van den Berg, *Biomolecules*, 2021, **11**, 1098.
41. G. Davies and B. Henrissat, *Structure*, 1995, **3**, 853-859.
42. D. G. Naumoff, *Biochemistry (Moscow)*, 2011, **76**, 622-635.
43. L. L. Lairson, B. Henrissat, G. J. Davies and S. G. Withers, *Annu Rev Biochem*, 2008, **77**, 521-555.
44. E. T. Reese, R. G. H. Siu and H. S. Levinson, *J Bacteriol*, 1950, **59**, 485-497.
45. G. Vaaje-Kolstad, B. Westereng, S. J. Horn, Z. Liu, H. Zhai, M. Sorlie and V. G. H. Eijsink, *Science*, 2010, **330**, 219-222.
46. R. J. Quinlan, M. D. Sweeney, L. Lo Leggio, H. Otten, J.-C. N. Poulsen, K. S. Johansen, K. B. R. M. Krogh, C. I. Jørgensen, M. Tovborg, A. Anthonsen, T. Tryfona, C. P. Walter, P. Dupree, F. Xu, G. J. Davies and P. H. Walton, *Proceedings of the National Academy of Sciences*, 2011, **108**, 15079-15084.
47. G. R. Hemsworth, B. Henrissat, G. J. Davies and P. H. Walton, *Nature Chemical Biology*, 2014, **10**, 122-126.
48. V. V. Vu, W. T. Beeson, E. A. Span, E. R. Farquhar and M. A. Marletta, *Proc Natl Acad Sci U S A*, 2014, **111**, 13822-13827.
49. M. Couturier, S. Ladevèze, G. Sulzenbacher, L. Ciano, M. Fanuel, C. Moreau, A. Villares, B. Cathala, F. Chaspoul, K. E. Frandsen, A. Labourel, I. Herpoël-Gimbert, S. Grisel, M. Haon, N. Lenfant, H. Rogniaux, D. Ropartz, G. J. Davies, M. N. Rosso, P. H. Walton, B. Henrissat and J. G. Berrin, *Nat Chem Biol*, 2018, **14**, 306-310.
50. F. Sabbadin, G. R. Hemsworth, L. Ciano, B. Henrissat, P. Dupree, T. Tryfona, R. S. Marques, S. T. Sweeney, K. Besser, L. Elias, G. Pesante, Y. Li, A. A. Dowle, R. Bates, L. D. Gomez, R. Simister, G. J. Davies, P. H. Walton, N. C. Bruce and S. J. McQueen-Mason, *Nature Communications*, 2018, **9**.
51. C. Filiatrault-Chastel, D. Navarro, M. Haon, S. Grisel, I. Herpoël-Gimbert, D. Chevret, M. Fanuel, B. Henrissat, S. Heiss-Blanquet, A. Margeot and J.-G. Berrin, *Biotechnology for Biofuels*, 2019, **12**, 55.
52. F. Sabbadin, S. Urresti, B. Henrissat, A. O. Avrova, L. R. J. Welsh, P. J. Lindley, M. Csukai, J. N. Squires, P. H. Walton, G. J. Davies, N. C. Bruce, S. C. Whisson and S. J. McQueen-Mason, *Science*, 2021, **373**, 774-779.

53. P. Sun, C. V. F. P. Laurent, V. J. P. Boerkamp, G. van Erven, R. Ludwig, W. J. H. van Berkel and M. A. Kabel, *ChemSusChem*, **n/a**.
54. C. H. Kjaergaard, M. F. Qayyum, S. D. Wong, F. Xu, G. R. Hemsworth, D. J. Walton, N. A. Young, G. J. Davies, P. H. Walton, K. S. Johansen, K. O. Hodgson, B. Hedman and E. I. Solomon, *Proceedings of the National Academy of Sciences*, 2014, **111**, 8797-8802.
55. G. Vaaje-Kolstad, D. R. Houston, A. H. K. Riemen, V. G. H. Eijsink and D. M. F. van Aalten, *Journal of Biological Chemistry*, 2005, **280**, 11313-11319.
56. W. T. Beeson, C. M. Phillips, J. H. D. Cate and M. A. Marletta, *Journal of the American Chemical Society*, 2012, **134**, 890-892.
57. H. B. Gray and J. R. Winkler, *Proceedings of the National Academy of Sciences of the United States of America*, 2005, **102**, 3534-3539.
58. A. Hemschemeier and T. Happe, *Nature Reviews Chemistry*, 2018, **2**, 231-243.
59. A. M. McDonnell and C. H. Dang, *J Adv Pract Oncol*, 2013, **4**, 263-268.
60. J. R. Winkler and H. B. Gray, in *Molecular Electronic Structures of Transition Metal Complexes I*, eds. D. M. P. Mingos, P. Day and J. P. Dahl, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012, DOI: 10.1007/430\_2011\_55, pp. 17-28.
61. H. B. Gray, *Chemistry International*, 2019, **41**, 16-19.
62. A. Tiessen, P. Pérez-Rodríguez and L. J. Delaye-Arredondo, *BMC Research Notes*, 2012, **5**, 85.
63. R. C. Gregory, G. R. Hemsworth, J. P. Turkenburg, S. J. Hart, P. H. Walton and G. J. Davies, *Dalton Transactions*, 2016, **45**, 16904-16912.
64. J. Kramer, Ö. Özkaya and R. Kümmerli, *Nature Reviews Microbiology*, 2020, **18**, 152-163.
65. K. E. H. Frandsen, T. J. Simmons, P. Dupree, J.-C. N. Poulsen, G. R. Hemsworth, L. Ciano, E. M. Johnston, M. Tovborg, K. S. Johansen, P. von Freiesleben, L. Marmuse, S. Fort, S. Cottaz, H. Driguez, B. Henrissat, N. Lenfant, F. Tuna, A. Baldansuren, G. J. Davies, L. Lo Leggio and P. H. Walton, *Nat Chem Biol*, 2016, **12**, 298-303.
66. T. J. Simmons, K. E. H. Frandsen, L. Ciano, T. Tryfona, N. Lenfant, J. C. Poulsen, L. F. L. Wilson, T. Tandrup, M. Tovborg, K. Schnorr, K. S. Johansen, B. Henrissat, P. H. Walton, L. Lo Leggio and P. Dupree, *Nature Communications*, 2017, **8**, 1064.
67. T. Tandrup, T. Tryfona, K. E. H. Frandsen, K. S. Johansen, P. Dupree and L. Lo Leggio, *Biochemistry*, 2020, **59**, 3347-3358.
68. E. I. Solomon, D. E. Heppner, E. M. Johnston, J. W. Ginsbach, J. Cirera, M. Qayyum, M. T. Kieber-Emmons, C. H. Kjaergaard, R. G. Hadt and L. Tian, *Chemical Reviews*, 2014, **114**, 3659-3853.
69. M. J. Appel, K. K. Meier, J. Lafrance-Vanasse, H. Lim, C.-L. Tsai, B. Hedman, K. O. Hodgson, J. A. Tainer, E. I. Solomon and C. R. Bertozzi, *Proceedings of the National Academy of Sciences*, 2019, **116**, 5370-5375.
70. T. Krüger, T. Dierks and N. Sewald, *Biological Chemistry*, 2019, **400**, 289-297.
71. D. A. Miarzlou, F. Leisinger, D. Joss, D. Häussinger and F. P. Seebeck, *Chemical Science*, 2019, **10**, 7049-7058.
72. A. Paradisi, M. J. Steward, P. Lindley, G. J. Davies and P. H. Walton, in *Comprehensive Coordination Chemistry III*, eds. E. C. Constable, G. Parkin and L. Que Jr, Elsevier, Oxford, 2021, DOI: <https://doi.org/10.1016/B978-0-12-409547-2.14928-5>, pp. 500-523.
73. P. L. Holland and W. B. Tolman, *Journal of the American Chemical Society*, 2000, **122**, 6331-6332.
74. C. Zong, C. J. Wilson, T. Shen, P. Wittung-Stafshede, S. L. Mayo and P. G. Wolynes, *Proceedings of the National Academy of Sciences*, 2007, **104**, 3159-3164.
75. E. I. Solomon and R. G. Hadt, *Coordination Chemistry Reviews*, 2011, **255**, 774-789.
76. H. Nar, A. Messerschmidt, R. Huber, M. van de Kamp and G. W. Canters, *Journal of Molecular Biology*, 1991, **221**, 765-772.
77. P. M. Colman, H. C. Freeman, J. M. Guss, M. Murata, V. A. Norris, J. A. M. Ramshaw and M. P. Venkatappa, *Nature*, 1978, **272**, 319-324.

78. E. Vijgenboom, J. E. Busch and G. W. Canters, *Microbiology (Reading)*, 1997, **143 ( Pt 9)**, 2853-2863.
79. R. Crichton, in *Biological Inorganic Chemistry (Third Edition)*, ed. R. Crichton, Academic Press, 2019, DOI: <https://doi.org/10.1016/B978-0-12-811741-5.00014-X>, pp. 405-433.
80. M. Nojiri, Y. Xie, T. Inoue, T. Yamamoto, H. Matsumura, K. Kataoka, Deligeer, K. Yamaguchi, Y. Kai and S. Suzuki, *Proceedings of the National Academy of Sciences*, 2007, **104**, 4315-4320.
81. M. A. McGuirl and D. M. Dooley, in *Encyclopedia of Inorganic and Bioinorganic Chemistry*, DOI: <https://doi.org/10.1002/9781119951438.eibc0274>.
82. J. Finney, H.-J. Moon, T. Ronnebaum, M. Lantz and M. Mure, *Arch Biochem Biophys*, 2014, **546**, 19-32.
83. S. i. Kishishita, T. Okajima, M. Kim, H. Yamaguchi, S. Hirota, S. Suzuki, S. i. Kuroda, K. Tanizawa and M. Mure, *Journal of the American Chemical Society*, 2003, **125**, 1041-1055.
84. M. Y. M. Pau, J. D. Lipscomb and E. I. Solomon, *Proceedings of the National Academy of Sciences of the United States of America*, 2007, **104**, 18355-18362.
85. M. Sendovski, M. Kanteev, V. S. Ben-Yosef, N. Adir and A. Fishman, *Journal of Molecular Biology*, 2011, **405**, 227-237.
86. M. Pretzler and A. Rompel, *Inorganica Chimica Acta*, 2018, **481**, 25-31.
87. N. Kitajima, K. Fujisawa, Y. Morooka and K. Toriumi, *Journal of the American Chemical Society*, 1989, **111**, 8975-8976.
88. C. A. Ramsden and P. A. Riley, *Bioorganic & Medicinal Chemistry*, 2014, **22**, 2388-2395.
89. D. M. Petrović, B. Bissaro, P. Chylenski, M. Skaugen, M. Sørli, M. S. Jensen, F. L. Aachmann, G. Courtade, A. Várnai and V. G. H. Eijsink, *Protein Sci*, 2018, **27**, 1636-1650.
90. B. Wang, E. M. Johnston, P. Li, S. Shaik, G. J. Davies, P. H. Walton and C. Rovira, *ACS Catalysis*, 2018, **8**, 1346-1351.
91. P. J. Lindley, A. Parkin, G. Davies and P. H. Walton, *Faraday Discussions*, 2021, DOI: 10.1039/D1FD00068C.
92. G. Jagadeeswaran, L. Gainey, R. Prade and A. J. Mort, *Appl Microbiol Biotechnol*, 2016, **100**, 4535-4547.
93. S. Kim, J. Ståhlberg, M. Sandgren, R. S. Paton and G. T. Beckham, *Proceedings of the National Academy of Sciences*, 2014, **111**, 149-154.
94. Z. Forsberg, Å. K. Røhr, S. Mekasha, K. K. Andersson, V. G. H. Eijsink, G. Vaaje-Kolstad and M. Sørli, *Biochemistry*, 2014, **53**, 1647-1656.
95. K. E. H. Frandsen, M. Haon, S. Grisel, B. Henrissat, L. Lo Leggio and J. G. Berrin, *J Biol Chem*, 2021, **296**, 100086.
96. X. Zhou, X. Qi, H. Huang and H. Zhu, *Int J Mol Sci*, 2019, **20**, 4594.
97. G. R. Hemsworth, B. Henrissat, G. J. Davies and P. H. Walton, *Nat Chem Biol*, 2014, **10**, 122-126.
98. M. Gudmundsson, S. Kim, M. Wu, T. Ishida, M. H. Momeni, G. Vaaje-Kolstad, D. Lundberg, A. Royant, J. Ståhlberg, V. G. Eijsink, G. T. Beckham and M. Sandgren, *J Biol Chem*, 2014, **289**, 18782-18792.
99. L. Lo Leggio, T. J. Simmons, J.-C. N. Poulsen, K. E. H. Frandsen, G. R. Hemsworth, M. A. Stringer, P. von Freiesleben, M. Tovborg, K. S. Johansen, L. De Maria, P. V. Harris, C.-L. Soong, P. Dupree, T. Tryfona, N. Lenfant, B. Henrissat, G. J. Davies and P. H. Walton, *Nature Communications*, 2015, **6**, 5961.
100. B. Misof, S. Liu, K. Meusemann, R. S. Peters, A. Donath, C. Mayer, P. B. Frandsen, J. Ware, T. Flouri, R. G. Beutel, O. Niehuis, M. Petersen, F. Izquierdo-Carrasco, T. Wappler, J. Rust, A. J. Aberer, U. Aspöck, H. Aspöck, D. Bartel, A. Blanke, S. Berger, A. Böhm, T. R. Buckley, B. Calcott, J. Chen, F. Friedrich, M. Fukui, M. Fujita, C. Greve, P. Grobe, S. Gu, Y. Huang, L. S. Jermin, A. Y. Kawahara, L. Krogmann, M. Kubiak, R. Lanfear, H. Letsch, Y. Li, Z. Li, J. Li, H. Lu, R. Machida, Y. Mashimo, P. Kapli, D. D. McKenna, G. Meng, Y. Nakagaki, J. L. Navarrete-Heredia, M. Ott, Y. Ou, G. Pass, L. Podsiadlowski, H. Pohl, B. M. v. Reumont, K. Schütte, K. Sekiya, S. Shimizu, A.

- Slipinski, A. Stamatakis, W. Song, X. Su, N. U. Szucsich, M. Tan, X. Tan, M. Tang, J. Tang, G. Timelthaler, S. Tomizuka, M. Trautwein, X. Tong, T. Uchifune, M. G. Walzl, B. M. Wiegmann, J. Wilbrandt, B. Wipfler, T. K. F. Wong, Q. Wu, G. Wu, Y. Xie, S. Yang, Q. Yang, D. K. Yeates, K. Yoshizawa, Q. Zhang, R. Zhang, W. Zhang, Y. Zhang, J. Zhao, C. Zhou, L. Zhou, T. Ziesmann, S. Zou, Y. Li, X. Xu, Y. Zhang, H. Yang, J. Wang, J. Wang, K. M. Kjer and X. Zhou, *Science*, 2014, **346**, 763-767.
101. R. Peterson and H. Nevalainen, *Microbiology*, 2012, **158**, 58-68.
102. E. M. Goss, J. F. Tabima, D. E. L. Cooke, S. Restrepo, W. E. Fry, G. A. Forbes, V. J. Fieland, M. Cardenas and N. J. Grünwald, *Proceedings of the National Academy of Sciences*, 2014, **111**, 8791-8796.
103. G. Courtade, S. B. Le, G. I. Sætrom, T. Brautaset and F. L. Aachmann, *Carbohydr Res*, 2017, **448**, 212-219.
104. G. R. Hemsworth, L. Ciano, G. J. Davies and P. H. Walton, *Methods Enzymol*, 2018, **613**, 63-90.
105. H. Jeong, V. Barbe, C. H. Lee, D. Vallenet, D. S. Yu, S.-H. Choi, A. Couloux, S.-W. Lee, S. H. Yoon, L. Cattolico, C.-G. Hur, H.-S. Park, B. Ségurens, S. C. Kim, T. K. Oh, R. E. Lenski, F. W. Studier, P. Daegelen and J. F. Kim, *Journal of Molecular Biology*, 2009, **394**, 644-652.
106. K. K. Meier, S. M. Jones, T. Kaper, H. Hansson, M. J. Koetsier, S. Karkehabadi, E. I. Solomon, M. Sandgren and B. Kelemen, *Chemical reviews*, 2018, **118**, 2593-2635.
107. G. Bertani, *J Bacteriol*, 1951, **62**, 293-300.
108. K. E. Kram and S. E. Finkel, *Appl Environ Microbiol*, 2015, **81**, 4442-4450.
109. G. S. Gray and W. M. Fitch, *Mol Biol Evol*, 1983, **1**, 57-66.
110. S. Brander, I. Horvath, J. Ø. Ipsen, A. Peculyte, L. Olsson, C. Hernández-Rollán, M. H. H. Nørholm, S. Mossin, L. L. Leggio, C. Probst, D. J. Thiele and K. S. Johansen, *Scientific Reports*, 2020, **10**, 16369.
111. B. Bennett and J. M. Kowalski, *Methods Enzymol*, 2015, **563**, 341-361.
112. N. J. Bunce, *Journal of Chemical Education*, 1987, **64**, 907.
113. M. M. Roessler and E. Salvadori, *Chemical Society Reviews*, 2018, **47**, 2534-2553.
114. E. Garribba and G. Micera, *Journal of Chemical Education*, 2006, **83**, 1229.
115. C. M. Marian, *WIREs Computational Molecular Science*, 2012, **2**, 187-203.
116. E. I. L. A. B. P. Solomon, *Inorganic electronic structure and spectroscopy*, Wiley, New York, 1999.
117. J. Bendix, *Angewandte Chemie International Edition*, 2017, **56**, 2247-2247.
118. M. Bennati and D. M. Murphy, in *Electron Paramagnetic Resonance*, 2009, DOI: <https://doi.org/10.1002/9780470432235.ch6>, pp. 195-250.
119. B. Caliskan and A. Caliskan, 2020, DOI: 10.5772/intechopen.91791.
120. A. Abragam, M. H. L. Pryce and F. E. Simon, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 1951, **205**, 135-153.
121. S. Shehada, M. dos Santos Dias, F. S. M. Guimarães, M. Abusaa and S. Lounis, *npj Computational Materials*, 2021, **7**, 87.
122. M. L. Munzarová, P. Kubáček and M. Kaupp, *Journal of the American Chemical Society*, 2000, **122**, 11900-11913.
123. J. Peisach and W. E. Blumberg, *Arch Biochem Biophys*, 1974, **165**, 691-708.
124. A. Schweiger and G. Jeschke, 2001.
125. J. A. Cieslak, P. J. Focia and A. Gross, *Biochemistry*, 2010, **49**, 1486-1494.
126. G. Courtade, L. Ciano, A. Paradisi, P. J. Lindley, Z. Forsberg, M. Sørliie, R. Wimmer, G. J. Davies, V. G. H. Eijsink, P. H. Walton and F. L. Aachmann, *Proceedings of the National Academy of Sciences*, 2020, **117**, 19178-19189.
127. S. Stoll and A. Schweiger, *Journal of Magnetic Resonance*, 2006, **178**, 42-55.
128. C. Mendoza, A. Désert, L. Khrouz, C. A. Páez, S. Parola and B. Heinrichs, *Environmental Science and Pollution Research*, 2021, **28**, 25124-25129.

129. B. Bissaro, I. Isaksen, G. Vaaje-Kolstad, V. G. H. Eijsink and Å. K. Røhr, *Biochemistry*, 2018, **57**, 1893-1906.
130. L. Ciano, A. Paradisi, G. R. Hemsworth, M. Tovborg, G. J. Davies and P. H. Walton, *Dalton Transactions*, 2020, **49**, 3413-3422.
131. C. A. Fowler, F. Sabbadin, L. Ciano, G. R. Hemsworth, L. Elias, N. Bruce, S. McQueen-Mason, G. J. Davies and P. H. Walton, *Biotechnology for Biofuels*, 2019, **12**, 232.
132. B. Bissaro, B. Streit, I. Isaksen, V. G. H. Eijsink, G. T. Beckham, J. L. DuBois and Å. K. Røhr, *Proceedings of the National Academy of Sciences*, 2020, **117**, 1504-1513.
133. A. S. Borisova, T. Isaksen, M. Dimarogona, A. A. Kognole, G. Mathiesen, A. Várnai, Å. K. Røhr, C. M. Payne, M. Sørli, M. Sandgren and V. G. H. Eijsink, *Journal of Biological Chemistry*, 2015, **290**, 22955-22969.
134. Y. A. Theibich, S. P. A. Sauer, L. L. Leggio and E. D. Hedegård, *Computational and Structural Biotechnology Journal*, 2021, **19**, 555-567.
135. F. Neese, F. Wennmohs, U. Becker and C. Riplinger, *J Chem Phys*, 2020, **152**, 224108.
136. S. G. Balasubramani, G. P. Chen, S. Coriani, M. Diedenhofen, M. S. Frank, Y. J. Franzke, F. Furche, R. Grotjahn, M. E. Harding, C. Hättig, A. Hellweg, B. Helmich-Paris, C. Holzer, U. Huniar, M. Kaupp, A. Marefat Khah, S. Karbalaei Khani, T. Müller, F. Mack, B. D. Nguyen, S. M. Parker, E. Perlt, D. Rappoport, K. Reiter, S. Roy, M. Rückert, G. Schmitz, M. Sierka, E. Tapavicza, D. P. Tew, C. van Wüllen, V. K. Voora, F. Weigend, A. Wodyński and J. M. Yu, *J Chem Phys*, 2020, **152**, 184107.
137. P. W. Abegg and T.-K. Ha, *Molecular Physics*, 1974, **27**, 763-767.
138. F. Neese, *WIREs Computational Molecular Science*, 2012, **2**, 73-78.
139. M. Y. Amusia, A. Z. Msezane and V. R. Shaginyan, *Physica Scripta*, 2003, **68**, C133-C140.
140. B. Helmich-Paris, *The Journal of Chemical Physics*, 2019, **150**, 174121.
141. R. K. Nesbet, *The Journal of Physical Chemistry*, 1996, **100**, 6104-6106.
142. P. Hohenberg and W. Kohn, *Physical Review*, 1964, **136**, B864-B871.
143. P. Ziesche, S. Kurth and J. P. Perdew, *Computational Materials Science*, 1998, **11**, 122-127.
144. T. J. Giese and D. M. York, *The Journal of Chemical Physics*, 2010, **133**, 244107.
145. J. P. Perdew, K. Burke and M. Ernzerhof, *Physical Review Letters*, 1998, **80**, 891-891.
146. C. Lee, W. Yang and R. G. Parr, *Phys Rev B Condens Matter*, 1988, **37**, 785-789.
147. A. D. Becke, *The Journal of Chemical Physics*, 1993, **98**, 5648-5652.
148. V. Barone and M. Cossi, *The Journal of Physical Chemistry A*, 1998, **102**, 1995-2001.
149. S. Grimme, S. Ehrlich and L. Goerigk, *Journal of Computational Chemistry*, 2011, **32**, 1456-1465.
150. M. W. Parker, *J Biol Phys*, 2003, **29**, 341-362.
151. E. D. Hedegård and U. Ryde, *ACS Omega*, 2017, **2**, 536-545.
152. A. Paradisi, E. M. Johnston, M. Tovborg, C. R. Nicoll, L. Ciano, A. Dowle, J. McMaster, Y. Hancock, G. J. Davies and P. H. Walton, *Journal of the American Chemical Society*, 2019, **141**, 18585-18599.
153. F. Neese, *Current Opinion in Chemical Biology*, 2003, **7**, 125-135.
154. R. K. Szilagyi, M. Metz and E. I. Solomon, *The Journal of Physical Chemistry A*, 2002, **106**, 2994-3007.
155. E. D. Larsson, G. Dong, V. Veryazov, U. Ryde and E. D. Hedegård, *Dalton Transactions*, 2020, **49**, 1501-1512.
156. F. Neese, *JBIC Journal of Biological Inorganic Chemistry*, 2006, **11**, 702-711.
157. E. D. Hedegård and U. Ryde, *J Biol Inorg Chem*, 2017, **22**, 1029-1037.
158. D. Dhar, G. M. Yee, A. D. Spaeth, D. W. Boyce, H. Zhang, B. Dereli, C. J. Cramer and W. B. Tolman, *Journal of the American Chemical Society*, 2016, **138**, 356-368.
159. L. Bertini, R. Breglia, M. Lambrugh, P. Fantucci, L. De Gioia, M. Borsari, M. Sola, C. A. Bortolotti and M. Bruschi, *Inorganic Chemistry*, 2018, **57**, 86-97.

160. B. Wang, Z. Wang, G. J. Davies, P. H. Walton and C. Rovira, *ACS Catalysis*, 2020, **10**, 12760-12769.
161. B. Bissaro, K. Røhr Å, G. Müller, P. Chylenski, M. Skaugen, Z. Forsberg, S. J. Horn, G. Vaaje-Kolstad and V. G. H. Eijsink, *Nat Chem Biol*, 2017, **13**, 1123-1128.
162. J. A. Hangasky, A. T. Iavarone and M. A. Marletta, *Proceedings of the National Academy of Sciences*, 2018, **115**, 4915-4920.
163. G. Müller, P. Chylenski, B. Bissaro, V. G. H. Eijsink and S. J. Horn, *Biotechnology for Biofuels*, 2018, **11**, 209.
164. B. Wang, P. H. Walton and C. Rovira, *ACS catalysis*, 2019, **9**, 4958-4969.
165. R. Kont, B. Bissaro, V. G. H. Eijsink and P. Våljamäe, *Nature Communications*, 2020, **11**, 5786.
166. M. Hobisch, D. Holtmann, P. Gomez de Santos, M. Alcalde, F. Hollmann and S. Kara, *Biotechnology Advances*, 2021, **51**, 107615.
167. A. Bakac, *Inorganic Chemistry*, 2010, **49**, 3584-3593.
168. V. G. H. Eijsink, D. Petrovic, Z. Forsberg, S. Mekasha, Å. K. Røhr, A. Várnai, B. Bissaro and G. Vaaje-Kolstad, *Biotechnology for Biofuels*, 2019, **12**, 58.
169. A. W. Munro, K. J. McLean, J. L. Grant and T. M. Makris, *Biochem Soc Trans*, 2018, **46**, 183-196.
170. C. M. Phillips, W. T. Beeson, J. H. Cate and M. A. Marletta, *ACS Chemical Biology*, 2011, **6**, 1399-1406.
171. C. R. Bagshaw, in *Encyclopedia of Biophysics*, ed. G. C. K. Roberts, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, DOI: 10.1007/978-3-642-16712-6\_59, pp. 2460-2466.
172. S. M. Jones, W. J. Transue, K. K. Meier, B. Kelemen and E. I. Solomon, *Proceedings of the National Academy of Sciences*, 2020, **117**, 11916-11922.



## 2 Research Papers

Herein, we have collected a series of original research articles which seek to identify the various factors affecting the copper electronic structure in chitin-active AA10 LPMOs. By exploring the factors that influence the copper coordination, we attempt to build on the current understanding of how chitin-active AA10s perform a difficult C-H bond activation under relatively benign conditions. In these articles, the principal techniques used to explore the electronic and geometric structure of these enzymes were a combination of electron paramagnetic resonance (EPR) spectroscopy, x-ray crystallography and supporting density function theory (DFT) calculations. We have found this combination of experimental methods to be both complementary and insightful. The articles presented in chapters 3, 4, and 5, are studies which focus on two members of the chitin-active AA10 family from the bacteria: *Bacillus amyloliquefaciens* (*BaAA10*) and *Bacillus licheniformis* (*BIAA10*). The findings from both enzymes show great parallels, indicating the conclusions drawn from these analyses are likely representative of the chitin-active AA10 family as a whole.

The article detailed in chapter 3 “Mapping the protonation states of the histidine brace in an AA10 lytic polysaccharide monooxygenase using CW-EPR spectroscopy and DFT calculations” is an original first-author publication in which we identified the variable behaviour of the copper coordination geometry of *BIAA10* with response to changing pH. The findings from this article are important in understanding the nature of the exogenous and endogenous ligands in the AA10 active site.

The article presented in chapter 4 “Mechanistic basis of substrate–O<sub>2</sub> coupling within a chitin-active lytic polysaccharide monooxygenase: An integrated NMR/EPR study” is an original research article to which I made a significant contribution to. It is a large body of work combining lots of expertise from multiple authors (joint-first authored by Dr. Luisa Ciano and Dr. Gaston Courtade). This article explored the affect of substrate binding on the ligand field of the chitin-active AA10 enzyme, *BIAA10*. The findings of this article were significant in unveiling a potential coupling mechanism existing between substrate binding and activation of dioxygen.

Finally, the research article detailed in chapter 5 “On the Mechanism of Substrate-O<sub>2</sub> Coupling in Lytic Polysaccharide Monooxygenases: EPR, X-Ray Crystallography, and DFT Studies of Azide Binding to a Chitin-Active LPMO” is a near-final draft, first author manuscript soon to be submitted for peer-review. The findings outlined in this manuscript build upon the aforementioned PNAS publication and make new advances in uncovering the substrate-coupled O<sub>2</sub> activation mechanism displayed by AA10 LPMOs. The findings from this manuscript are important in advancing the current understanding of

how AA10 LPMOs can oxidise strong C-H bonds with perfect selectivity by generating the most potent intermediates only when the substrate is correctly positioned.

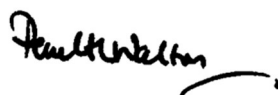
### 3 Mapping the protonation states of the histidine brace in an AA10 lytic polysaccharide monooxygenase using CW-EPR spectroscopy and DFT calculations

#### Declaration

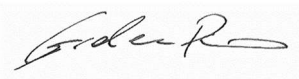
The following research paper describes the unique electronic structures that are adopted by an AA10 LPMO, *B/A*A10, depending on the variable pH of the buffering solution. The enzymes dependency on pH is of considerable industrial and biotechnological importance in reference to maximising their efficacy in biofuel production. A combined approach of EPR spectroscopy, UV-Visible spectroscopy and DFT calculations was used to characterise 2 single-deprotonation steps that occur at the active site of an AA10 LPMO enzyme between the pH values of 6.5-12.5 ( $pK_{a,s} = 8.7$  and  $\sim 11.5$ ). This study provides meaningful insight as to the various  $pK_a$  values of key active site residues of which may be significant considering the wider catalytic cycle.

For this work, I produced and purified *B/A*A10 protein samples; collected, analysed, and simulated the EPR spectra of *B/A*A10 samples at various pH values; conceived and executed DFT and TD-DFT calculations; performed UV-Vis calculations; performed SVD analyses. Z. Forsberg, G. Courtade, V. G. H. Eijsink and F. L. Aachmann generously provided the clone/cell colonies and purification protocol for the *B/A*A10 enzyme. Prof Paul Walton and I conceived the initial idea for and wrote the paper. All authors contributed to the final draft of the manuscript.

Signed Prof Paul H. Walton







Signed Prof Gideon J. Davies



Signed Dr Alison Parkin



# Mapping the protonation states of the histidine brace in an AA10 lytic polysaccharide monooxygenase using CW-EPR spectroscopy and DFT calculations†

Peter J. Lindley,  Alison Parkin,  Gideon J. Davies   
and Paul H. Walton \* 

Received 13th October 2021, Accepted 4th November 2021

DOI: 10.1039/d1fd00068c

The active site of the polysaccharide-degrading lytic polysaccharide monooxygenase (LPMO) enzyme features a single copper ion coordinated by a histidine brace. The primary coordination sphere of the copper contains several ligating atoms which are bonded to ionisable protons (e.g. OH<sub>2</sub>, NH), the pK<sub>a</sub>s of which are unknown. Using a combination of CW-EPR X-band spectroscopy over a range of pH values and DFT calculations, we show that the active site of a chitin-active AA10 LPMO can exist in three different protonation states (pK<sub>a1</sub> = 8.7, pK<sub>a2</sub> ~ 11.5), representing the ionisation of the coordinating groups. The middle pH species (fully formed at pH ~ 10.5) is proposed to be Cu(II)(His)<sub>2</sub>(OH)<sub>2</sub> (N<sub>2</sub>O<sub>2</sub> coordination) with a decoordinated R-NH<sub>3</sub><sup>+</sup> group at the amino terminus. This species also sees a rotation of the SOMO equatorial plane from the canonical histidine brace plane, whereby the nominal Cu d(x<sup>2</sup> - y<sup>2</sup>)-orbital has rotated some 45° along the His-Cu(II)-His axis, driven by the elongation and decoordination of the amino group. The highest pH species (>12) is proposed to exist as a Cu(II)-azanide, in which the NH<sub>2</sub> of the amino terminus has been deprotonated. The high pH means that this species is unlikely to be biologically relevant in the catalytic cycle of AA10 LPMOs.

## Introduction

Lytic polysaccharide monooxygenases (LPMOs) are copper-containing enzymes which catalyse the oxidative cleavage of glycosidic bonds within poly-saccharides.<sup>1-3</sup> The active site of LPMOs has been shown by crystallography to contain a single copper ion coordinated in a T-shaped configuration by a 'histidine brace', which consists of an N-terminal histidine chelating to the copper *via* its NH<sub>2</sub> and π-N imidazole atoms, along with the τ-N atom of a further histidine.<sup>4</sup>

Department of Chemistry, University of York, Heslington, York, YO10 5DD, UK. E-mail: paul.walton@york.ac.uk

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d1fd00068c

336 | *Faraday Discuss.*, 2022, 234, 336-348 This journal is © The Royal Society of Chemistry 2022

In the copper(II) resting state, the structures further show that one or two exogenous ligands, usually water or hydroxide, occupy the remaining copper coordination sites.

LPMOs are divided into auxiliary activity (AA) families in the Carbohydrate Active enZyme (CAZy) database ([www.cazy.org](http://www.cazy.org)) based on sequence similarity, where so far eight distinct classes have been characterised (AA9–11, AA13–17).<sup>1,2,5–10</sup> The number of exogenous ligands is partly dependent on the genomic class of the LPMO. For instance, structures of AA9-classified enzymes exhibit a single water/hydroxide to give a planar N<sub>3</sub>O coordination geometry around the copper, whereas chitin-active AA10-classified enzymes appear to show the copper(II) in a distorted square-pyramidal N<sub>3</sub>O<sub>2</sub> geometry, where two exogenous water/hydroxide ligands complete the coordination sphere, as depicted in Fig. 1. In both the AA9 and AA10 cases, EPR spectroscopy has established that the SOMO is dominated by the d(x<sup>2</sup> – y<sup>2</sup>) orbital, requiring the ligand field of the copper to have a large sigma-bonding component in the equatorial plane of the metal ion. In the case of AA9 LPMOs, the coordination geometry of the Cu is best described as axially elongated, whereas in the case of AA10 LPMOs, EPR spectroscopy has shown *ca.* 3% of d(z<sup>2</sup>) mixing into the largely d(x<sup>2</sup> – y<sup>2</sup>) SOMO, commensurate with a distorted square-pyramidal geometry.<sup>11,12</sup>

In chitin-active AA10 LPMOs, the large number of ionisable protons of the ligands, which coordinate to the copper ion in the histidine brace (NH<sub>2</sub>, 2 × H<sub>2</sub>O), opens up the possibility that the copper(II) centre exhibits a range of different protonation states, some of which may be relevant to the enzyme's catalytic cycle. Indeed, this reasoning has extended as far as suggesting that the amino terminus can deprotonate itself to form a copper(II)-azanide bond, the σ-donating power of which is expected to be high and possibly capable of stabilising a high valent copper species.<sup>13</sup> Such a species may be a key intermediate in the catalytic cycle which is capable of abstracting a hydrogen atom from the polysaccharide substrate.<sup>2,14</sup> As such, determining the different degrees of protonation of the active site is germane to any discussion about the catalytic mechanisms of LPMOs.

It is in this context that we aimed to establish the nature of the different protonation states of the active site of an AA10 LPMO from the bacterium *Bacillus*

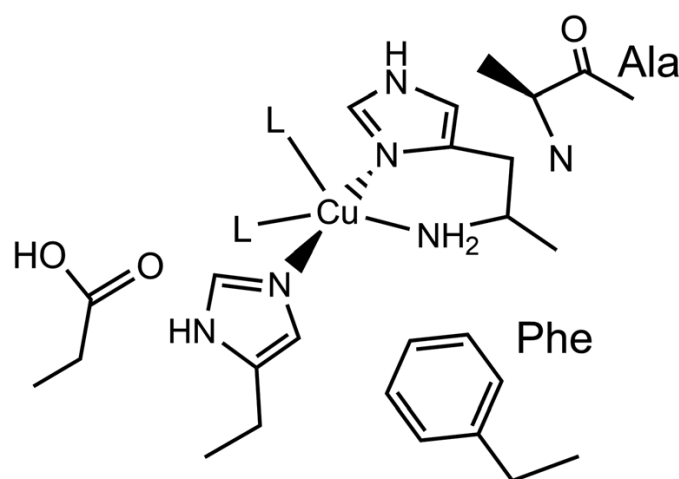


Fig. 1 Active site of a chitin-active AA10 LPMO, depicting the histidine brace exogenous ligands (L) ligating to the Cu(II) ion, and conserved residues.

*licheniformis*, *Bl*(AA10).<sup>15</sup> To this end, we employed CW-EPR spectroscopy in conjunction with DFT calculations to show that the active site exists in three distinct protonation states across the pH range 6.5 to 12.5. We further show that the highest pH species is most likely to be the copper(II)-azanide species proposed earlier,<sup>2,13</sup> but that it only exists at pH outside the usual biological range, and is not a biologically relevant species for the resting state of the enzyme.

## Results and discussion

### Preparation and purity of *Bacillus licheniformis* AA10 LPMO

Samples of this enzyme were prepared in *Escherichia coli* (BL21 competent cells) *via* periplasmic expression, previously reported by Courtade *et al.*<sup>16</sup> Crude protein was obtained from the periplasm by osmotic shock and further purified using a 5 mL HiTrap® DEAE FF anion exchanger, eluting with a NaCl gradient from 0–500 mM over 90 column volumes. LPMO-containing fractions were collected and further purified on a Superdex S75 gel filtration column (GE Life Sciences).

The purity of *Bl*(AA10) was established by gel electrophoresis, which displayed a clear and single band at *ca.* 20 kDa (ESI<sup>†</sup>). For the subsequent EPR experiments, it was important to determine that all copper in the sample was bound stoichiometrically to the active site of the enzyme. On this basis, a single equivalent of Cu was added to the enzyme from a solution of 10 mM CuSO<sub>4</sub> prior to size exclusion chromatography, from which any uncoordinated metal ions were separated from the purified enzyme.

### CW-EPR X-band spectroscopy

A sample of *Bl*(AA10) (0.4 mM) was prepared in a mixed buffer system of MES, CHES, HEPES and CAPS, each at 5 mM. These buffers were chosen for their known poor-coordination properties toward metal ions and, when used synchronously, buffer a wide range of pHs. The pH of the solution was adjusted up and down by the addition of dilute solutions of NaOH and H<sub>2</sub>SO<sub>4</sub>, respectively. The pH of the resulting solution at room temperature was determined using a calibrated Fisherbrand™ Accumet™ AE150 Benchtop pH Meter, from which 12 samples from pH 6.5 to 12.5 were prepared in 0.5 pH unit steps. To verify that the pH extremes did not denature the protein, EPR spectra were collected on samples starting at pH 6.5 and then raising to pH 12.5, and then recollected on the same sample at a lower pH, indeed restoring the previous species.

At each pH, the sample was rapidly (<2 s) cooled<sup>17</sup> to 150 K and its CW-EPR X band spectrum collected.‡ The trend of the EPR spectra shows a qualitative change across the pH range, in which it is evident that all species have a spectral envelope which is consistent with a predominantly  $d(x^2 - y^2)$  SOMO, showing that the Cu(II) ion at all pHs has a ligand field dominated by sigma donation in its equatorial plane (Fig. 2). Spin quantification by double integration of the spectra

---

‡ Continuous wave (CW) X-band EPR spectra were obtained using a 100  $\mu$ L frozen solution of *Bl*AA10 at a concentration of 0.4 mM. Data collection was carried out on a Bruker micro EMX spectrometer using a microwave frequency of *ca.* 9.3 GHz, modulation amplitude of 4G, modulation frequency of 100 kHz and microwave power of 10.02 mW. Spectra were intensity averaged over three scans and simulated using the EasySpin 5.2.28 (ref. 27) open-source toolbox implemented by MATLAB R2020a software on a desktop PC. Raw EPR data are available at DOI: 10.15124/7a485d50-ea4e-4cd1-8fc2-650057c076af.

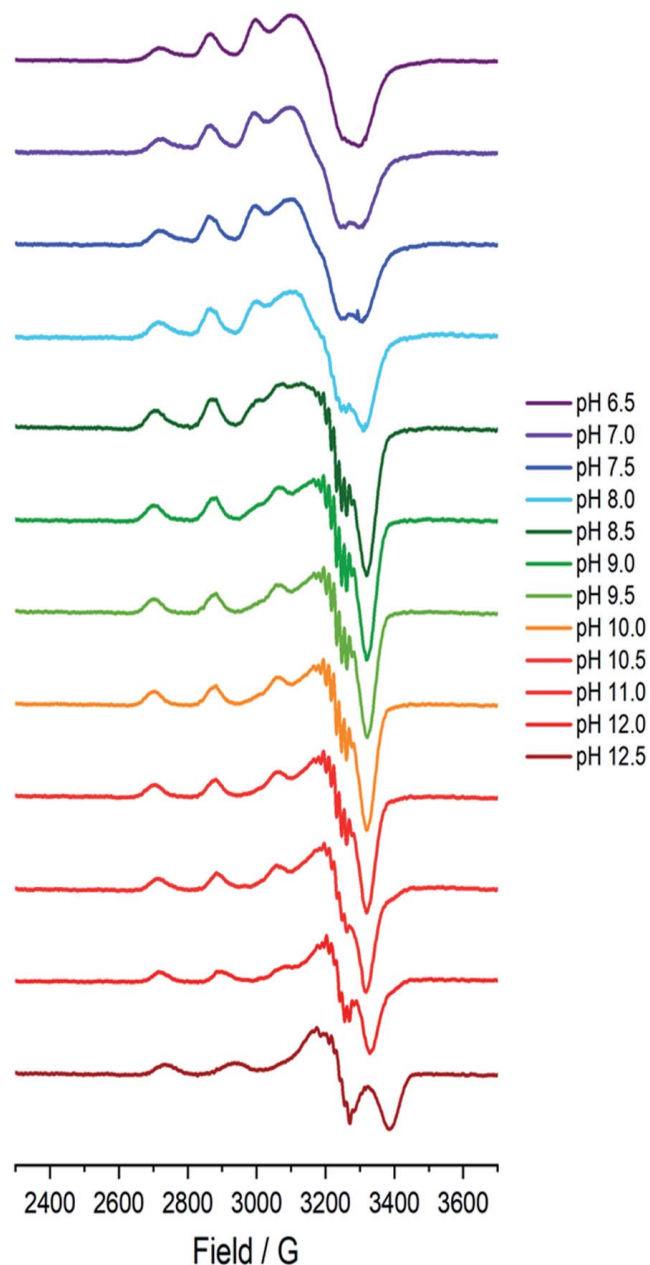


Fig. 2 Frozen solution X-band EPR spectra (150 K) of BI(AA10) (0.4 mM) in a mixed buffer of MES, CHES, HEPES and CAPS (5 mM each) between pH 6.5 and 12.5.

also shows that there are no redox processes that occur at the active site across the pH range (ESI<sup>†</sup>).

#### Analysis of EPR spectra and distribution of pH dependent species

Between pHs 6.5 and 8.0, a single species is seen to dominate the spectrum, as evidenced by a single set of Cu hyperfine split peaks at low field ( $g_3 = 2.260$ ,  $A_3 = 445$  MHz). Increasing the pH leads to the formation of a new set of peaks in the low field region ( $g_3 = 2.233$ ,  $A_3 = 554$  MHz) along with distinctive superhyperfine

coupling around  $g = 2.07$ , commensurate with the formation of a second distinct species, the concentration of which is maximised at pH 10.5. A further increase in pH to 12.5 sees the formation of a further set of peaks in the low field region ( $g_3 = 2.180$ ,  $A_3 = 614$  MHz), consistent with the formation of a third distinct species. This final species is not fully formed at pH 12.5 but is dominant at that pH. The total number of species across the pH range was assessed using singular value decomposition analysis of the collected spectra (ESI<sup>†</sup>). From this analysis, three principal components emerged, corroborating the qualitative assessment of the pH dependence of the EPR spectra.

On the basis of the three components (species 1, 2 and 3), we sought to deconvolute the series of spectra. Given the clear separation between the different species across the pH range, it was assumed that the spectrum at pH 6.5 represented 100% of a single complex. Indeed, it was possible to simulate (with EasySpin) the lowest pH species (pH 6.5) with a single set of spin-Hamiltonian parameters (Fig. 3a and Table 1). The same was also true of the spectra collected at pH 10.5 and 12.5 (Fig. 3b and c, respectively, Table 1). In performing the simulation of species 3, we noted the appearance of a significant 'overshoot' feature which appears at high field in the species at pH 12.5. Such a feature has in the past been confused with an anomalously low  $g_1$  value.<sup>18</sup>

Once spin Hamiltonian parameters of each species were obtained, it was possible to simulate the mixed species spectra (pHs 7 to 9.5 and >10.5) to determine the species distributions at each pH, from which the  $pK_a$ s of species 1 (low pH) and species 2 (middle pH), and the stoichiometry of protons for each ionisation, were determined using standard methods (ESI<sup>†</sup>). The deprotonation of species 1 to 2 occurs with a  $pK_a$  of 8.7 and is a single proton step. The deprotonation of species 2 to 3 occurs with a  $pK_a$  of  $\sim 11.5$  and is also a single proton step.

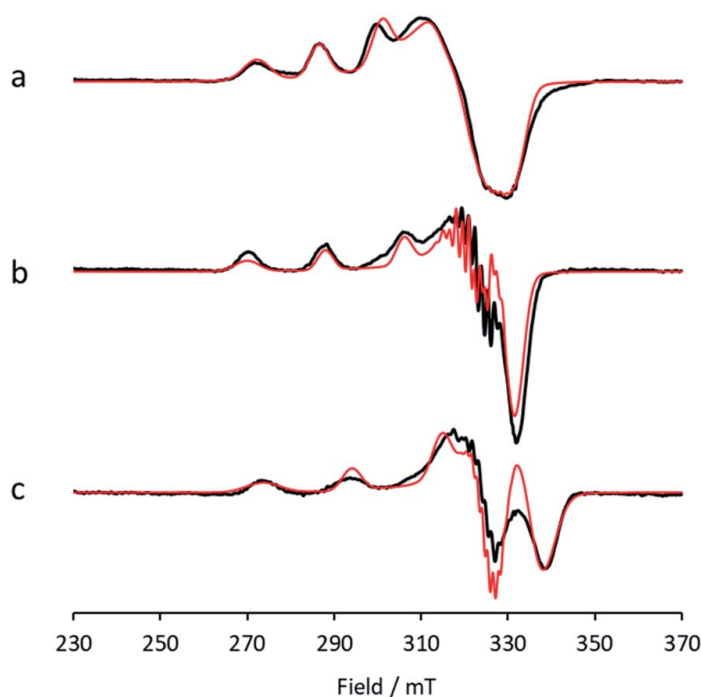


Fig. 3 X-band EPR (a) low pH species 1, with simulation (red), (b) intermediate pH species 2, with simulation (red), (c) high pH species 3, with simulation (red).



**Table 1** Spin Hamiltonian parameters of the pH dependent species. It is assumed that the  $A_3$  value is negative in sign, although this is not determined experimentally

Species	g values			Hyperfine coupling constants/MHz			$b^2/\%$
	$g_1$	$g_2$	$g_3$	$ A_1 $	$ A_2 $	$A_3$	
1	2.030	2.125	2.260	60	90	−445	3.6
2	2.044	2.069	2.233	60	76	−554	0.8
3	2.038	2.062	2.180	70	90	−614	0.9

### Ligand field analyses of spin Hamiltonian parameters

The collected spin Hamiltonian parameters of the three species are shown in Table 1. In all cases, the  $g$  values follow the pattern of  $g_1 < g_2 \ll g_3$ , confirming a SOMO with a high  $d(x^2 - y^2)$  character. Some  $d(z^2)$  mixing is evident for all species from the differences in the values of  $g_1$  and  $g_2$ . The degree of mixing,  $b^2$ , can be calculated from the standard perturbative EPR formulae as described by Hitchman *et al.*,<sup>19</sup> which give the molecular orbital coefficients,  $a$  and  $b$ , for the SOMO defined in eqn (1) (Table 1).

$$\psi(\text{SOMO}) = a\psi(x^2 - y^2) + b\psi(z^2) \quad (1)$$

This mixing is low for species 2 and 3, revealing that the ligand field around the Cu in each of these is close to axial four-fold symmetry (note that this does not necessarily translate into a four-fold axial symmetry of the identity of the co-ordinating atoms, as different coordinating atoms may have similar ligand field parameters). For species 1, more mixing is evident ( $b^2 = 3.6\%$ ), which is in-line with previous EPR studies on AA10 LPMOs<sup>11</sup> and the subsequent structures of these enzymes, which exhibit a distorted square pyramidal geometry around the copper ion (Fig. 1), where the exogenous ligands are two water/hydroxide molecules, one of which forms a Cu–O bond of *ca.* 1.9 Å, and the other a more distant bond of 2.2 Å.<sup>20</sup>

Taking the value of  $g_3$  as a measure of the relative ligand field strengths in the equatorial plane of the copper, it is evident that this ligand field increases significantly moving from species 1 to 2 to 3, and—moreover, from the values of  $d(z^2)$  mixing—that the ligand field in species 2 and 3 is relatively evenly distributed amongst the ligating atoms of the equatorial plane, suggestive of a near four-fold symmetric planar arrangement of the equatorial ligand field. Indeed, corroborating such a ligand field, the increase in  $A_3$  value from species 1 to 2 is also known to be associated with a change in the Fermi coupling as the coordination number of the copper drops from five to four.<sup>15</sup> Without access to the d–d electronic transition energies, it is not possible to be more precise about the nature of the ligand field; nevertheless the trend in increasing ligand field around the equatorial plane is evident from the relative spin-Hamiltonian parameters.

Further information about the ligand field in species 2 can be obtained from the clearly delineated Cu–N superhyperfine coupling (Fig. 3b). This coupling could be accurately simulated with the inclusion of two nitrogen atoms ( $^{14}\text{N}$ ,  $I = 1$ ) in the spin Hamiltonian, in which each N atom couples to the copper with

a principal coupling value of 40 MHz. This value is similar to those already known for the Cu–His coupling constants seen in other LPMOs.<sup>17,21</sup> The simulation of species 2 could be adequately performed with the inclusion of two N nuclei, incommensurate with the expected N<sub>3</sub> coordination of the histidine brace. Additionally, the  $g_3$  and  $A_3$  values of species 2 place it squarely within those expected for a CuN<sub>2</sub>O<sub>2</sub> coordination geometry on a canonical Peisach–Blumberg plot.<sup>22</sup>

In assigning a putative structure to species 2, it is informative to consider previous EPR studies on the substrate-bound state of chitin-active AA10 enzymes. These species have been well characterised as having a planar CuN<sub>3</sub>O coordination geometry with a  $g_z$  value of 2.205 and an  $A_z$  value of ~620 MHz.<sup>15,23</sup> On this basis, a similar axial CuN<sub>3</sub>O (sq. pl) ligand field is unlikely for species 2, given the discrepancies in the  $g_z$  and  $A_z$  values. This difference in the spin Hamiltonian parameters necessitates an alternative ligand field assignment for species 2. Thus, taking together the Cu spin Hamiltonian values and superhyperfine couplings, a more likely scenario is one in which the NH<sub>2</sub> group of the histidine brace is decoordinated from the copper (see DFT calculations below) and the coordination sphere around the Cu is made up of two histidines and two water/hydroxide ligands (see below for further discussion).

The spin Hamiltonian parameters for species 3 are distinct from those of other known LPMOs, which is particularly evident in the value of  $g_3$  (2.18), which is lower than all known values, including the situation in which a chloride ion has entered the coordination sphere of the copper ( $g_3 = 2.23$ ). Given the low value, we are directed towards the proposal that the ligand field around the copper in species 3 is one that is highly  $\sigma$ -donating, causing a large separation in the  $d(x^2 - y^2)$  and  $d(xy)$  orbital energies. Such a strong  $\sigma$  donation might be expected of a Cu–azanide interaction, which would arise from the recoordination of the  $-\text{NH}_3^+$  amino terminus in species 2 which had been doubly deprotonated to give a  $-\text{NH}^-$  ligating group. Commensurate with this proposal is the simulation of the superhyperfine coupling in species 3, which could only be satisfactorily achieved with the inclusion of three nitrogen atoms in the spin Hamiltonian with principal coupling constants of 35, 35 and 33 MHz (see DFT calculations below).

From the overall ligand field analysis, a hypothesis emerges for the nature of species 1, 2 and 3 and the pH-dependent ionisation processes between them. The hypothesis is depicted in Fig. 4, in which it is shown that species 2 forms *via* the deprotonation of a coordinating water molecule, followed by the decoordination and subsequent protonation of the  $-\text{NH}_2$  group, driven by an internal proton

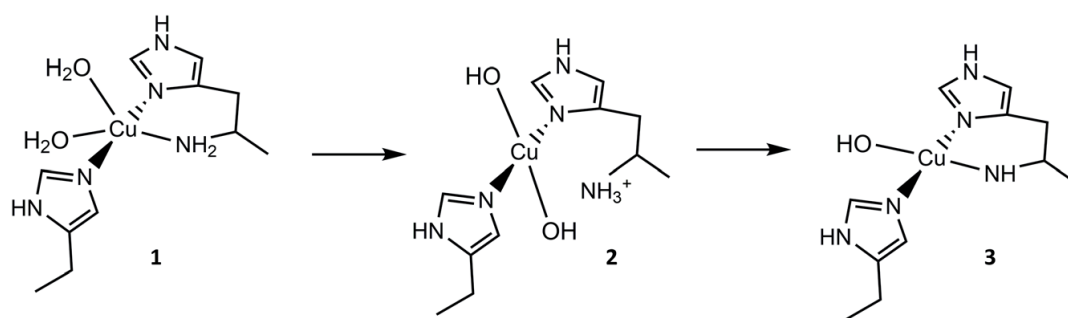


Fig. 4 Hypothetical protonation states and coordination geometries of species 1, 2 and 3.

transfer process from one of two coordinating water molecules in species 1. The formation of species 3 from species 2, on the other hand, sees the reverse proton transfer from the de-coordinated  $\text{-NH}_3^+$  to a coordinated hydroxide with the loss of the resulting water molecule, and the subsequent deprotonation of an N-H proton from the  $\text{-NH}_2$  and recoordination to the metal as an azanide. In support of such an assignment, the formation of species 3 was accompanied by an unusual colour change from colourless to pale pink (ESI†) due to the formation of a semi-intense absorption band ( $\epsilon \approx 2500 \text{ M}^{-1} \text{ cm}^{-1}$ ) centred at 520 nm. The intensity of this band is in keeping with that of a charge transfer transition – likely arising from a ligand to metal charge transfer (LMCT) from the proposed azanide ( $\text{-NH}^-$ ) group to the copper (see TD-DFT calculations, ESI†).

Thus, the proposed ligand field changes depicted in Fig. 4 are revealed by the differences in the spin Hamiltonian parameters across the species, in which  $g_z$  decreases and  $A_z$  increases (going from species 1 to 3), commensurate with an increasing equatorial ligand field.

### DFT calculations

Building on the hypothesis depicted in Fig. 4, density functional theory geometry optimisations (uBP86) and EPR calculations (uB3LYP with 38% HF exchange) were performed on the proposed species 1, 2 and 3.

In the absence of a high-resolution crystal structure for the *Bacillus licheniformis* AA10, crystallographic coordinates of the spectroscopically similar enzyme, *Bacillus amyloliquefaciens* AA10 (PDB 5IJU),<sup>20</sup> were used as the basis for these calculations, in keeping with the experiments of Courtade *et al.*<sup>15</sup> Importantly in this regard, the structure of BaAA10 is one of the few LPMO structures in which the copper ion has not been photoreduced from Cu(II) to Cu(I) during the X-ray experiment.<sup>24</sup> Accordingly, key residues from the crystallographic coordinates were used for the basis of the calculations for species 1, retaining the primary coordination sphere and supporting residues (full information and coordinates are given in the ESI†). Following standard practice, the spatial coordinates of certain atoms at the periphery of the model were kept fixed throughout the calculation to prevent the unfettered movement of residues. These atoms are depicted in Fig. S7 (ESI†) and were selected on the basis of earlier DFT studies on a closely related LPMO.<sup>15</sup>

For species 1, the geometry was initially modelled with the inclusion of two ligating water molecules with bond distances based on the positioning of the oxygen atoms in the crystal structure. Notably in this regard, the water situated beneath the histidine brace plane is expected to participate in strong hydrogen bonding with a nearby aspartate residue (glutamate in BaAA10), the potential of which can affect the protonation state and ligating properties of the water molecule. Indeed, upon optimisation, despite the fact that Glu and Asp residues are expected to be deprotonated across the studied pH range owing to the low  $\text{p}K_{\text{a}}$ s of their carboxylate groups ( $\text{p}K_{\text{a}} = 2\text{--}4.5$ ),<sup>25</sup> one of the protons bound to the water molecule migrated to the neighbouring carboxylate group (Fig. 5) to give a mixed hydroxide/water coordination at the Cu. Confidence in this structural assignment comes from the resulting geometry of the optimised structure, which is a close match to the crystal structure coordinates. Thus, for species 1, the coordination sphere of the Cu is best described as a distorted square-pyramid, where water and hydroxide act as exogenous ligands.

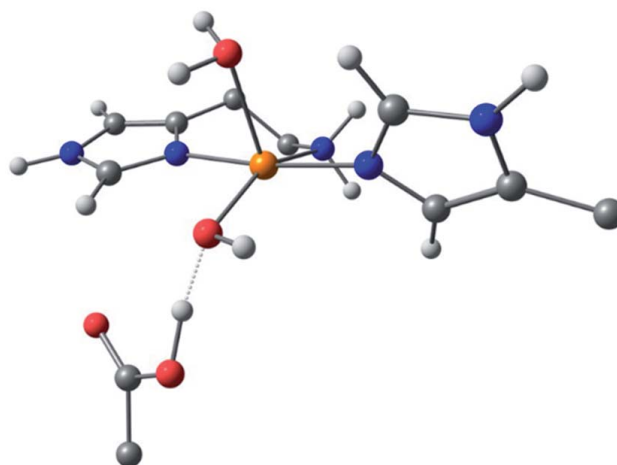


Fig. 5 DFT model of species 1 (truncated for clarity) showing the distorted square-pyramidal geometry and hydroxide/water coordination of the exogenous ligands.

While a good match was found between the optimised and experimental structures, the match between the calculated and experimental EPR parameters (Table 2) was only moderate (*e.g.* calculated  $g_3 = 2.216$ , experimental  $g_3 = 2.26$ ). Moreover, the calculated Cu hyperfine values,  $A_1$  and  $A_2$ , have relative magnitudes which are opposite to those observed experimentally. Notwithstanding the known difficulties in DFT-calculated EPR parameters matching with those from experiments, the relatively large differences seen between the experimental and calculated EPR parameters for species 1 are a recurring theme in DFT calculations on AA10 LPMO active sites in their Cu(II) resting states, even with careful selection of the basis sets and functionals.<sup>26</sup> As such, while the optimised structure of 1 matches well that of the experiment and is likely a good representation of the actual structure, caution must be applied in interpreting the calculated EPR data for species 1, where the reasons for the discrepancies between theoretical and experimental results remain unclear. This situation is in contrast to the calculations of the EPR parameters for species 2 and 3, which give a somewhat closer match to the experimental data (see below).

Following the ligand field analysis above, species 2 was initially modelled by removing the proton shared between the water and glutamate, and by transferring the proton attached to the second water ligand to the  $-NH_2$  group (representing an overall removal of a single proton from species 1, as per the experimental

Table 2 Calculated spin Hamiltonian parameters of the pH dependent species. The  $x$  axis of the calculated  $g$  tensor lies along the His–Cu–His direction

Species	$g$ values			Hyperfine coupling constants/MHz		
	$g_x$	$g_y$	$g_z$	$ A_x $	$ A_y $	$A_z$
1	2.049	2.081	2.216	110	29	–598
2	2.070	2.079	2.244	94	3	–582
3	2.037	2.070	2.188	15	140	–496

findings). Upon optimisation, the angle between the hydroxide ligands,  $\angle\text{O-Cu-O}$ , increased from *ca.*  $90^\circ$  to  $157^\circ$ . The divergence of the hydroxide ligands was also accompanied by the expected decoordination of the now protonated  $-\text{NH}_3^+$  group from the metal and a rotation of the  $d(x^2 - y^2)$  SOMO by *ca.*  $45^\circ$  (Fig. 6), such that the equatorial plane was now defined by the two hydroxides and the nitrogen atoms of two histidine ligands, thus completing the  $\text{CuN}_2\text{O}_2$  coordination geometry expected from the ligand field analysis. The calculated EPR parameters (Table 2) for species 2 reflect the experimental data in that  $g_x$  and  $g_y$  are closer in values, and that the value of  $g_z$  is *ca.* 2.24.

Species 3 was modelled as a Cu(II)-azanide species formed by the expulsion of water from species 2, followed by the further removal of an  $-\text{NH}_2$  proton, resulting in a short Cu-NH bond of 1.94 Å (full computational details can be found in the ESI†). The optimised geometry is best described as  $\text{CuN}_3\text{O}$  planar, albeit distorted ( $\angle\text{HN-Cu-O} \sim 160^\circ$ ). This distortion away from planar has been previously re-reported for four-coordinate LPMO structures when modelled by DFT.<sup>15</sup> The calculated EPR parameters for species 3 match well with those from experiments,

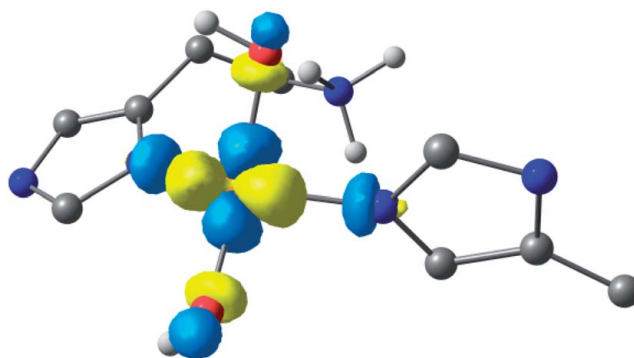


Fig. 6 Hypothetical  $d(x^2 - y^2)$  dominated SOMO of species 2 showing that the orbital plane has rotated (the colours denote the orbital phases).

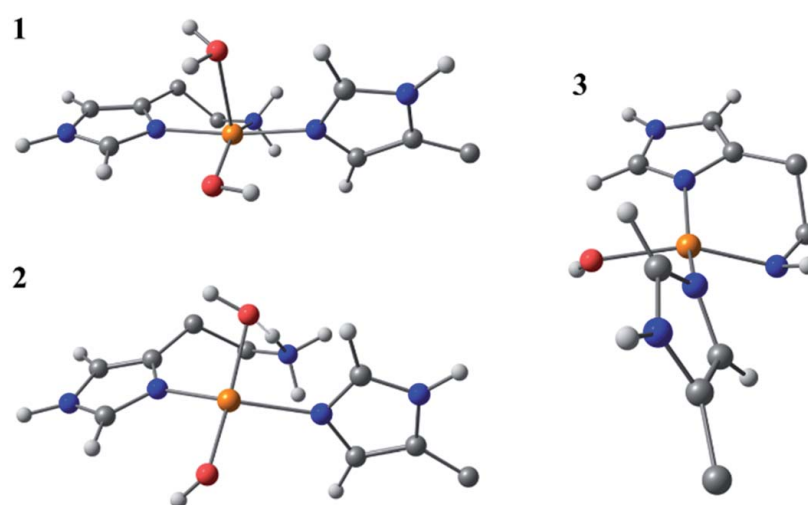


Fig. 7 DFT optimised geometries of species 1, 2 and 3. The models are truncated, and certain atoms are hidden for clarity.

most notably in the  $g_{3,z}$  value of 2.18. Also, from a spectroscopic perspective, Species 3 exhibits a semi-intense UV/visible absorption band centred at 520 nm ( $\epsilon \approx 2500 \text{ M}^{-1} \text{ cm}^{-1}$ ). This band is proposed to arise from a LMCT transition from the proposed azanide group to the Cu, which is corroborated with the TD-DFT calculations (ESI†). Both the d–d and LMCT transition energies appear to be significantly blue-shifted with respect to the experimental transitions. Exact transition energies are known to carry significant error when calculated in this way for open shell systems. Nevertheless, these calculations reveal that an intense absorption band is expected for the Cu(II)–azanide complex proposed as species 3 and its energy is predicted to be positioned between the  $\pi$ – $\pi^*$  and the d–d transitions. The three optimised structures for species 1, 2 and 3 are shown in Fig. 7.

## Conclusions

Herein, we establish three discrete protonation states for an AA10 LPMO within the pH range of 6.5 to 12.5, two of which are biologically relevant and may be important in the context of the LPMO mechanism. The first species exists at acidic–neutral pH and is believed to be consistent with a 5-coordinate ( $\text{N}_3\text{O}_2$ ), in keeping with the current understanding of the chitin-active AA10 coordination sphere. A second species is formed by deprotonating the ligating water molecule ( $\text{p}K_{\text{a}} 8.7$ ); yielding a near-axial EPR spectral envelope ( $g_1 = 2.044$ ,  $g_2 = 2.069$  and  $g_3 = 2.233$ ). We hypothesise that this species exists as a  $\text{N}_2\text{O}_2$ , 4-coordinate ligandfield with two hydroxide ligands and an elongated/decoordinated  $-\text{R-NH}_3^+$  following proton transfer from the coordinating hydroxide to the amino terminus. This gives rise to a rotated SOMO which is no longer defined by the chelating histidine brace. The third species formed at a high pH ( $\text{p}K_{\text{a}} \sim 11.5$ ) is proposed to exist as a 4-coordinate,  $\text{N}_3\text{O}$  complex, whereby species 2 has lost a water molecule (from the interaction of an  $\text{OH}^-$  and  $\text{NH}_3^+$ ) and the  $-\text{NH}_2$  group has been deprotonated to form a Cu(II)–NH azanide complex. The high pH at which this species becomes accessible is outside of that which is biologically relevant and, as such, is unlikely to be pertinent in catalysis.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We acknowledge the University of York for the support of PJJ's PhD studentship. We thank Z. Forsberg, G. Courtade, V. G. H. Eijsink and F. L. Aachmann for providing the clone/cell colonies and purification protocol for the B/AA10 enzyme. GJD thanks the Royal Society for the Ken Murray Research Professorship.

## Notes and references

1 G. Vaaje-Kolstad, B. Westereng, S. J. Horn, Z. Liu, H. Zhai, M. Sørli and V. G. H. Eijsink, *Science*, 2010, **330**, 219–222.

2 R. J. Quinlan, M. D. Sweeney, L. Lo Leggio, H. Otten, J.-C. N. Poulsen, K. S. Johansen, K. B. R. M. Krogh, C. I. Jørgensen, M. Tovborg,

- A. Anthonsen, T. Tryfona, C. P. Walter, P. Dupree, F. Xu, G. J. Davies and P. H. Walton, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 15079–15084.
- 3 C. M. Phillips, W. T. Beeson, J. H. Cate and M. A. Marletta, *ACS Chem. Biol.*, 2011, **6**, 1399–1406.
- 4 L. Ciano, G. J. Davies, W. B. Tolman and P. H. Walton, *Nat. Catal.*, 2018, **1**, 571–577.
- 5 G. R. Hemsworth, B. Henrissat, G. J. Davies and P. H. Walton, *Nat. Chem. Biol.*, 2014, **10**, 122–126.
- 6 V. V. Vu, W. T. Beeson, E. A. Span, E. R. Farquhar and M. A. Marletta, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 13822–13827.
- 7 M. Couturier, S. Ladevèze, G. Sulzenbacher, L. Ciano, M. Fanuel, C. Moreau, A. Villares, B. Cathala, F. Chaspoul, K. E. Frandsen, A. Labourel, I. Herpoël-Gimbert, S. Grisel, M. Haon, N. Lenfant, H. Rogniaux, D. Ropartz, G. J. Davies, M.-N. Rosso, P. H. Walton, B. Henrissat and J.-G. Berrin, *Nat. Chem. Biol.*, 2018, **14**, 306.
- 8 F. Sabbadin, G. R. Hemsworth, L. Ciano, B. Henrissat, P. Dupree, T. Tryfona, R. D. S. Marques, S. T. Sweeney, K. Besser, L. Elias, G. Pesante, Y. Li, A. A. Dowe, R. Bates, L. D. Gomez, R. Simister, G. J. Davies, P. H. Walton, N. C. Bruce and S. J. McQueen-Mason, *Nat. Commun.*, 2018, **9**, 756.
- 9 C. Filiatrault-Chastel, D. Navarro, M. Haon, S. Grisel, I. Herpoël-Gimbert, D. Chevret, M. Fanuel, B. Henrissat, S. Heiss-Blanquet, A. Margeot and J.-G. Berrin, *Biotechnol. Biofuels*, 2019, **12**, 55.
- 10 F. Sabbadin, S. Urresti, B. Henrissat, A. O. Avrova, L. R. J. Welsh, P. Lindley, M. Csukai, J. N. Squires, P. H. Walton, G. J. Davies, N. C. Bruce, S. C. Whisson and S. J. McQueen Mason, *Science*, 2021, **373**, 774–779.
- 11 G. R. Hemsworth, E. J. Taylor, R. Q. Kim, R. C. Gregory, S. J. Lewis, J. P. Turkenburg, A. J. Parkin, G. J. Davies and P. H. Walton, *J. Am. Chem. Soc.*, 2013, **135**, 6069–6077.
- 12 F. L. Aachmann, M. Sørli, G. Skjåk-Bræk, V. G. H. Eijsink and G. Vaaje-Kolstad, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 18779–18784.
- 13 N. Gagnon and W. B. Tolman, *Acc. Chem. Res.*, 2015, **48**, 2126–2131.
- 14 J.-P. Bacik, S. Mekasha, Z. Forsberg, A. Y. Kovalevsky, G. Vaaje-Kolstad, V. G. H. Eijsink, J. C. Nix, L. Coates, M. J. Cuneo, C. J. Unkefer and J. C. H. Chen, *Biochemistry*, 2017, **56**, 2529–2532.
- 15 G. Courtade, L. Ciano, A. Paradisi, P. J. Lindley, Z. Forsberg, M. Sørli, R. Wimmer, G. J. Davies, V. G. H. Eijsink, P. H. Walton and F. L. Aachmann, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 19178–19189.
- 16 G. Courtade, S. B. Le, G. I. Sætrom, T. Brautaset and F. L. Aachmann, *Carbohydr. Res.*, 2017, **448**, 212–219.
- 17 G. R. Hemsworth, L. Ciano, G. J. Davies and P. H. Walton, in *Methods Enzymol.*, ed. F. Armstrong, Academic Press, 2018, vol. 613, pp. 63–90.
- 18 B. Bissaro, I. Isaksen, G. Vaaje-Kolstad, V. G. H. Eijsink and Å. K. Røhr, *Biochemistry*, 2018, **57**, 1893–1906.
- 19 M. A. Hitchman and R. L. Belford, *Inorg. Chem.*, 1969, **8**, 958–965.
- 20 R. C. Gregory, G. R. Hemsworth, J. P. Turkenburg, S. J. Hart, P. H. Walton and G. J. Davies, *Dalton Trans.*, 2016, **45**, 16904–16912.
- 21 K. E. H. Frandsen, T. J. Simmons, P. Dupree, J.-C. N. Poulsen, G. R. Hemsworth, L. Ciano, E. M. Johnston, M. Tovborg, K. S. Johansen, P. von Freiesleben, L. Marmuse, S. Fort, S. Cottaz, H. Driguez, B. Henrissat,

- N. Lenfant, F. Tuna, A. Baldansuren, G. J. Davies, L. Lo Leggio and P. H. Walton, *Nat. Chem. Biol.*, 2016, **12**, 298.
- 22 J. Peisach and W. E. Blumberg, *Arch. Biochem. Biophys.*, 1974, **165**, 691–708.
- 23 B. Bissaro, B. Streit, I. Isaksen, V. G. H. Eijssink, G. T. Beckham, J. L. DuBois and Å. K. Røhr, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 1504.
- 24 A. Paradisi, M. J. Steward, P. Lindley, G. J. Davies and P. H. Walton, in *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*, Elsevier, 2020, DOI: 10.1016/B978-0-12-409547-2.14928-5.
- 25 M. J. Harms, C. A. Castañeda, J. L. Schlessman, G. R. Sue, D. G. Isom, B. R. Cannon and B. E. García-Moreno, *J. Mol. Biol.*, 2009, **389**, 34–47.
- 26 Y. A. Theibich, S. P. A. Sauer, L. L. Leggio and E. D. Hedegård, *Comput. Struct. Biotechnol. J.*, 2021, **19**, 555–567.
- 27 S. Stoll and A. Schweiger, *J. Magn. Reson.*, 2006, **178**, 42–55.



# Supplementary Information

## Contents

Figure S1. SDS PAGE Electrophoresis of <i>B/AA10</i> enzyme.....	2
Figure S2 Spin quantification by double-integration of EPR spectra of <i>B/AA10</i> collected at varying pH. ....	3
Table S1. Singular Values obtained from SVD analysis of the pH titration EPR spectra .....	4
Figure S3. Singular Value Decomposition (SVD) abstract EPR spectra .....	5
Figure S4. pH titration speciation plot .....	6
Figure S5 pH titration regression analysis to determine pKa1 and proton stoichiometry .....	7
Figure S6. UV-Visible spectrum of <i>B/AA10</i> ( <i>ca.</i> 10 $\mu$ M) at pH 12.5.....	8
DFT Geometry optimisations.....	9
Figure S7. DFT model used to describe species 1.....	10
Figure S8. Comparison on the DFT optimised coordinates of model 1 against the crystallographic coordinates .....	11
DFT EPR property calculations.....	11
Time Dependent-DFT (TD-DFT).....	13
DFT cartesian coordinates .....	14
References.....	21

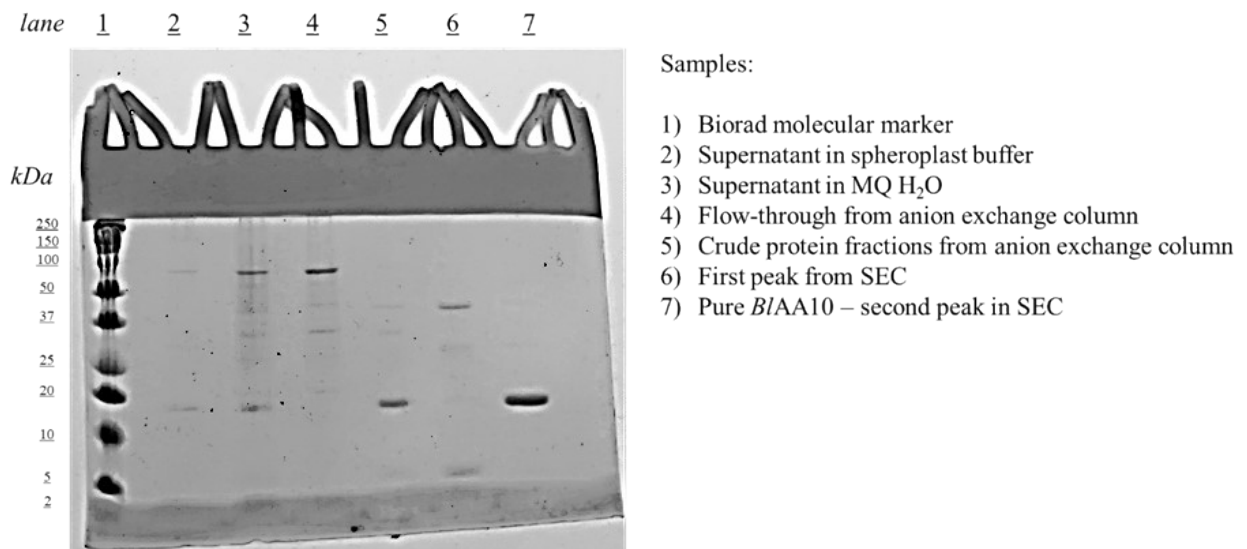


Figure S1. SDS PAGE Electrophoresis of *B/AA10* enzyme.

Gel electrophoresis was carried out using a 12% SDS-PAGE gel to confirm purity of *B/AA10* enzyme (lane 7) showing single band slightly under 20 kDa.

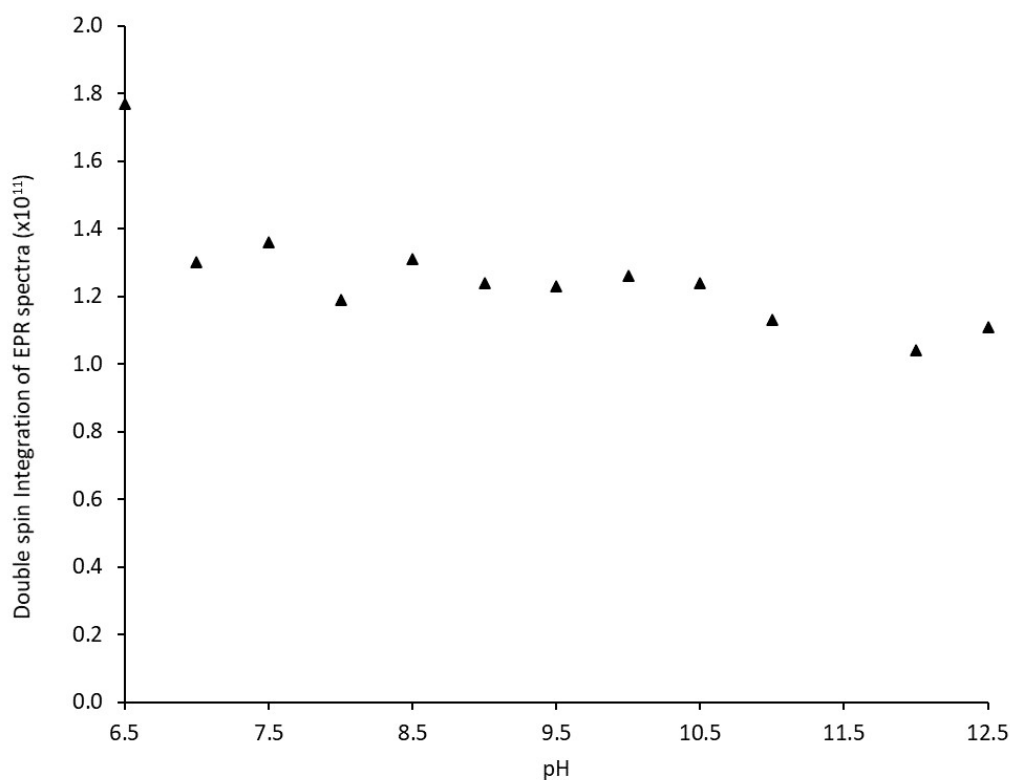


Figure S2 Spin quantification by double-integration of EPR spectra of *B/AA10* collected at varying pH.

The first derivative spectra were integrated twice with the maximum value of the resulting curve reporting on the total area under each absorption spectrum. Experimental conditions (such as microwave power) can affect signal intensity and as such were fixed between experiments. However, other factors such as slightly differing sample alignment in the EPR cavity along with variations in concentration due to altering pH were not able to be controlled for introducing some uncertainty in these values. Notwithstanding these caveats, however, the data shows a relatively consistent double-integration spin quantification (albeit with a slight downward trend likely due to diluting samples slightly with NaOH with increasing pH) indicating no redox processes occurred throughout the experiment.

Table S1. Singular Values obtained from SVD analysis of the pH titration EPR spectra

Principal Component	Singular Value
1	59
2	20
3	10

Singular Value Decomposition (SVD) was carried out in MATLAB R2020a by importing the EPR pH titration data set as a matrix (A) using the function:

$$A = \begin{bmatrix} a & \cdots & b \\ \vdots & \ddots & \vdots \\ y & \cdots & z \end{bmatrix}$$

Whereby the dataset is copied inside the square brackets.

A series of singular values were obtained using the functions:

$$s = svd(A)$$

$$[U,S,V] = svd(A)$$

The latter performs a singular value decomposition of matrix A, such that  $A = U \times S \times V'$ . It was possible to obtain abstract spectra from the U ( $m \times m$ ) matrix by the function:

$$G = U(1:m, n)$$

Where G contains the generated abstract spectrum, m is the rank of the U matrix (in this case 1024 due to the selected scan width of the EPR spectrum) and n is the selected principal component (in this case 1-3).

From the SVD analysis, three large non-zero singular values were obtained (Table S1) with their corresponding abstract spectra containing relatively little noise. In the absence of noiseless data, other non-zero singular values were generated. However, subsequent values were disregarded due to their comparatively small singular values and noisy abstract spectra. It was determined that there were 3 major principal components within the dataset.

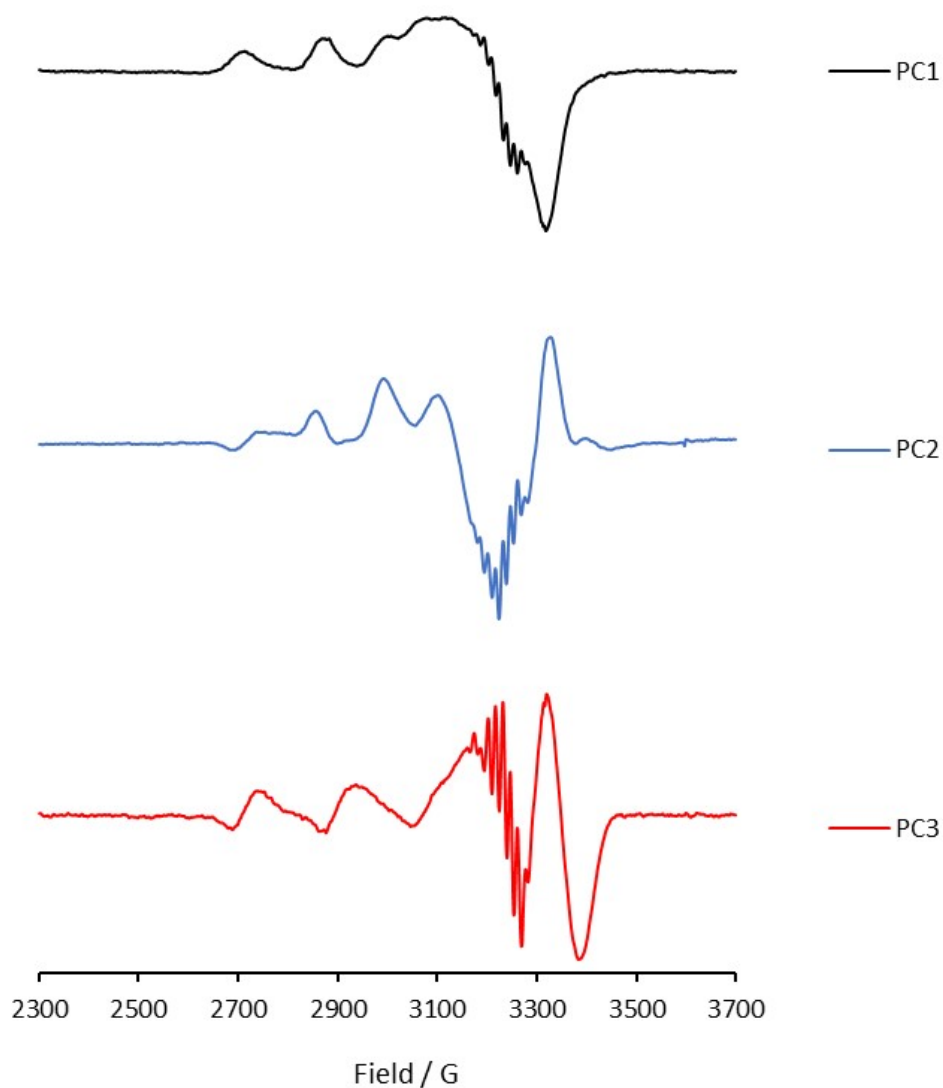


Figure S3. Singular Value Decomposition (SVD) abstract EPR spectra of the three large principal components.

PC3 (red) appears to somewhat resemble species **3** formed at high pH notable by the low field hyperfine peaks and overshoot feature. However, the abstract spectra of both PC1 and PC2 do not seem to match the spectra obtained for species **1** & **2** suggesting the SVD analysis struggled to separate these two species. This would likely improve with increasing number of data sets by performing the pH titration with smaller incremental steps.

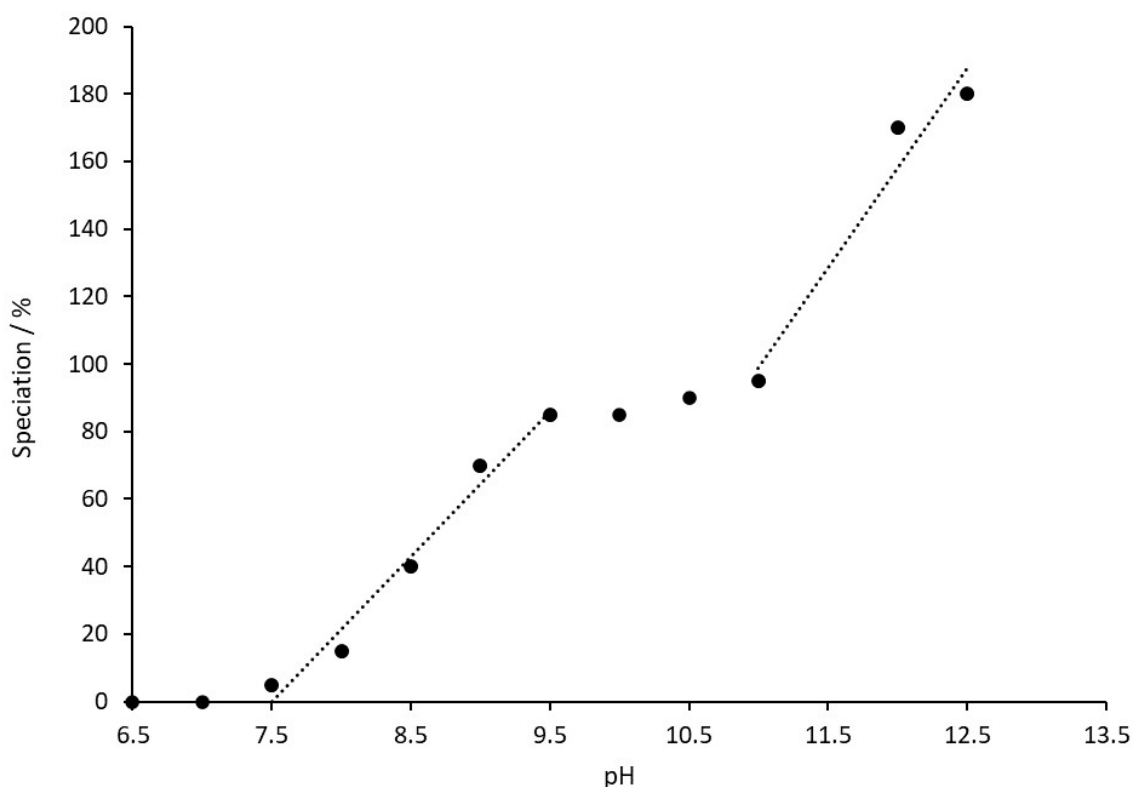


Figure S4. pH titration speciation plot. Speciation determined by comparative weighting of each species within the EPR spectra ranging from pH 6.5-12.5 with 0% representing all species 1, 100% being all species 2, and 200% being all species 3.

The speciation of the sample at each pH value was determined by assuming the spectrum collected at pH 6.5 describes 100% species **1**. Each spectrum was intensity-normalised to a maximum value of 1, to remove error from sample-cavity alignment between experiment and the slightly varying sample concentration from adjusting the pH. A percentage of the intensity of species **1** was then subtracted from each 'mixed' EPR spectrum and to determine the relative ratio of species **1**: species **2**. Speciation was determined when no shoulders corresponding to the 'subtracted spectrum' were present, yielding a flat baseline between peaks within the parallel region of the spectrum. By pH 10.5, there was no/negligible signal intensity corresponding to species **1** and the spectrum was assumed to be  $\approx$  100% species **2**. Subsequent speciation between pH 10.5-12.5 were derived using the same method, but now by subtracting species **3** intensity from the mixed spectra assuming the spectrum at pH 12.5  $\approx$  100% species **3**.

Assumptions were made on basis that each of the species at pH 6.5, 10.5 and 12.5 could be well described using a single set of spin Hamiltonian parameters in their respective simulations.

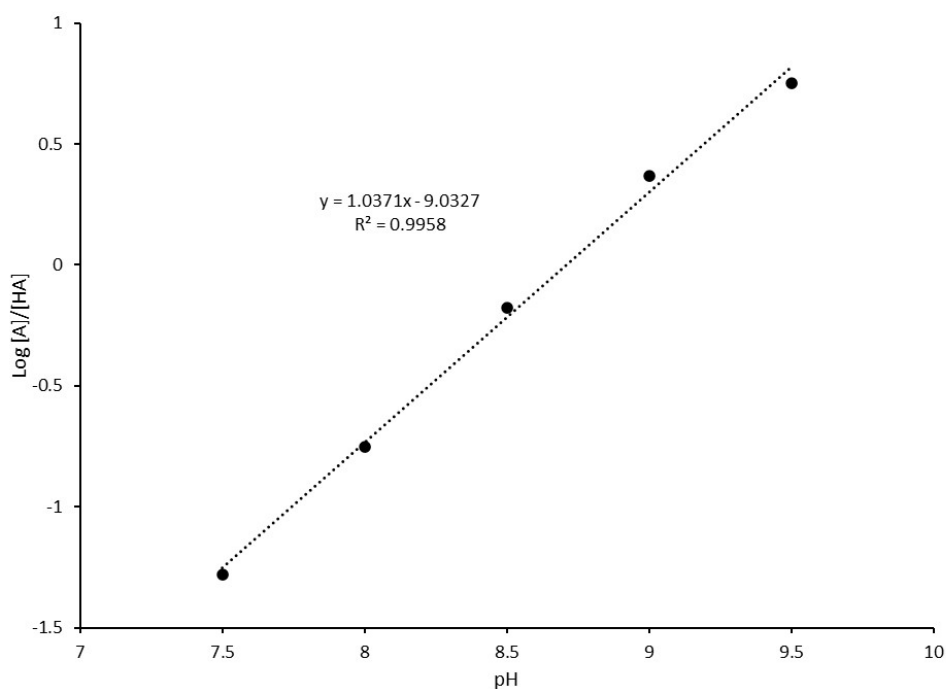


Figure S5 pH titration regression analysis to determine pKa1 and proton stoichiometry.

$$pH = pK_a - \log \frac{[A^-]}{[HA]}$$

At 50% of each species,  $-\log \frac{[A^-]}{[HA]} = 0, \therefore pH = pK_a$

$$\log \frac{[A^-][H^+]}{[HA]} = 1.037 \times pH - 9.033$$

At 50% of each species:

$$0 = 1.037 \times pH - 9.033$$

$$pH = pK_a = \frac{9.033}{1.037} = 8.71$$

The proton stoichiometry is reported by the gradient. Here, this is 1.037. Since the value needs to be an integer number, it was assumed a single proton was involved in the first deprotonation process.

The same process was performed on the second deprotonation step to find an approximate pKa2 of 11.5 and proton stoichiometry of 1.3. Again, this was adjudged to be a single deprotonation event. Due to the comparatively few data points in this part of the titration, the value of pKa2 is expected to carry more error than pKa1.

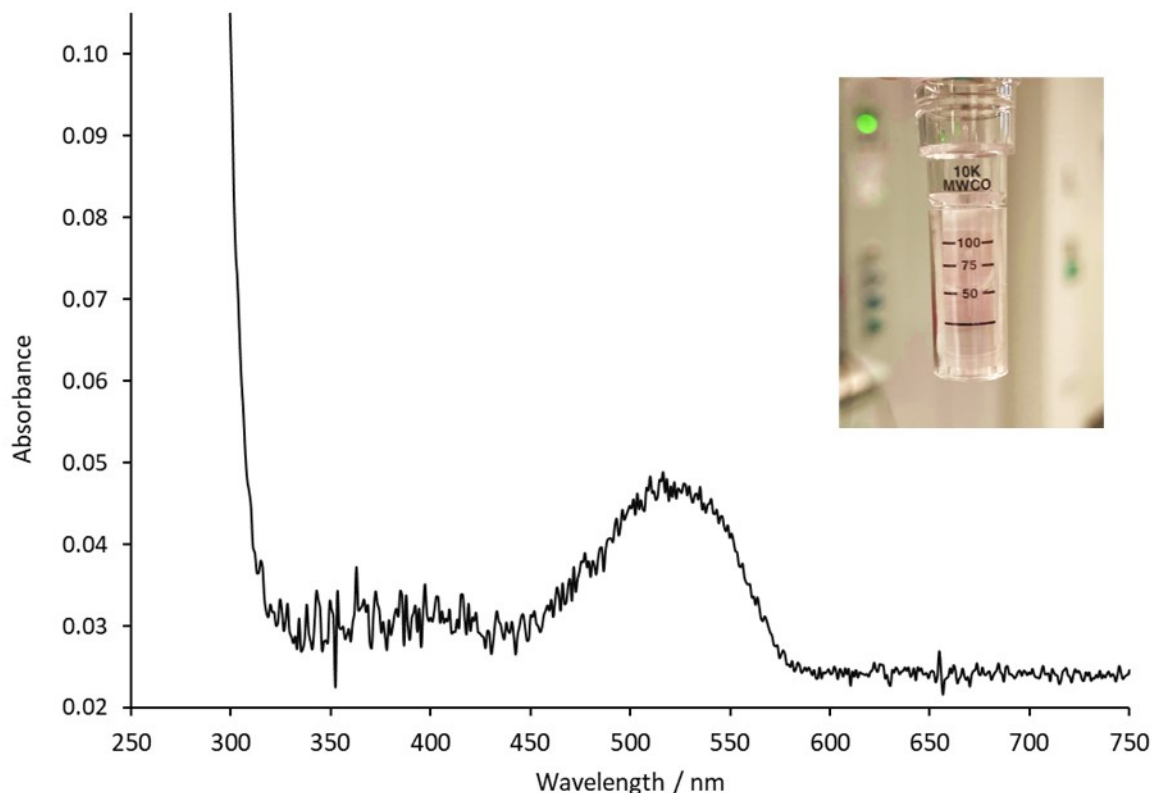


Figure S6. UV-Visible spectrum of *BIAA10* (ca. 10  $\mu\text{M}$ ) in MES, CHES, HEPES and CAPS, each at 5 mM (pH 12.5). Insert showing 10 kDa mass cut-off centrifugal filter containing *BIAA10* (ca. 10  $\mu\text{M}$ ) MES, CHES, HEPES and CAPS, each at 5 mM (pH 12.5) with a pale pink colour.

The addition of NaOH to the *BIAA10* enzyme afforded a colour change of the solution to from colourless to pale pink when approaching pH  $> 12$ . A UV/Visible spectrum was collected of *BIAA10* at a concentration of 10  $\mu\text{M}$  and pH 12.5. The absorption band centred at 520 nm has a molar absorptivity  $\epsilon \approx 2500 \text{ M}^{-1} \text{ cm}^{-1}$ . The absorption band is centred in the cyan/green region of the electromagnetic spectrum, giving rise to a complementary colour of red/pink; in accordance with the appearance of the sample in the centrifugal concentrator.



## DFT Geometry optimisations

Initial atomic coordinates to model species **1**, **2** and **3** were generated from the crystallographic structure of the similar enzyme *BaAA10* (PDB 5IJU, resolution 1.7 Å) from the Carbohydrate-Active enZymes (CAZy) database.<sup>1</sup> In all models the crystallographic coordinates were truncated to include the Cu(II) ion and 9 residues deemed to be key in the active site geometry: His28, Glu68, Gln92, Ala123, Pro124, His125, Thr127, Trp187 and Phe196 (numbering starting at the first histidine is position 28; His 28 and His 125 are analogous to His32 and His121 in BILPMO10A). Hydrogen atoms were added to appropriate positions as were not attainable from the crystallographic coordinate file. The following modifications were made to reduce the size of the computational models, thus decreasing computational cost/time:

His28 and His125 were truncated at the carbonyl carbon, which was replaced by a methyl group, Glu68 and Gln92 were truncated with methyl substitution of the C $\gamma$ , the nitrogen of the amide bond between Ala123 and Thr122 was replaced by methyl groups, Trp187 and Phe196 were truncated with methyl substitution of the C $\beta$ , Thr127 was truncated with methyl substitution of the C $\alpha$  and the methyl group of C $\beta$  was removed. For the species **1** model, two water molecules were retained with Cu-O distances fixed to 1.9 and 2.2 Å in keeping with the crystallographic coordinates. These Cu-O constraints were lifted for models **2** and **3**. Upon optimisation, the proton attached to the water nearest the Glu side chain was deprotonated by the carboxylate group. For the species **2** model, the proton shared by the ligating hydroxide and glutamate was removed and a proton from the additional water molecule was transferred to the -NH<sub>2</sub> group. For the species **3** model, the -NH<sub>3</sub><sup>+</sup> proton was transferred back to the hydroxide and an additional proton was removed from the -NH<sub>2</sub> group, giving an -NH- (azanide) group. By significantly truncating these models, certain structural constraints needed to be applied in order prevent atomic movement that would otherwise be impossible if the remainder of the proteins structure was imposed. In order to account for this, numerous atoms were kept frozen throughout the optimisations denoted by asterisks in figure S7.

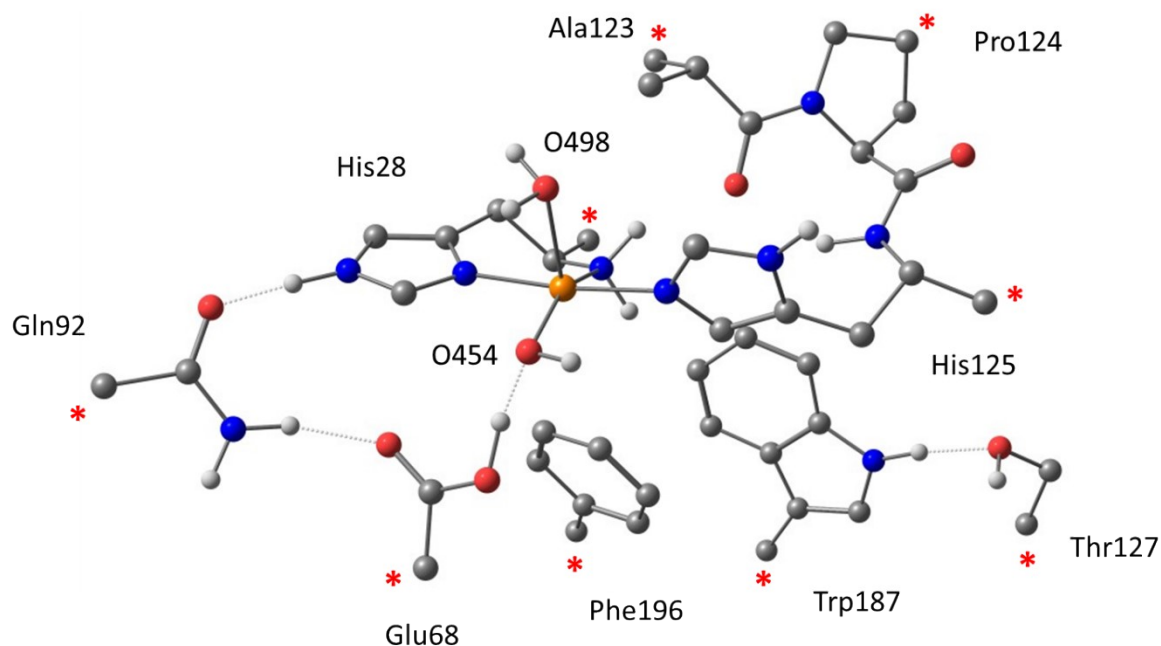


Figure S7. DFT model used to describe species 1. Residue numbers correspond to the labels from the crystal structure (PDB 5IJU) with His28 being the first amino acid in the polypeptide chain. Atoms kept frozen throughout the calculation are denoted by a red asterisk. Hydrogens connected to carbon atoms are omitted from the figure for clarity. Analogous positions were also frozen for models 2 and 3.

Geometry optimisations were performed using ORCA 4.2.0 program at the DFT level of theory. Optimisations of all models were performed using the spin-unrestricted, generalised gradient approximation (GGA) functional uBP86. Ahlrichs's Def-2-TZVP basis set was used to treat the copper and the first coordination sphere nitrogen atoms and oxygen atoms. On all remaining atoms a Def2-SVP basis set was used. Solvation effects were accounted for using the polarizable continuum model with water as the desired solvent (CPCM(water)), as implemented by ORCA 4.2.0. To assess the validity of the optimised geometries, the atomic positions were compared against the crystallographic coordinates, with the copper ion the origin of both models as in figure S8.

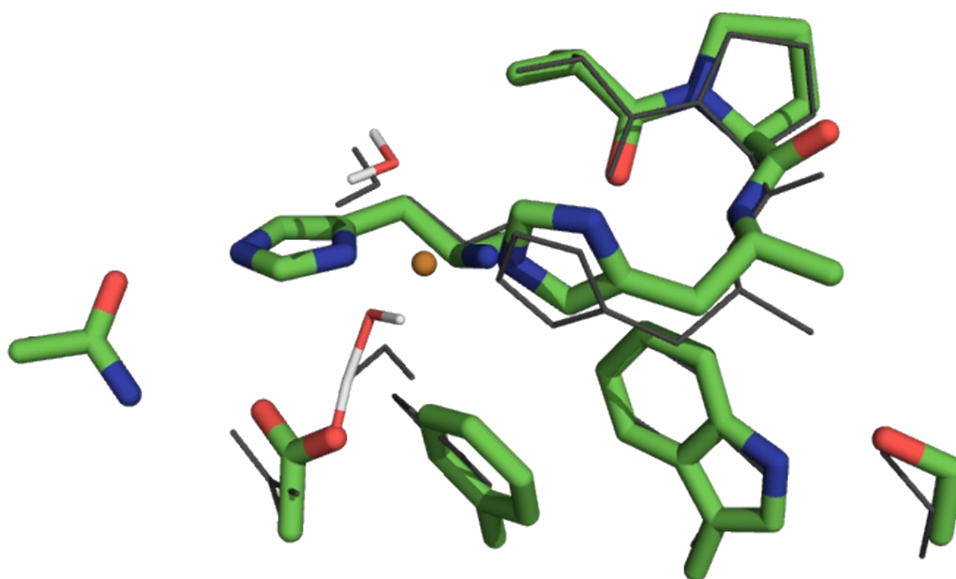


Figure S8. Comparison on the DFT optimised coordinates of model 1 (sticks coloured according to atom type) against the crystallographic coordinates of *BaAA10* (PDB 5IJU) (grey wires). Both models were aligned using the copper ion as the origin.

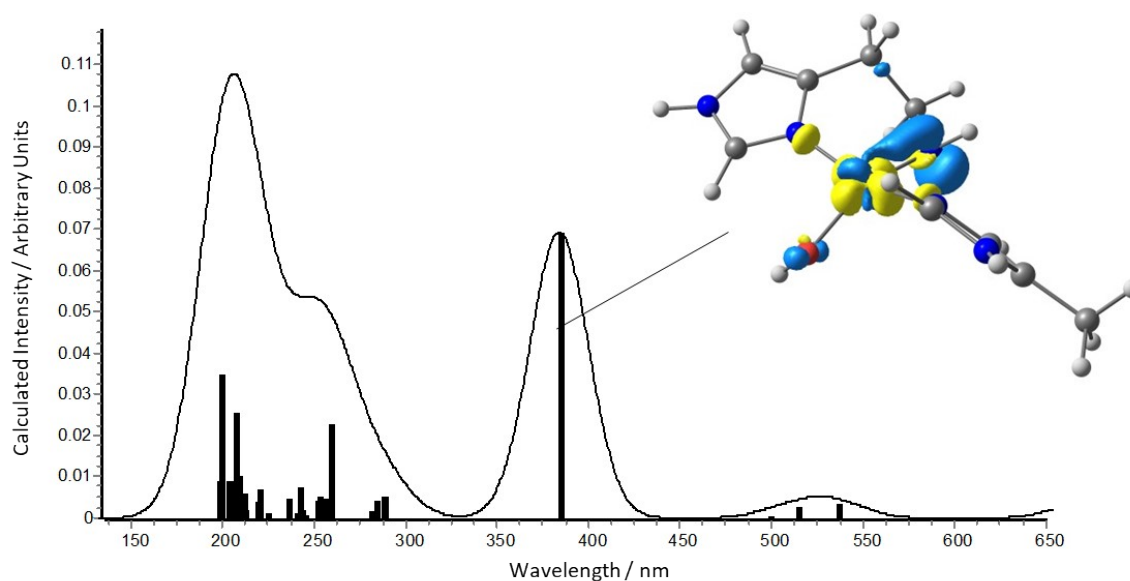
## DFT EPR property calculations

EPR property calculations were performed on the optimized geometries of **1**, **2** and **3** using the ORCA 4.2.0 program at the DFT level of theory. The cartesian reference system was oriented as such that the NH<sub>2</sub>-Cu(II)-O axis was aligned with the y-axis and the N-Cu(II)-N axis was oriented along x. The integration grid size was kept large (AngularGrid = 7 for all atoms and IntAcc = 7 for the Cu(II) ion) to ensure that the core density was correctly described. Solvation effects were accounted for in the property calculations by implementing the conductor-like polarized continuum model (CPCM) with a dielectric constant of 80.0 and a refractive index of 1.33 (water). The hyperfine coupling calculations included the Fermi-contact, spin dipolar and spin orbit contributions. The hybrid functional, uB3LYP, was used for these calculations with an adjusted fraction (38%) of Hartree-Fock exchange following studies showing improvements in the EPR property calculations.<sup>2</sup> The copper ions were described using the core-property “CP-PPP” basis set having shown good accuracy with EPR calculations on d-block metals.<sup>3</sup> The IGLO-III basis set was applied to all ligating atoms and the ring atoms of the imidazole groups due to increased flexibility in the core region, making it more suited for EPR properties calculations with respect to the Def2-TZVP basis set. All remaining atoms were treated with Ahlrich’s def2-SVP basis set.

**Table S2. EPR spin Hamiltonian parameters derived *via* simulation and DFT for species 1, 2 and 3.**

		Species 1		Species 2		Species 3	
		Simulation	DFT	Simulation	DFT	Simulation	DFT
<i>g</i> -factors	$g_1$	2.030	2.049	2.044	2.070	2.038	2.037
	$g_2$	2.125	2.081	2.069	2.079	2.062	2.070
	$g_3$	2.260	2.216	2.233	2.244	2.180	2.188
Cu HFC / MHz	$ A_1 $	60	110	50	94	70	15
	$ A_2 $	90	29	66	4	90	140
	$A_3$	-445	-598	-554	-582	-614	-496
Principal N SHFC / MHz	N1	40	36	N/A	1	33	43
	N2	40	47	40	40	35	44
	N3	32	44	40	40	35	39

## Time Dependent-DFT (TD-DFT)



**Figure S9. TD-DFT UV/vis spectrum of truncated model of species 3 and transition density of the proposed LMCT transition from the azanide (blue) to the metal (yellow). Spectrum generated by Gaussian broadening of the TD-DFT vectors.**

A TD-DFT calculation was performed on truncated coordinates of species **3** (model size reduced to lower computational cost) using the hybrid exchange-correlation functional, CAM-B3LYP, to provide more accurate energies of charge transfer excitations in which standard B3LYP significantly underestimates.<sup>4</sup> All atoms were treated with the def2-TZVP basis set with the inclusion of dispersion correction using the D3(BJ) scheme. Solvation effects were modelled using the conductor-like polarisable continuum model (CPCM) using water as the desired solvent. To improve the speed of these calculations, the RIJCOSX approximation was employed with the auxiliary basis set, def2/J. A total of 30 roots were calculated for the spectrum and a Gaussian broadening scheme was applied to produce the calculated spectrum. The vectors calculated at wavelengths < 300 nm showed to be  $\pi$ - $\pi^*$  transitions from the imidazole groups and would be expected to make up part of the large 280 nm absorption band synonymous with a protein's UV-Visible absorption profile. The vectors calculated at wavelengths > 500 nm belonged to d-d transitions (albeit considerably blue shifted). Finally, the vector calculated at 385 nm was considerably more intense than those in the d-d region and the transition density showed that this transition occurs via density from the azanide nitrogen to the copper as a formal LMCT transition (albeit significantly blue-shifted as with the d-d values). These bands appear to be considerably blue shifted compared to the experimental data. However, calculating exact energies with TD-DFT is known to be problematic in open-shell transition metal complexes due to heavy spin contaminations of excited states.<sup>5</sup> Notwithstanding this caveat, however, the calculation does reveal that for the proposed Cu(II)-azanide species that an intense LMCT band is expected to appear between the d-d region and the protein's characteristic band at 280 nm

# DFT cartesian coordinates

## Species 1

Single point energy uB3LYP (38%HF): -4415.903404286993 Hartree's

N	0.000000000	-2.106299000	0.000000000	C	-4.177704000	-0.674004000	0.065444000
C	1.285261000	-2.847226000	-0.152034000	C	-2.930088000	-0.819181000	-0.520022000
C	1.122078000	-4.350704000	0.041703000	N	-3.964911000	0.129792000	1.174747000
C	2.336038000	-2.287052000	0.820419000	C	-2.648652000	0.446659000	1.245175000
C	2.845544000	-0.947966000	0.407208000	N	-1.997756000	-0.113122000	0.222818000
C	4.138269000	-0.453258000	0.327257000	C	-8.957055000	-5.259243000	-3.500860000
N	1.982180000	0.060670000	0.000000000	C	-8.595048000	-5.747141000	-4.902956000
C	2.724471000	1.133715000	-0.290291000	O	-7.977147000	-4.360466000	-2.938994000
N	4.031509000	0.855908000	-0.104730000	C	-3.066742000	-5.816685000	-6.614897000
C	1.631917000	1.915757000	-5.298686000	C	-3.756518000	-5.390754000	-5.358724000
C	1.606374000	1.912440000	-3.778504000	C	-5.122775000	-5.268517000	-5.136719000
O	2.622213000	1.682122000	-3.103618000	C	-3.121457000	-5.032139000	-4.114885000
O	0.413896000	2.127533000	-3.263347000	C	-4.166982000	-4.718600000	-3.175313000
C	7.366924000	3.807846000	-2.235129000	C	-1.775105000	-4.921238000	-3.712727000
C	6.123790000	2.996081000	-1.944977000	N	-5.369813000	-4.874719000	-3.830946000
N	5.285393000	2.768311000	-2.972645000	C	-3.880113000	-4.342045000	-1.848153000
O	5.929856000	2.577606000	-0.777208000	C	-1.493909000	-4.497629000	-2.412103000
C	-1.237577000	-3.314669000	4.017779000	C	-2.530154000	-4.232100000	-1.478554000
C	-2.326742000	-3.718388000	3.015369000	C	0.654185000	-3.672477000	-6.185474000
O	-2.245557000	-3.386867000	1.806101000	C	0.174375000	-2.654386000	-5.179963000
C	-0.784796000	-1.872228000	3.752311000	C	-1.165963000	-2.215792000	-5.164509000
N	-3.370718000	-4.462920000	3.473447000	C	1.061682000	-2.130257000	-4.213566000
C	-4.372727000	-5.030393000	2.541718000	C	-1.613004000	-1.286190000	-4.211669000
C	-5.498029000	-4.004295000	2.244676000	C	0.623948000	-1.190203000	-3.268833000
O	-6.558380000	-3.991514000	2.885384000	C	-0.717473000	-0.766831000	-3.264858000
C	-4.908392000	-6.251818000	3.304949000	Cu	0.000000000	0.000000000	0.000000000
C	-4.868706000	-5.803219000	4.777582000	O	-0.062899000	1.700025000	-0.846144000
C	-3.566685000	-4.974428000	4.854900000	O	0.133766000	0.900513000	2.002794000
N	-5.231319000	-3.148614000	1.218215000	H	0.937252000	2.670457000	-5.713171000
C	-6.157546000	-2.073320000	0.828242000	H	1.286244000	0.917746000	-5.642701000
C	-7.522375000	-2.600893000	0.368845000	H	7.355898000	4.711777000	-1.594650000
C	-5.513642000	-1.231457000	-0.291537000	H	7.457620000	4.112504000	-3.294201000

H	5.501788000	3.155806000	-3.892369000	H	2.056037000	-4.878231000	-0.237294000
H	4.381152000	2.268454000	-2.883026000	H	3.186700000	-2.991328000	0.887664000
H	-0.029686000	-1.562411000	4.501599000	H	1.887069000	-2.248630000	1.839636000
H	-1.634247000	-1.162349000	3.795137000	H	5.102806000	-0.921532000	0.546896000
H	-0.325872000	-1.774204000	2.749604000	H	2.335610000	2.081927000	-0.660864000
H	-1.667611000	-3.358953000	5.038394000	H	-4.236695000	-3.072281000	0.949144000
H	-4.224551000	-7.108200000	3.136930000	H	-6.325517000	-1.421568000	1.717072000
H	-5.920548000	-6.537875000	2.964384000	H	-8.188473000	-1.744850000	0.140760000
H	-5.743952000	-5.160508000	4.990465000	H	-6.219383000	-0.411500000	-0.542515000
H	-4.859422000	-6.646591000	5.494132000	H	-5.400639000	-1.847848000	-1.205962000
H	-3.635754000	-4.138403000	5.578310000	H	-2.660073000	-1.388507000	-1.414484000
H	-2.702258000	-5.608840000	5.142787000	H	-4.679722000	0.440279000	1.838002000
H	-3.864934000	-5.294688000	1.592874000	H	-2.197250000	1.053982000	2.033197000
H	-9.018824000	-6.115604000	-2.798768000	H	-0.992232000	2.000355000	-0.826284000
H	-8.447209000	-4.895176000	-5.597785000	H	0.426401000	1.787814000	1.706695000
H	-7.674248000	-6.362188000	-4.888057000	H	0.902548000	0.568413000	2.507640000
H	-9.415336000	-6.376732000	-5.306071000	H	0.310483000	1.920003000	-2.186687000
H	-8.111062000	-3.479499000	-3.342798000	H	4.820253000	1.533606000	-0.351460000
H	-2.382209000	-5.028729000	-6.999499000	H	-6.320622000	-4.719417000	-3.431948000
H	-2.439634000	-6.720426000	-6.453806000	H	-3.792661000	-6.050094000	-7.418589000
H	-5.951177000	-5.437580000	-5.833694000	H	1.428961000	-3.245579000	-6.857440000
H	-0.959938000	-5.146987000	-4.413863000	H	8.255801000	3.213779000	-1.943812000
H	-0.443246000	-4.367808000	-2.125658000	H	2.659130000	2.080103000	-5.669793000
H	-4.683663000	-4.154429000	-1.120978000	H	-9.952172000	-4.762834000	-3.503324000
H	-2.285359000	-3.952182000	-0.442045000	H	0.908759000	-4.587282000	1.103907000
H	1.118293000	-4.546685000	-5.682197000	H	0.298392000	-4.772965000	-0.559476000
H	-0.176229000	-4.041652000	-6.818446000	C	-0.049322000	-4.304750000	3.955766000
H	-1.875583000	-2.626351000	-5.899595000	H	0.733518000	-4.005084000	4.679757000
H	2.112617000	-2.461984000	-4.206464000	H	-0.365215000	-5.340431000	4.192215000
H	-2.666568000	-0.966990000	-4.210047000	H	0.398585000	-4.309316000	2.942947000
H	1.335196000	-0.765472000	-2.544520000	H	-0.584157000	-2.303455000	-0.822216000
H	-1.045018000	-0.007958000	-2.540812000	H	-7.428870000	-3.218953000	-0.548658000
H	-0.554931000	-2.458342000	0.801229000	H	-7.988307000	-3.209102000	1.164688000
H	1.624751000	-2.647667000	-1.188349000				

## Species 2

Single point energy uB3LYP (38%HF): -4415.403254849613 Hartree's

N	28.395404000	2.179706000	75.635634000	C	27.478423000	-1.036725000	74.497215000
C	27.811571000	3.496969000	76.083355000	N	29.522981000	-1.834849000	74.353654000
C	28.544004000	4.658001000	75.421022000	C	29.313074000	-1.406179000	75.625907000
C	27.877057000	3.610779000	77.619221000	N	28.076110000	-0.925080000	75.744045000
C	26.841106000	2.891959000	78.435160000	C	26.540957000	0.049601000	66.553510000
C	25.971636000	3.512244000	79.320212000	C	25.237999000	0.404003000	65.839996000
N	26.586223000	1.515216000	78.492020000	O	26.347946000	-0.677501000	67.780772000
C	25.603850000	1.335639000	79.387336000	C	23.034995000	4.172974000	69.639914000
N	25.208756000	2.524186000	79.900007000	C	24.137895000	3.163669000	69.751562000
C	21.657999000	0.989999000	78.099997000	C	24.189391000	1.905975000	69.158240000
C	22.781287000	0.111559000	78.666636000	C	25.359927000	3.287999000	70.515736000
O	22.955070000	0.087729000	79.929129000	C	26.094261000	2.060769000	70.343405000
O	23.481864000	-0.524468000	77.823315000	C	25.900683000	4.301478000	71.340944000
C	23.453002000	2.661002000	84.412001000	N	25.351987000	1.243493000	69.514896000
C	23.396775000	2.354310000	82.921147000	C	27.330557000	1.839174000	70.982309000
N	22.892789000	1.168528000	82.534492000	C	27.126764000	4.076978000	71.976527000
O	23.834279000	3.197247000	82.097541000	C	27.832514000	2.855943000	71.808236000
C	32.489747000	1.804847000	75.165148000	C	22.458000000	4.792022000	73.872074000
C	31.675255000	1.557910000	73.896161000	C	23.282112000	3.643830000	74.411169000
O	30.414506000	1.557445000	73.879677000	C	23.923959000	2.736101000	73.541217000
C	32.067780000	0.869360000	76.301614000	C	23.448963000	3.459422000	75.802236000
N	32.378155000	1.385363000	72.756372000	C	24.713253000	1.684059000	74.039670000
C	31.724643000	1.364034000	71.427906000	C	24.214147000	2.396294000	76.304581000
C	31.283034000	-0.071723000	71.054050000	C	24.850022000	1.501623000	75.426774000
O	31.999896000	-0.830925000	70.387896000	Cu	27.373532000	-0.073807000	77.488236000
C	32.832591000	1.875332000	70.498405000	O	26.039517000	-1.364684000	77.732942000
C	34.104999000	1.306000000	71.147000000	O	29.157404000	0.728152000	77.652753000
C	33.853637000	1.457586000	72.643923000	H	21.566210000	0.876066000	77.004209000
N	30.046874000	-0.424785000	71.503088000	H	21.874603000	2.053791000	78.333413000
C	29.526651000	-1.790111000	71.319179000	H	24.516303000	2.680652000	84.726421000
C	29.291002000	-2.135003000	69.835995000	H	22.904017000	1.927588000	85.031403000
C	28.245528000	-1.969751000	72.154566000	H	22.560320000	0.499823000	83.230873000
C	28.378199000	-1.605172000	73.603843000	H	22.845357000	0.891254000	81.523605000



H	32.682306000	1.073007000	77.201581000	H	28.017445000	5.608303000	75.633908000
H	32.230087000	-0.191075000	76.017772000	H	27.762385000	4.684833000	77.864649000
H	31.003217000	0.989444000	76.581409000	H	28.906205000	3.344725000	77.944287000
H	33.550844000	1.593536000	74.940841000	H	25.849456000	4.570279000	79.574681000
H	32.839175000	2.984381000	70.517349000	H	25.129121000	0.385284000	79.652887000
H	32.687382000	1.539103000	69.455253000	H	29.638916000	0.179986000	72.231160000
H	34.200037000	0.236383000	70.879374000	H	30.303267000	-2.495578000	71.696741000
H	35.027473000	1.829041000	70.829644000	H	28.988305000	-3.197899000	69.750386000
H	34.337181000	0.660917000	73.241403000	H	27.935603000	-3.032530000	72.057203000
H	34.210710000	2.440765000	73.014921000	H	27.419484000	-1.367878000	71.721937000
H	30.833891000	2.022218000	71.461282000	H	26.444151000	-0.728851000	74.321006000
H	27.088553000	0.970112000	66.842448000	H	30.388044000	-2.262549000	74.013894000
H	24.672278000	-0.510856000	65.565065000	H	30.062319000	-1.440198000	76.420353000
H	24.590562000	1.039443000	66.476992000	H	29.152607000	1.272858000	78.463150000
H	25.455411000	0.960353000	64.905187000	H	25.127761000	-0.980657000	77.872525000
H	25.884607000	-1.514262000	67.575312000	H	24.513957000	2.690804000	80.687925000
H	22.522375000	4.336690000	70.613517000	H	25.690415000	0.382006000	69.035129000
H	23.416410000	5.165036000	69.316621000	H	22.266543000	3.850630000	68.910179000
H	23.458572000	1.422085000	68.498050000	H	21.691283000	5.123910000	74.599774000
H	25.357292000	5.247262000	71.491094000	H	23.042389000	3.675027000	84.583167000
H	27.550754000	4.863850000	72.617706000	H	20.690844000	0.749236000	78.589118000
H	27.891257000	0.905744000	70.831282000	H	27.208815000	-0.531015000	65.877193000
H	28.788041000	2.701363000	72.333118000	H	29.580073000	4.748419000	75.806468000
H	23.101758000	5.670104000	73.646515000	H	28.594094000	4.527621000	74.324168000
H	21.947272000	4.516931000	72.927033000	C	32.394498000	3.296120000	75.535720000
H	23.808229000	2.848429000	72.453431000	H	33.028117000	3.510175000	76.418980000
H	22.951190000	4.149748000	76.502727000	H	32.728442000	3.946734000	74.702290000
H	25.207764000	1.002954000	73.328386000	H	31.352248000	3.576718000	75.789870000
H	24.302122000	2.242548000	77.386927000	H	27.683340000	1.613456000	75.160184000
H	25.408985000	0.650036000	75.848803000	H	28.487218000	-1.510256000	69.393795000
H	29.207308000	2.225705000	74.970940000	H	30.224161000	-1.988242000	69.263890000
H	26.751032000	3.492615000	75.764815000	H	28.775015000	1.575655000	76.515717000

### Species 3

Single point energy uB3LYP (38%HFX): -4414.815119531047 Hartree's

N	0.000000000	-1.938135000	0.000000000	C	-4.064333000	-0.154887000	0.738706000
C	1.240492000	-2.690738000	-0.046188000	C	-2.857416000	-0.647698000	0.265013000
C	0.991023000	-4.176505000	0.233300000	N	-3.721635000	0.836122000	1.649364000
C	2.342901000	-2.132572000	0.910903000	C	-2.361263000	0.935660000	1.686066000
C	2.851549000	-0.818822000	0.415186000	N	-1.823280000	0.046347000	0.859639000
C	4.143345000	-0.380175000	0.153431000	C	-8.817712000	-4.710922000	-2.547817000
N	1.974356000	0.176907000	0.000000000	C	-9.113530000	-5.117550000	-3.989370000
C	2.708427000	1.176426000	-0.500934000	O	-8.271811000	-3.385129000	-2.426137000
N	4.024367000	0.873710000	-0.421514000	C	-3.732152000	-5.557309000	-6.074339000
C	1.483186000	1.913459000	-5.309096000	C	-4.325528000	-4.899112000	-4.867840000
C	1.471590000	2.332830000	-3.824413000	C	-5.539046000	-4.225614000	-4.769100000
O	2.591557000	2.440757000	-3.221391000	C	-3.729951000	-4.836900000	-3.553122000
O	0.342767000	2.510674000	-3.283615000	C	-4.639379000	-4.111371000	-2.708250000
C	7.518436000	3.538220000	-2.703036000	C	-2.512845000	-5.309997000	-3.012327000
C	6.140784000	2.966173000	-2.429703000	N	-5.724456000	-3.745451000	-3.481271000
N	5.296969000	2.849261000	-3.469503000	C	-4.350104000	-3.861248000	-1.351963000
O	5.853123000	2.631152000	-1.253449000	C	-2.228093000	-5.052378000	-1.668275000
C	-1.038899000	-2.919202000	4.271800000	C	-3.129215000	-4.327946000	-0.846127000
C	-2.208289000	-3.341004000	3.370546000	C	0.126307000	-3.627455000	-5.964013000
O	-2.192215000	-3.122714000	2.140317000	C	-0.217743000	-2.598341000	-4.912058000
C	-0.553702000	-1.521896000	3.856943000	C	-1.560813000	-2.378786000	-4.536352000
N	-3.269396000	-3.969821000	3.951323000	C	0.787244000	-1.839246000	-4.271768000
C	-4.365981000	-4.520263000	3.121763000	C	-1.890265000	-1.422362000	-3.562100000
C	-5.463392000	-3.452215000	2.870872000	C	0.459570000	-0.875940000	-3.303054000
O	-6.500253000	-3.408448000	3.551128000	C	-0.882150000	-0.658452000	-2.952457000
C	-4.894465000	-5.679638000	3.975652000	Cu	0.000000000	0.000000000	0.000000000
C	-4.725433000	-5.152697000	5.409803000	O	-0.389619000	1.874310000	-0.650852000
C	-3.384013000	-4.400057000	5.362755000	H	0.550030000	2.226426000	-5.817112000
N	-5.207199000	-2.587642000	1.851623000	H	1.536464000	0.805043000	-5.355475000
C	-6.100529000	-1.457627000	1.551199000	H	7.654747000	4.444451000	-2.079134000
C	-7.494176000	-1.909759000	1.102709000	H	7.688901000	3.797726000	-3.765099000
C	-5.465412000	-0.574275000	0.457075000	H	5.604658000	3.174431000	-4.387864000

H	4.280394000	2.581487000	-3.372037000	H	1.924176000	-4.777454000	0.190213000
H	0.311963000	-1.215284000	4.479204000	H	3.195674000	-2.839555000	0.988467000
H	-1.355761000	-0.765935000	3.976918000	H	1.898845000	-2.039668000	1.927289000
H	-0.252491000	-1.509390000	2.789921000	H	5.116617000	-0.852523000	0.324638000
H	-1.401696000	-2.869991000	5.318372000	H	2.324659000	2.061717000	-1.017100000
H	-4.261068000	-6.574421000	3.805384000	H	-4.230768000	-2.556788000	1.512686000
H	-5.941606000	-5.933412000	3.727808000	H	-6.210929000	-0.853866000	2.481542000
H	-5.549723000	-4.450332000	5.638364000	H	-8.143503000	-1.021101000	0.967821000
H	-4.721273000	-5.952270000	6.175375000	H	-6.121424000	0.309448000	0.316197000
H	-3.358555000	-3.530633000	6.049707000	H	-5.472005000	-1.129619000	-0.502938000
H	-2.537215000	-5.069246000	5.624819000	H	-2.654324000	-1.447706000	-0.455915000
H	-3.946991000	-4.842120000	2.147539000	H	-4.375997000	1.410045000	2.186303000
H	-8.055818000	-5.383307000	-2.102617000	H	-1.812156000	1.669356000	2.283689000
H	-9.857761000	-4.434848000	-4.450964000	H	0.054777000	2.071978000	-1.516242000
H	-8.193854000	-5.104735000	-4.607739000	H	4.786897000	1.523657000	-0.769602000
H	-9.534060000	-6.143800000	-4.017907000	H	-6.632934000	-3.392624000	-3.114417000
H	-8.922749000	-2.750850000	-2.788232000	H	-4.390876000	-5.458577000	-6.959625000
H	-2.745623000	-5.117742000	-6.338835000	H	-0.420005000	-3.438614000	-6.912280000
H	-3.553378000	-6.641411000	-5.905637000	H	8.281071000	2.803349000	-2.376889000
H	-6.295730000	-4.042171000	-5.542418000	H	2.365556000	2.320166000	-5.842874000
H	-1.798987000	-5.861680000	-3.644161000	H	-9.736073000	-4.796326000	-1.923205000
H	-1.282736000	-5.410861000	-1.236253000	H	0.537940000	-4.306988000	1.237282000
H	-5.051950000	-3.317249000	-0.704895000	H	0.274692000	-4.579903000	-0.505884000
H	-2.849589000	-4.106025000	0.194539000	C	0.102311000	-3.963466000	4.215665000
H	1.210489000	-3.633960000	-6.192105000	H	0.922595000	-3.662257000	4.897552000
H	-0.155268000	-4.651081000	-5.632752000	H	-0.246039000	-4.971966000	4.517911000
H	-2.361878000	-2.970372000	-5.005889000	H	0.512567000	-4.035397000	3.189035000
H	1.843162000	-2.004657000	-4.544836000	H	-7.452517000	-2.451052000	0.134214000
H	-2.943961000	-1.285028000	-3.271573000	H	-7.952216000	-2.566013000	1.864306000
H	1.248905000	-0.289840000	-2.809556000	O	0.644523000	3.026240000	1.350747000
H	-1.122220000	0.112064000	-2.203728000	H	1.280061000	2.340145000	1.630071000
H	-0.564587000	-2.291138000	0.786724000	H	0.217768000	2.566455000	0.475477000
H	1.657773000	-2.596930000	-1.079636000				

### Truncated Species 3 for TD-DFT

Single point energy uB3LYP (38%HFx): -2341.605057235535 Hartree's

N	0.000000000	-1.938134958	0.000000000	H	1.657773148	-2.596929941	-1.079635890
C	1.240492015	-2.690738250	-0.046188178	H	3.195674419	-2.839555064	0.988467117
C	2.342901048	-2.132572142	0.910902962	H	1.898845135	-2.039668197	1.927288941
C	2.851549395	-0.818821887	0.415186119	H	5.116617466	-0.852523069	0.324638071
C	4.143345145	-0.380175223	0.153431239	H	2.324659250	2.061716896	-1.017099840
N	1.974356083	0.176907129	0.000000000	H	-6.121424519	0.309448038	0.316197165
C	2.708427058	1.176426113	-0.500934001	H	-5.472005498	-1.129619327	-0.502937995
N	4.024367049	0.873710268	-0.421514021	H	-2.654323976	-1.447706184	-0.455914776
C	-5.465412479	-0.574274793	0.457075261	H	-4.375997393	1.410045169	2.186303213
C	-4.064333161	-0.154887006	0.738706039	H	-1.812156376	1.669356308	2.283689289
C	-2.857416383	-0.647698136	0.265013025	H	0.054777254	2.071978172	-1.516242046
N	-3.721635328	0.836122279	1.649363992	H	4.786897177	1.523656878	-0.769602053
C	-2.361262970	0.935659990	1.686066138	H	1.004395352	-3.702462707	0.248891111
N	-1.823280211	0.046346931	0.859638916	H	-5.819682889	-1.169608194	1.285610726
Cu	0.000000000	0.000000000	0.000000000				
O	-0.389618921	1.874310361	-0.650852032				
H	-0.564587145	-2.291137988	0.786724112				

## References

1. R. C. Gregory, G. R. Hemsworth, J. P. Turkenburg, S. J. Hart, P. H. Walton and G. J. Davies, *Dalton Transactions*, 2016, **45**, 16904-16912.
2. F. Neese, *Magnetic Resonance in Chemistry*, 2004, **42**, S187-S198.
3. E. D. Hedegård, J. Kongsted and S. P. A. Sauer, *Physical Chemistry Chemical Physics*, 2012, **14**, 10669-10676.
4. T. Yanai, D. P. Tew and N. C. Handy, *Chemical Physics Letters*, 2004, **393**, 51-57.
5. B. Suo, K. Shen, Z. Li and W. Liu, *The Journal of Physical Chemistry A*, 2017, **121**, 3929-3942.

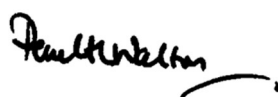
# 4 Mechanistic basis of substrate–O<sub>2</sub> coupling within a chitin-active lytic polysaccharide monooxygenase: An integrated NMR/EPR study

## Declaration

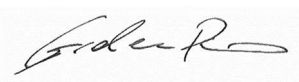
The following research paper describes the substrate-induced perturbations to the active site electronic structure of a bacterial AA10 LPMO, *B/AA10*. This paper contains a large and varied dataset that altogether describe the overall geometric structure and changes to the active site structure of the *B/AA10* enzyme following the binding of its native polysaccharide substrate. Many authors contributed to this work whose contributions are as follows:

Gaston Courtade, Luisa Ciano, Zara Forsberg, Vincent Eijsink, Paul Walton and Finn Achmann conceived the original research ideas; Gaston Courtade provided original enzyme samples/cell colonies/purification protocols for the *B/AA10* enzyme; Gaston Courtade and I produced *B/AA10* enzyme samples for analysis; Finn Achmann, Zara Forsberg and co-workers performed and characterised the NMR analyses of the <sup>15</sup>N-*B/AA10* enzyme; Luisa Ciano performed the CW-EPR experiments; Luisa Ciano and I contributed to the simulation analysis of the *B/AA10* CW-EPR spectra; I conceived and performed EPR and spin density DFT calculations; I identified a potential H-bond interaction between chitin N-acetyl group and active site that was later confirmed by HYSORE EPR experimentation; Paul Walton and Alessandro Paradisi developed the ligand field theory analysis of the *B/AA10* enzyme and devised the novel DELFT methodology; The EPSRC EPR National Service in Manchester executed the HYSORE EPR experiments. Luisa Ciano performed the HYSORE EPR simulations; Gaston Courtade, Luisa Ciano, Alessandro Paradisi, Zara Forsberg, Morten Sorlie, Gideon Davies, Vincent Eijsink, Paul Walton and Finn Achmann and I wrote the paper. All authors contributed to the final manuscript.

Signed Prof Paul H. Walton



Signed Prof Gideon J. Davies



Signed Dr Alison Parkin



# Mechanistic basis of substrate-O<sub>2</sub> coupling within a chitin-active lytic polysaccharide monooxygenase: an integrated NMR/EPR study

Gaston Courtade<sup>a,1</sup>, Luisa Ciano<sup>b,c,1,2</sup>, Alessandro Paradisi<sup>b</sup>, Peter Lindley<sup>b</sup>, Zarah Forsberg<sup>d</sup>, Morten Sørli<sup>d</sup>, Reinhard Wimmer<sup>e</sup>, Gideon J. Davies<sup>b</sup>, Vincent G. H. Eijssink<sup>d</sup>, Paul H. Walton<sup>b,3</sup>, Finn L. Aachmann<sup>a,3</sup>

<sup>1</sup> these authors contributed equally to this work.

## Author affiliation

<sup>a</sup> NOBIPOL, Department of Biotechnology and Food Science, NTNU Norwegian University of Science and Technology, Sem Sælands vei 6/8, N-7491 Trondheim, Norway

<sup>b</sup> Department of Chemistry, University of York, Heslington, York YO10 5DD, UK

<sup>c</sup> School of Chemistry and Photon Science Institute, University of Manchester, Oxford Road, Manchester, M13 9PL, UK

<sup>d</sup> Faculty of Chemistry, Biotechnology and Food Science, NMBU Norwegian University of Life Sciences, N-1432 Ås, Norway

<sup>e</sup> Department of Chemistry and Bioscience, Aalborg University, Frederik Bajers vej 7H, 9220 Aalborg Ø, Denmark

<sup>2</sup> Current address: School of Chemistry, University of Nottingham, University Park, Nottingham, NG7 2RD, UK

## Corresponding authors

<sup>3</sup> To whom correspondence should be addressed.

Finn L. Aachmann

NOBIPOL, Department of Biotechnology and Food Science, NTNU Norwegian University of Science and Technology, Sem Sælands vei 6/8, N-7491 Trondheim, Norway

Tel.: +47 73593317

Paul H. Walton

Department of Chemistry, University of York, Heslington, York YO10 5DD, UK

Tel: +44 1904 324457

E-mail: [finn.l.aachmann@ntnu.no](mailto:finn.l.aachmann@ntnu.no), [paul.walton@york.ac.uk](mailto:paul.walton@york.ac.uk)

## Keywords

lytic polysaccharide monooxygenase, LPMO, AA10, chitin, copper, EPR, NMR

## Abstract

Lytic polysaccharide monooxygenases (LPMOs) have a unique ability to activate molecular oxygen for subsequent oxidative cleavage of glycosidic bonds. To provide insight into the mode of action of these industrially important enzymes, we have performed an integrated NMR/EPR study into the detailed aspects of an AA10 LPMO-substrate interaction. Using NMR spectroscopy, we have elucidated the solution-phase structure of *apo-B/LPMO10A* from *Bacillus licheniformis*, along with the first solution-phase structural characterization of the Cu(I)-LPMO, showing that the presence of the metal has minimal effects on the overall protein structure. We have, moreover, used paramagnetic relaxation enhancement (PRE) to characterize the Cu(II)-LPMO by NMR spectroscopy. In addition, a multi-frequency continuous wave (CW) EPR and  $^{15}\text{N}$ -HYSCORE spectroscopy study on the uniformly isotope-labeled  $^{63}\text{Cu(II)}$ -bound  $^{15}\text{N-B/LPMO10A}$  along with its natural abundance isotopologue determined copper spin-Hamiltonian parameters for LPMOs to markedly improved accuracy. The data demonstrate that large changes in the Cu(II) spin-Hamiltonian parameters are induced upon binding of the substrate. These changes arise from a rearrangement of the copper coordination sphere from a five-coordinate distorted square pyramid to one which is four-coordinate near-square planar. There is also a small reduction in metal-ligand covalency and an attendant increase in the  $d(x^2-y^2)$  character/energy of the singly-occupied molecular orbital (SOMO), which we propose from DFT calculations predisposes the copper active site for the formation of a stable Cu-O<sub>2</sub> intermediate. This switch in orbital character upon addition of chitin provides a basis for understanding the coupling of substrate-binding with O<sub>2</sub> activation in chitin-active AA10 LPMOs.



## **Significance statement**

Lytic polysaccharide monooxygenases (LPMOs) have unique catalytic centers, at which a single copper catalyzes the oxidative cleavage of a glycosidic bond. The mechanism by which LPMOs activate molecular oxygen is key to understanding copper (bio)catalysis, but remains poorly understood, largely because the insoluble and heterogeneous nature of LPMO substrates precludes the use of usual laboratory techniques. Using an integrated NMR/EPR approach, we have unraveled structural and electronic details of the interactions of an LPMO from *Bacillus licheniformis* and  $\beta$ -chitin. EPR spectroscopy on uniformly isotope  $^{15}\text{N}$ -labeled  $^{63}\text{Cu(II)}$ -LPMO provided new insight into substrate-driven rearrangement of the copper coordination sphere that predisposes the enzyme for  $\text{O}_2$ -activation.

## Background

The sustainable use of polysaccharides from lignocellulosic biomass as a feedstock in the production of biofuels and biomaterials is key to reducing dependency on fossil fuels. In this regard, chitin, an abundant insoluble polysaccharide found in the exoskeletons of arthropods and the cell walls of fungi, has often been proposed as a potential feedstock for conversion into high-value biomaterials (1-3). Given this impetus, the efficient processing of chitin through enzymatic breakdown into its constituent sugars is an attractive means of realizing its full chemical and calorific potential, and indeed that of other polysaccharides such as cellulose. It is unsurprising therefore that the commercial use of enzyme cocktails for this purpose is widespread, where the content of these cocktails includes a range of glycoside hydrolases (GHs) and, more recently, copper-dependent redox enzymes known as lytic polysaccharide monooxygenases (LPMOs).

LPMOs are currently classified as auxiliary activity (AA) families 9–11 and 13–16 in the CAZy database (4, 5). They have been shown to augment dramatically the activity of GHs, probably by reducing the crystallinity of their substrates (6-12), and now, alongside GH enzymes, are seen as key components in the efficient processing of abundant biomass. Accordingly, there is much interest in increasing the efficiency of LPMOs through a deeper understanding of their molecular and electronic features. From previous studies it is known that LPMOs have an oxidative mode of action on their substrates. This oxidation proceeds through hydrogen-atom abstraction from either the C1 or C4 carbon in 1-4 linked polysaccharides such as cellulose or chitin (8, 9, 13, 14) to generate the respective hydroxylated product, from which elimination leads to cleavage of the glycosidic bond. The mechanism likely involves the formation of a copper-bound reactive oxygen species that arises from the reaction of the copper active site of the LPMO with O<sub>2</sub> and a reducing agent, or with hydrogen peroxide (9, 14-18).

In the context of the O<sub>2</sub> mechanism there is consensus regarding the initial step of the catalytic cycle in which Cu(I)-LPMO reacts with O<sub>2</sub> to give a Cu(II) and superoxide (14, 19). Any Cu(II)-superoxide complex formed in this manner may then oxidise the substrate or, with the addition of further electrons and protons, go on to form ‘high valent’ copper-oxygen intermediates, which have not yet been observed experimentally in LPMOs. Proposals for these intermediates, however, have come from computational studies and include Cu(II)-oxyl and Cu(III)-hydroxide (19-25). An

equivalent outcome is achieved by the direct reaction of Cu(I)-LPMO with H<sub>2</sub>O<sub>2</sub>. Indeed, DFT calculations on AA9 LPMOs have shown that the O<sub>2</sub> and H<sub>2</sub>O<sub>2</sub> reaction pathways converge on a common intermediate (26, 27). Recent work on the mechanism of H<sub>2</sub>O<sub>2</sub>-driven LPMO reactions has led to suggestions that hydroxyl radicals could also play a role (15, 26, 28). If such species are the oxidants or, indeed part of an overall catalytic cycle, then the role of the substrate is essential (15). For instance, in the presence of substrate, it has been shown by DFT calculations for both AA9 LPMOs (26) and, more recently AA10 LPMOs (28), that any hydroxyl generated in this manner is re-directed by active site residues back towards the copper ion. This redirection forms a protein-bound Cu(II)-oxyl species, which then acts as the key oxidizing intermediate in the catalytic cycle, thus avoiding the deleterious oxidative effects of the hydroxyl intermediate.

In this context, a major question facing the LPMO world is what mechanisms are employed by the enzymes to couple the presence of substrate to the generation of oxidising intermediates, thus avoiding the deleterious oxidation of the protein? Following initial proposals on the mechanism of LPMOs (8, 19) and the importance of considering the substrate's role, a growing number of studies are indeed now showing that substrate binding and LPMO catalysis are coupled (18, 22, 29-32). For instance, both Borisova *et al.* (29) and Frandsen *et al.* (22) demonstrated that the spin-Hamiltonian parameters of the Cu(II) in some AA9 LPMOs shift significantly upon the LPMO binding to the substrate, potentially indicating a “trigger” mechanism in which the substrate regulates the catalytic mechanism and protects from inactivation pathways (33). For family AA10 LPMOs, the dissociation constants ( $K_d$ ) of copper binding have been determined to be 6–55 nM for Cu(II) and approximately 1 nM for Cu(I) (34, 35). In this respect, Kracher *et al.* (30) linked the oxidation state of the copper to binding affinity for the substrate. More recently, based on modeling studies, Bissaro *et al.* (32) showed that chitin-binding to *Sm*LPMO10A results in a constrained copper site geometry that includes a tunnel through which small co-substrates could diffuse in the presence of substrate. It has also been shown that the presence of substrate enhances the stability of LPMOs (15, 33, 36, 37). Evidently, whether an LPMO is bound to its polysaccharide substrate not only affects the stability of the LPMO, but it also likely determines the mechanism(s) by which the oxidative intermediates are generated (19).

In order to understand better the mode of action of LPMOs, with particular focus on the role of copper, we describe herein an integrated NMR/EPR spectroscopy approach designed to investigate the effect of copper- and substrate-binding to a chitin-active LPMO from *Bacillus licheniformis* (hereinafter called *B/LPMO10A*) and in particular the enzyme's ability to activate O<sub>2</sub> at the copper center. It is the first study of its kind on LPMOs, in which we have taken advantage of the <sup>15</sup>N-labelling that is required for the NMR study to simplify and constrain the analysis of the spin-Hamiltonian parameters obtainable from EPR spectroscopy. Using this approach, we have solved the NMR structure of both *apo-B/LPMO10A* and Cu(I)-*B/LPMO10A*, assessed dynamic features derived from relaxation data ( $T_1$ ,  $T_2$  and  $\{^1\text{H}\}$ -<sup>15</sup>N NOE) and evaluated the structural effects of Cu(I) and Cu(II) binding. In tandem, EPR spectroscopy performed on both natural isotopic abundance samples of <sup>63</sup>Cu(II)-*B/LPMO10A* and on uniformly isotope-labeled <sup>63</sup>Cu(II)-<sup>15</sup>N-*B/LPMO10A* has allowed the determination of hyperfine couplings. These couplings provide insights into the rearrangement of the copper coordination sphere upon substrate binding at a high level of detail, leading to the proposal of a potential substrate-O<sub>2</sub> coupling mechanism in chitin-active AA10 LPMOs. The work further sets the scene for future integrated NMR/EPR studies and in-depth EPR investigation of solid state samples, *i.e.* investigations performed with the protein bound to its natural solid state substrate—an important aspect of all LPMO-substrate studies.

## Results and discussion

### Functional characterization of *B/LPMO10A*

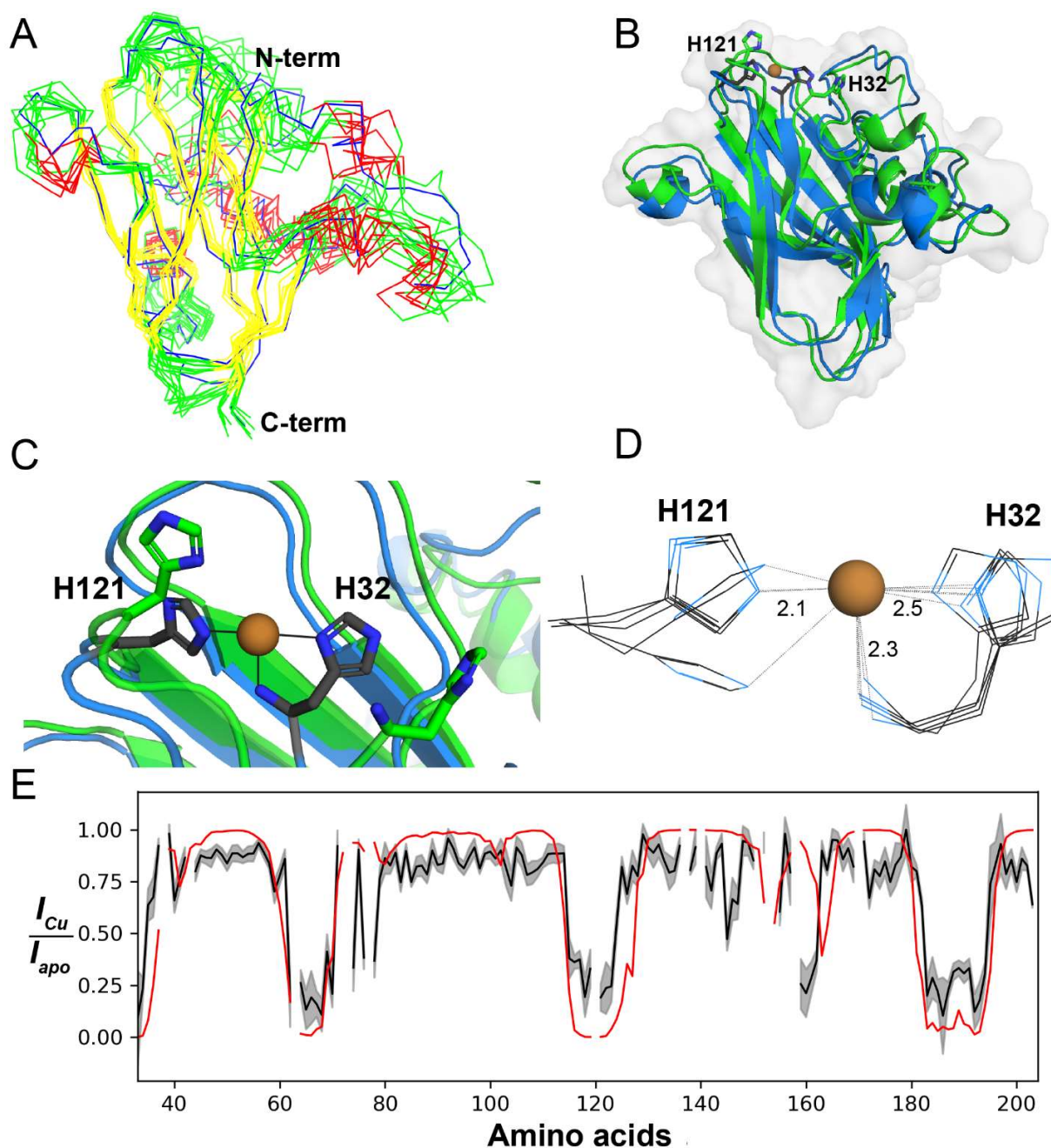
*B/LPMO10A* was recombinantly produced in *E. coli* and copper saturated using previously established methods (38, 39). Activity assays in reactions with 2 mM ascorbic acid as reductant showed that the enzyme is active on  $\beta$ -chitin and also has low activity on  $\alpha$ -chitin (Figure S1A). In reactions without added reductants product formation was not observed. The profile of oxidized oligomers (Figure S1A) showed a dominance of products with an even number of sugar units, which is a feature that is typical for LPMOs acting on crystalline chitin (9, 40, 41). As expected, binding to  $\beta$ -chitin was observed (Figure S1B).

### *apo*- and Cu(I) structures of *B/LPMO10A*

The solution structure of *apo-B/LPMO10A* (PDB ID: 5LW4) (Figure 1A) was elucidated using NMR spectroscopy. The structure was calculated in CYANA using 1623 NOE-derived distance

constraints, 264 TALOS-N determined torsion angle constraints, and one disulfide bridge (Cys45–Cys56) constraint (see Table S1). Like all other known LPMOs, *B/LPMO10A* has the typical fibronectin type III-like  $\beta$ -sandwich fold, in this case composed of eight  $\beta$ -strands that form a three-stranded and a five-stranded  $\beta$ -sheet, connected through loops of various lengths. The three-stranded sheet is composed solely of antiparallel strands, whereas the five-stranded sheet contains four antiparallel strands and one short parallel strand. The stretch of 66 amino acids that connects the first (Phe35–Lys38) and the second (His105–Met107)  $\beta$ -strands is composed of irregular loop regions, two  $\alpha$ -helices and one  $3_{10}$ -helix. A third short  $\alpha$ -helix occurs in the loop between the fifth (Phe146–Pro154) and the sixth (Gly176–Val185)  $\beta$ -strands.

NMR investigations of LPMOs are usually carried out using the *apo*-proteins, in order to avoid the detrimental signal reduction caused by the paramagnetic relaxation enhancement (PRE) effect brought about by the nature of the type II copper site (42). Here, we exploited the PRE effect to gain insights into the effects of Cu(II) on the structure of *apo-B/LPMO10A* (Figure S2). Cu(II) was added to a sample of *apo-B/LPMO10A* and the PRE effect was evaluated by comparing signal intensity reduction in  $^{15}\text{N}$ -HSQC spectra with PREs calculated using Cu(I)-*B/LPMO10A* structures, the  $g_{\text{iso}}$  values from Table 1 and relaxation parameters (Figure S3). Residues nearest the copper coordination site showed more than 80% reduction in signal intensity (Figure 1E). As expected, residues with the highest signal intensity reduction are located within a 12 Å radius from the Cu(II) coordination site (Figure S2). An exception is a short helix (Ala160-Arg162; Figure S3) that is further away than the expected 12 Å. This deviation could indicate structural differences beyond the copper-active site between copper-bound and *apo*-forms.



**Figure 1. Structures of *apo*- and Cu(I)-BILPMO10A.** (A) Ensemble of the 10 lowest energy conformers of *apo*-BILPMO10A (PDB:5LW4) in stereo representation. Helices are colored red, loops are colored green and strands are colored yellow; the lowest CYANA target energy conformer is colored blue. The overall backbone rmsd of the ensemble is 2.41 Å, while the rmsd of the regions containing  $\alpha$ -helices (residues 41-47, 57-61, 82-84, 89-94, 159-162) and  $\beta$ -sheets (residues 33-36, 37-40, 103-107, 110-117, 125-132, 150-154, 164-168, 175-184, 185-190 and 191-201) is 1.48 Å. (B) Overlay of *apo*-BILPMO10A (green) and Cu(I)-BILPMO10A (PDB:6TWE blue). The copper atom is shown as an orange sphere and the sidechains

of His32 and His121 are shown as sticks. The backbone ( $C^\alpha$ , N,  $C'$ ) rmsd between the *apo* ensemble and the Cu(I) ensemble is 0.9 Å. (C) Zoomed-in view of the overlay in (B) showing details of the copper site. (D) Ensemble of five lowest energy conformers of Cu(I)-*B/LPMO10A*, showing the copper site. Average distances from each N-atom to the Cu atom are indicated. (E) PRE effects upon adding Cu(II) to *apo-B/LPMO10A*. The black line shows the normalized  $H^N$ , N signal intensity upon addition of Cu(II) to  $^{13}C$ - and  $^{15}N$ -labeled *apo-B/LPMO10A* in a 1:2 ratio, relative to the intensity for the *apo*-enzyme, with errors shown in grey. The red line shows PREs calculated using the Cu(I)-*B/LPMO10A* ensemble. Gaps in the data represent missing assignments for amino acid residues (e.g. Pro).

Preparation of a Cu(I)-bound sample of  $^{13}C$  and  $^{15}N$  labeled *B/LPMO10A* enabled structure determination of the reduced Cu(I)-LPMO in solution (PDB ID: 6TWE and Figure 1). The structure was elucidated by using 1209 distance restraints derived from NOESY spectra of a Cu(I)-bound *B/LPMO10A* sample together with force-field parameters for the copper site of another AA10, *SmLPMO10A* (32). Overall, the Cu(I)-structure is similar to the *apo*-structure (Figure 1B), but there are clear differences in the internuclear  $^1H$ - $^1H$  distances at the copper site (Figure 1C). The overall backbone ( $C^\alpha$ , N,  $C'$ ) rmsd between the *apo*- and Cu(I)-structures is 0.9 Å. Analysis of NOESY spectra revealed that the presence of Cu(I) slightly affects internuclear distances between the  $\beta$ -strands, suggesting minor sliding movements in the  $\beta$ -strands upon copper binding. While Figure 1C shows that copper binding has a strong effect on the conformation of the copper-coordinating His side-chains, limitations inherent to NMR data preclude determination of exact atomic coordinates for the copper structure. These limitations are evident in Figure 1D, which shows that the histidine side chains do not coordinate to the Cu with the expected T-shape geometry, which is chemically unreasonable. Figures 1B and 1C show that copper binding has a major effect on the conformation of the copper site. This conformational change is a result of the combined effects of introducing force-field parameters for the copper site and of structural changes encoded in NOE-derived distance restraints. Prior to this study, Cu(I)-structures of LPMOs had only been obtained by photoreduction of Cu(II) in the X ray beam during data acquisition for crystallography or X-ray absorption studies (42).

### Heteronuclear relaxation

To gain insight into the motion of *B/LPMO10A* in solution,  $\{^1H\}$ - $^{15}N$  NOE,  $^{15}N$ - $T_1$ , and  $^{15}N$ - $T_2$  were measured (Figure S3), and the rotational correlation time,  $\tau_c$ , which relates to molecular

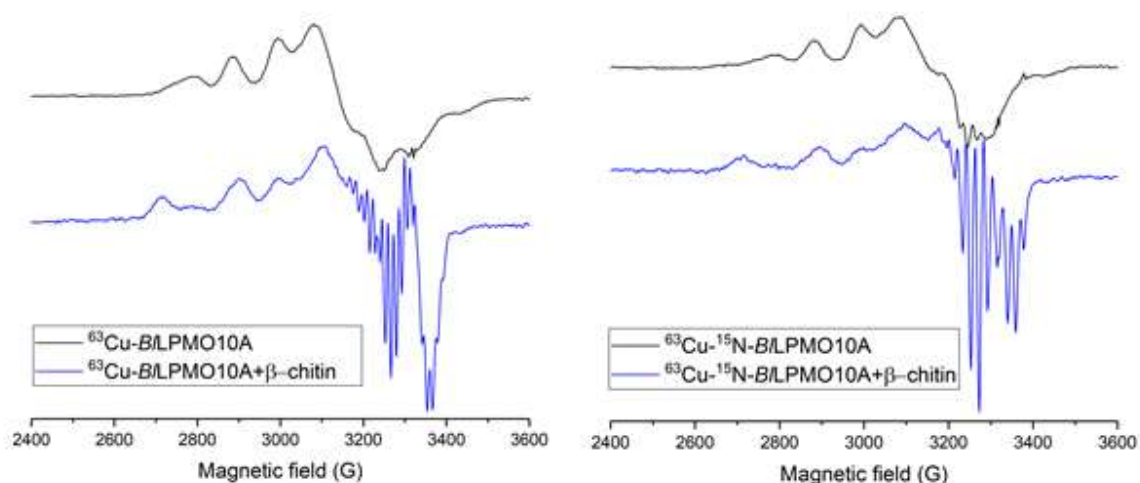
tumbling, was determined from the average  $T_1/T_2$ . The overall horizontal trends in Figure S3 suggest that the core of *B/LPMO10A* is rigid, while loops, particularly in the first half of the protein, have reduced  $\{^1\text{H}\}$ - $^{15}\text{N}$  NOE and increased  $T_2$  values that indicate some conformational flexibility. The rotational correlation time ( $\tau_c = 10.2 \pm 0.9$  ns) was found to be similar to what would be expected for a globular protein of similar molecular weight ( $\tau_c = 9.75 \pm 0.46$  ns for a 20 kDa globular protein (43)). This finding indicates that *B/LPMO10A* is a well-packed protein, as observed previously for other LPMOs.

### **EPR spectroscopy**

The availability of  $^{15}\text{N}$ -labeled enzyme offered the possibility of simultaneously determining the spin-Hamiltonian parameters of the Cu(II) unpaired electron for both  $^{14}\text{N}$  and  $^{15}\text{N}$  species, thereby providing a means of accurately determining their values. To this end, CW EPR spectra were collected at both X-band and Q-band frequencies with pure  $^{63}\text{Cu}$  isotopes of LPMOs, from which a simultaneous fit of the spectra at both frequencies provided a doubly-constrained and therefore reliable set of Cu(II)  $g$  values and hyperfine coupling constants ( $A_{\text{Cu}}$ ), along with nitrogen superhyperfine coupling constants. The data from these fits are presented in Table 1 and shown in Figures 2 and S4.

In the absence of chitin, simultaneous fits of X- and Q-band spectra afforded a consistent set of copper spin-Hamiltonian parameters, characterized by rhombic principal  $g$  matrix values, a reduced  $|A_3|$  value and large  $|A_{1,2}|$  values (with respect to typical  $|A|$  values for axial Cu(II) systems), where the overall spectral envelope indicates a SOMO with mostly  $d(x^2-y^2)$  character (Table 1). These values are similar to those obtained for other chitin-active AA10 LPMOs, but differ in that the  $|A_1|$  value, which is more accurately determined in the present study through the use of two frequencies, is larger than previously reported (32, 35, 38), as is the corresponding  $g_1$  value.





**Figure 2. CW-EPR spectra of *B/LPMO10A*.** Spectra for  $^{63}\text{Cu}$ -*B/LPMO10A* (left) and  $^{63}\text{Cu}$ - $^{15}\text{N}$ -*B/LPMO10A* (right) before and after addition of squid pen  $\beta$ -chitin (black and blue lines, respectively). The spectra were recorded with 0.29 mM of  $^{63}\text{Cu}$ -*B/LPMO10A* and 0.17 mM  $^{63}\text{Cu}$ - $^{15}\text{N}$ -*B/LPMO10A*, both in 20 mM MES buffer pH 5.5 with 10% glycerol.

**Table 1. Spin-Hamiltonian parameters for  $^{63}\text{Cu(II)}$ -*B/LPMO10A* and  $^{63}\text{Cu(II)}$ - $^{15}\text{N}$ -*B/LPMO10A* at pH 5.5 with and without squid pen  $\beta$ -chitin.** For coupled nitrogen nuclei, only the principal coupling value could be determined from the simulations of the superhyperfine (SHF) which we presume is the coupling along the Cu-N bond; the three values in each spectrum refer to three different N nuclei, with the smallest value in each set being assigned to the  $\text{NH}_2$ .

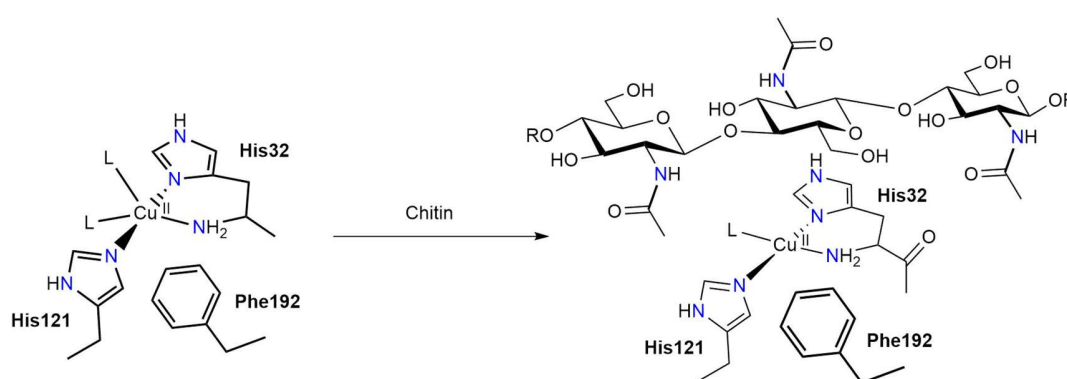
		$^{63}\text{Cu}$ - <i>B/LPMO10A</i>	$^{63}\text{Cu}$ - $^{15}\text{N}$ - <i>B/LPMO10A</i>	$^{63}\text{Cu}$ - <i>B/LPMO10A</i> + $\beta$ -chitin	$^{63}\text{Cu}$ - $^{15}\text{N}$ - <i>B/LPMO10A</i> + $\beta$ -chitin	
		X-band	X-band	X-band	X-band	Q-band
<b>g values</b>	$g_1$	2.027	2.029	2.032	2.042	2.046
	$g_2$	2.095	2.081	2.112	2.053	2.057
	$g_3$	2.261	2.261	2.260	2.205	2.208
	$g_{\text{iso}}$	2.128	2.124	2.135	2.101	2.104
<b><math>A_{\text{Cu}}</math> (/MHz)</b>	$ A_1 $	255	255	255	80	80
	$ A_2 $	110	115	115	85	90
	$ A_3 $	336	336	340	620	610
<b>calculated*</b>	$A_{\text{iso}}$	10	11	10	-208 or -262	-207 or -260
	<b>SHF <math>A_{\text{N}}</math> principal</b>	43, 43, 28 $\pm 5$	60, 60, 40 $\pm 5$	60, 60, 40	40, 40, 32 $\pm 2$	56, 56, 45 $\pm 2$

values** (/MHz)						
$g_{Cu}$ strains	0, 0.02, 0.007	0, 0.025, 0.005	0, 0.045, 0.007	0.004, 0, 0.007	0.005, 0, 0.009	0.005, 0, 0.003
$A_{Cu}$ strains (/MHz)	130, 55, 160	160, 60, 160	20, 10, 120	20, 20, 20	10, 10, 10	90, 40, 240
Linewidths	0.6, 0.6	0.6, 0.6	4.5, 4.5	0.4, 0.4	0.5, 0.6	1.3, 1.3
Frequency (/GHz)	9.3046	9.2973	35.00	9.2988	9.2884	35.05

\* signs of  $A_1$  and  $A_2$  calculated from DFT (see main text).

\*\* error estimated from quality of simulated fits

The rhombic spin-Hamiltonian parameters, particularly the reduced  $g_3$  along with increased  $g_2$  (as compared to  $g$  values for typical axial Cu(II) systems), derive from d-orbital mixing which occurs in copper complexes that possess a distorted square-pyramidal coordination geometry (44). Indeed, such a coordination geometry in Cu(II)-*B/LPMO10A* would be in accord with those previously observed in the crystal structures of AA10 LPMOs in the Cu(II) oxidation state (41, 45-47). In these structures the base of the pyramid is defined by the three nitrogen atoms of the histidine brace and a water molecule. A further water molecule in the nominal axial position, albeit slightly off axis with respect to the ideal geometry, completes the coordination sphere (Figure 3 left).



**Figure 3.** Schematic representation of the change in the coordination sphere of the copper ion upon binding of  $\beta$ -chitin ( $L = H_2O$  or  $OH^-$ ).

Addition of chitin flakes to the samples of Cu(II)-*B/LPMO10A* led to significant changes in both X-band and Q-band EPR spectra (Figure 2). Excellent simultaneous fits of the different frequency

spectra could be obtained, with a notably high correlation in the areas in which superhyperfine coupling is evident (Figure S4). In the presence of chitin, the spectra are well defined with axial copper spin-Hamiltonian parameters ( $g_1 \sim g_2 < g_3$ ) and a large  $|A_3|$  value. The latter value leads to the appearance of an ostensible “overshoot” feature (also known as “extra absorption peak” (48)) at high field in the X-band spectrum, which has been suggested in other studies on LPMOs to arise from a low  $g_1$  value (2.018) (32). However, Q-band spectra of Cu(II)-*B/LPMO10A* demonstrate that the high field feature in the X-band spectrum actually arises from the large  $|A_3|$  value and that, when simulated with two frequencies, the perpendicular hyperfine constants are those typical for an axial Cu(II) complex in which there is little mixing into the  $d(x^2-y^2)$  SOMO from other metal-based orbitals and the  $g$  values are typical for an axial Cu(II) complex. Thus, in the presence of chitin, the copper coordination sphere is one which has near axial coordination symmetry, similar to that seen in most AA9 LPMO structures (Figure 3 right) where the equatorial plane of the copper is defined by the three nitrogen atoms of the histidine brace and the coordinating atom of an exogenous molecule (*e.g.* water, hydroxide), all held within a near planar arrangement around the copper.

#### *Substrate-induced effects at the active site*

While the NMR solution structure of Cu(I)-*B/LPMO10A* offers qualitative insights into structural rearrangements induced by copper-binding, it does not provide a high-resolution structure of the active site (Figure 1D). It is thus not possible to corroborate the EPR parameters with DFT calculations based on the Cu(I)-*B/LPMO10A* model of the active site. Therefore, we used ligand field theory enhanced by DFT (acronym DELFT for DFT-enhanced ligand field theory) to analyze the nature of the SOMO and its magnetic interactions (see Supplementary Discussion). In the DELFT method, DFT calculations over a range of functionals (Table S3) are performed on the active site of a closely-related AA10 LPMO (*BaAA10* from *Bacillus amyloliquefaciens*, 58% sequence identity to *B/LPMO10A*) where the structure of the active site in its Cu(II) form is known from X-ray crystallography (41), and where *BaAA10* shows virtually identical CW-EPR spectra to those of *B/LPMO10A* (35, 49), demonstrating that the two enzymes have a comparable arrangement at the active site and a similar response to the addition of chitin. On this basis, the coordinates for *BaAA10* were used as a basis for DFT calculations (Figures S7-S9). The objective

of DELFT approach is not to obtain accurate spin-Hamiltonian parameters for the Cu and coordinating nitrogen atoms, not least because the calculation of such parameters with DFT is fraught with difficulty (50), but rather to determine the signs of the Cu hyperfine coupling constants which cannot be ascertained experimentally from CW EPR spectroscopy. Accordingly, it was found that, in the presence of substrate,  $A_2$  is negative and the sign of  $A_1$  is unclear, and that both  $A_1$  and  $A_2$  are positive in the absence of substrate (Tables S4-S7). Accordingly, it is possible to calculate experimental  $A_{iso}$  values of ca.  $-205$  or  $-265$  MHz (depending on the sign of  $A_1$ ) in the presence of substrate and ca.  $10$  MHz in its absence (Table 1), representing a significant shift in  $A_{iso}$  induced by the addition of substrate (Supporting Information).

In the context of ligand field theory  $A_{iso}$  is a useful measure since it is determined, to second order within a fixed  $g$  matrix, by three main factors: i) spin-orbit contributions, ii) the degree of Fermi contact of the unpaired electron with the copper nucleus, and iii) the spin density at the copper, which in turn can be related to the degree of metal-ligand covalency in the ground state. Which of these is the origin of the shift in  $A_{iso}$  value upon substrate binding in *B/LPMO10A* can then be evaluated from the experimentally well-defined  $A_3$  hyperfine coupling value of the copper, using the following equation, which applies to the hyperfine coupling constants of a Cu(II) ion with  $d(x^2-y^2)$  SOMO in a distorted square pyramidal geometry.

$$A_3 = -P_d \left[ K + \frac{4\alpha_{GS}^2(a^2 - b^2)}{7} - \frac{(3a - \sqrt{3}b)\Delta g_2}{14(a + \sqrt{3}b)} - \frac{(3a + \sqrt{3}b)\Delta g_1}{14(a - \sqrt{3}b)} - \Delta g_3 \right]$$

In this equation,  $P_d = 1180 \text{ cm}^{-1}$ ,  $a$  is the orbital coefficient of the  $d(x^2-y^2)$  orbital,  $b$  is the orbital coefficient of the  $d(z^2)$  orbital,  $\alpha_{GS}^2$  is spin density on Cu,  $\Delta g_n = g_n - 2.0023$  and the Fermi contact  $-P_d K$  which can be calculated from:

$$A_{iso} = P_d \left[ -K + \frac{1}{3}(\Delta g_x + \Delta g_y + \Delta g_z) \right]$$

Apart from the spin density,  $\alpha_{GS}^2$ , the only unknown in the first equation is the degree of  $d(z^2)$  mixing into the ground state, denoted by the orbital coefficient  $b$ . (This value can be estimated from the difference in  $g_1$  and  $g_2$  values (51), giving a value of  $b^2$  to be  $\sim 2\%$ ). The values for the

individual contributions to the  $A_{\text{iso}}$  can then be calculated (Table 2). Also shown in Table 2 are the equivalent theoretical values calculated using DFT of the active site of *BaAA10* in two forms, one with two exogenous water molecules coordinating to the copper (analogous to the active site structure in the absence of substrate) and the other with a single water as the exogenous ligand, mimicking the substrate-bound state (see Table S3, S4 and below for further discussion).

**Table 2.** Contributions to hyperfine coupling /MHz. Para = parallel direction, perp = perpendicular direction. + = presence of substrate, - = absence of substrate \* calculated with  $A_{\text{iso}} = -208$  MHz, \*\* calculated with  $A_{\text{iso}} = -262$  MHz

Method + chitin, - no chitin	Fermi contact /MHz	Dipolar para /MHz	Dipolar perp /MHz	Orbital para /MHz	Orbital perp /MHz	Spin density $\alpha$
DELFT -	-139	-533	281	332	70	0.82
DFT -	-178	-528	264	295	94	
DELFT +*	-324	-558	279	262	53	0.83
DELFT+**	-380	-504	252	264	50	0.75
DFT +	-348	-545	272	291	80	

This analysis shows that the DELFT and DFT approaches yield broadly consistent values and the same trends in how the contributors to hyperfine coupling constants change upon substrate addition, with the largest change appearing in the value of the Fermi contact. The DELFT approach also reveals that the spin density,  $\alpha_{\text{GS}}^2$ , in the presence of substrate takes one of two values (0.83 or 0.75) depending on the sign of  $A_1$  used in the calculation of  $A_{\text{iso}}$ . It is likely that the value of 0.83 is the more reliable one, given that there are no significant changes in the identity of coordinating atoms to the copper upon the addition of substrate (i.e. we do not expect a large increase in the covalency of the copper-ligand bonds as would be required by a spin density value of 0.75). Thus, it appears as if the addition of substrate affords either no change or a slight decrease in the overall metal-ligand covalency.

Most contributors to the value of the hyperfine coupling are not altered much by the addition of substrate, However, the change in the value of Fermi contact parameter,  $-P_{\text{Cu}\kappa}$  (see Supplementary Discussion), is significant, going from  $-139$  MHz to  $-324$  MHz. For comparison, the latter value

is similar to that calculated for the square-planar complex  $[\text{Cu}(\text{NH}_3)_4]^{2+}$  ( $-362$  MHz) using referenced configuration interaction calculations (52).

Any change in value of the Fermi contact upon a chemical perturbation to the copper coordination sphere arises from two principal sources: differences in the 4s mixing with the SOMO (a positive contributor to the value of the Fermi contact) and/or changes to spin polarization of the copper core (a negative contributor) or valence (a positive contributor) orbitals. In this case, it is unlikely the change in the value of the Fermi contact upon substrate binding is due only to a reduction in 4s orbital contribution to the SOMO, since the observed shift of  $+185$  MHz in Fermi contact upon substrate binding would require a 4s orbital content of  $\sim 4\%$  in the absence of substrate.(53, 54) This value is higher than expected when compared to the  $d(z^2)$  content of  $\sim 2\%$ . Thus, the changes in the value of the Fermi contact upon substrate binding are also caused by a significant increase in core orbital polarization and/or reduction in polarization of valence orbitals. Previous studies on copper(II) complex hyperfine values have emphasized the importance of valence shell polarization, especially when there is a change in coordination number at the copper (with polarization of the core orbitals being relatively insensitive to coordination changes and proportional to the overall spin density at the Cu) (55). As such, the large change in Fermi contact upon substrate binding likely arises from the change in coordination number of five to four at the copper with an attendant reduction in the polarization of the valence electrons, together with a small contribution due to the reduction of 4s orbital mixing with the SOMO.

#### *Substrate-induced changes in metal-ligand covalency*

Further indications about any changes in metal-ligand bonding upon substrate binding can be gleaned from the values of the largest nitrogen superhyperfine coupling values. Table 1 shows that upon substrate addition, the coupling values assigned to the two  $\text{Cu}-^{15}\text{N}(\text{His})$  interactions reduce from  $\sim 60(\pm 5)$  MHz to  $\sim 55(\pm 2)$  MHz, and that there is an increase in the coupling assigned to the  $\text{Cu}-^{15}\text{NH}_2$  interaction from  $\sim 40(\pm 5)$  to  $\sim 45(\pm 2)$  MHz. Analogous shifts, corrected for the difference in gyromagnetic ratio, are seen for the  $^{14}\text{N}$  isotopologue. Despite the excellence of the simulated fits however, the estimated errors in the values preclude a definitive conclusion about differences in their values before and after substrate addition. The major source of error lies in the CW-EPR spectrum of *B/LPMO10A* in the absence of substrate. Thus, Davies  $^{14}\text{N}$  ENDOR data

were collected on the enzyme before substrate addition, which were orientation-selected to the Cu perpendicular direction (Figure S5). (Unfortunately, we were unable to collect satisfactory spectra on the sample after substrate addition, see below.) The spin-Hamiltonian parameters obtained from the  $^{14}\text{N}$  ENDOR experiments (Table S2) match well the values derived from CW-EPR experiments, giving  $A_{\parallel,\text{N}}(\text{Cu},\perp) = 23 (\pm 1)$  MHz (*cf* 28(5) MHz from CW-EPR), assigned to the N-terminus nitrogen, and  $A_{\parallel,\text{N}}(\text{Cu},\perp) = 40 (\pm 2)$  MHz (*cf* 43(5) MHz from CW-EPR) for the two His nitrogen atoms, lending confidence in the values obtained from CW EPR. (It is notable that the experimental Cu- $^{14}\text{NH}_2$  coupling value is much less than that obtained from previous DFT calculations on LPMOs, ca 50 MHz, likely reflecting the well-known issues with DFT in calculating accurately such coupling values.) Given the lack of a satisfactory  $^{14}\text{N}$  ENDOR spectrum of the enzyme in the presence of substrate, orientation-selective  $^{15}\text{N}$ -HYSCORE spectra before and after chitin addition were also collected. The use of  $^{15}\text{N}$  in this regard significantly simplifies the spectra, allowing extraction of the key coupling values with high accuracy (see next section for further discussion of these data). These data show that, upon substrate addition, the dipolar coupling of the two remote nitrogen atoms of each of the two imidazole rings of the histidine brace *increases* from  $T = 0.30$  and  $0.34$  MHz to  $0.35$  and  $0.40$  MHz, respectively (Table 3 and Figure 4), while the  $a_{\text{iso}}$  couplings decrease from  $2.00$  to  $1.90$  and  $2.40$  to  $2.20$  MHz. The calculated  $A+2T$  values, which represent the axial coupling of the N atoms to the Cu, can thus be estimated to be  $2.6$  and  $3.1$  MHz in the absence of substrate and  $2.6$  and  $3.0$  MHz in the presence of substrate. The small decrease in  $a_{\text{iso}}$  upon substrate binding ( $\sim 10\%$ ; Table 3) is in accord with the decrease in the coupling constants of the coordinating nitrogen atoms of the histidine rings observed from simulations of the CW EPR data (Table 1). The small increase in dipolar coupling is in accord with an increase in spin density at the Cu upon substrate binding.

The overall picture which emerges is that substrate binding drives a structural rearrangement at the copper ion characterized by a greater Cu-NH<sub>2</sub> covalency and reduced Cu-His covalency, accompanied by a small but significant increase in spin density at the Cu. Such changes would be consistent with a structural rearrangement where the N<sub>3</sub> T-shaped coordination geometry provided by the histidine brace is shortened in the Cu-NH<sub>2</sub> direction and elongated in the Cu-N<sub>His</sub> directions upon substrate binding.

*Substrate-induced changes to hydrogen-bonding to the histidine brace*

It is known from crystal structure and modeling studies of LPMOs with oligosaccharides that the substrate does not form a direct coordination bond to the copper (22, 31, 32). Thus, the substrate-induced changes in the copper coordination sphere described above are due to changes in the primary coordination sphere brought about by the substrate or through interactions of the substrate with the copper's outer coordination sphere(s). In terms of these outer coordination sphere interactions, N-HYSCORE spectroscopy is a powerful tool in that it can provide hydrogen bonding information on the remote nitrogen atoms on the imidazole rings of the histidines. In particular, the values of the quadrupole tensor which are determined from  $^{14}\text{N}$ -HYSCORE spectra include an estimate of the electric field gradient that exists at the nitrogen atom, which—in turn—is directly related to the strength of any hydrogen bond formed at the N-H group.

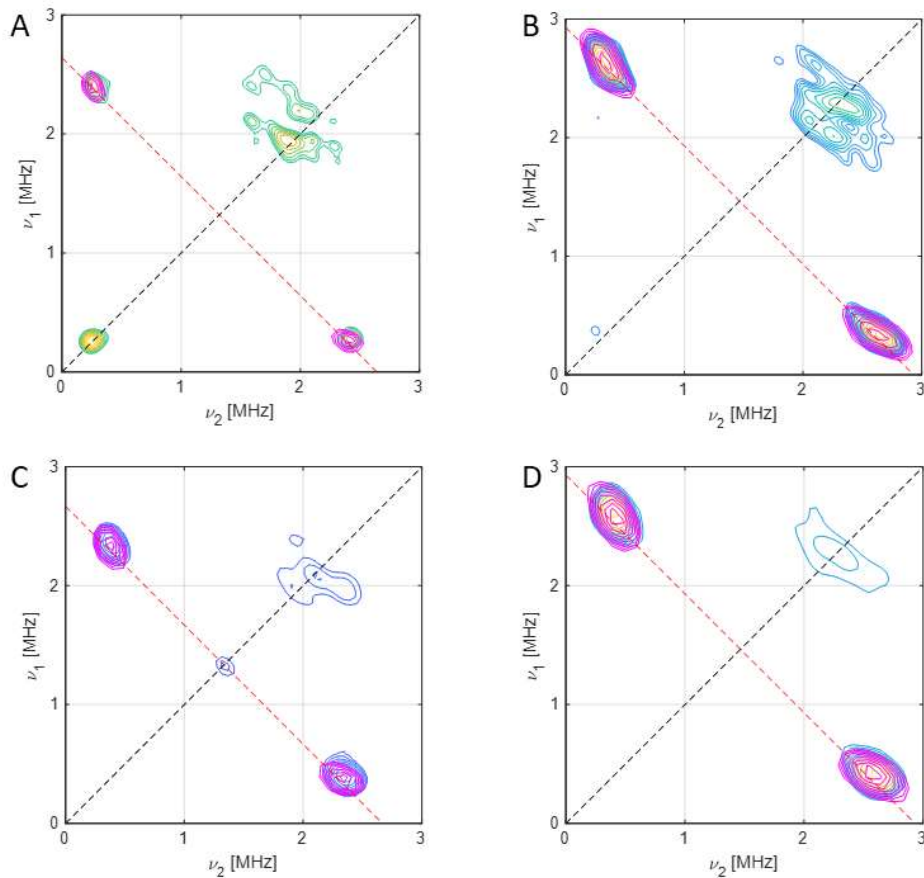
**Table 3.**  $^{15}\text{N}$  HYSCORE simulation parameters for  $^{63}\text{Cu}(\text{II})$ - $^{15}\text{N}$ -*B/LPMO10A* with and without squid pen  $\beta$ -chitin. The perpendicular and parallel symbols are used to define the set of simulation parameters used for the spectra collected at that field position ( $\parallel$  for 3060 or 3090 G and  $\perp$  for 3995 G). The numbers in brackets represent the error on the measurement estimated from the quality of simulated fits. The Euler angles define the  $zy'z''$  rotations with respect to the  $g$  matrix directions.

	$^{63}\text{Cu}$ - $^{15}\text{N}$ - <i>B/LPMO10A</i>			$^{63}\text{Cu}$ - $^{15}\text{N}$ - <i>B/LPMO10A</i> + $\beta$ -chitin		
	$a_{iso}$	T	A Frame Euler angles	$a_{iso}$	T	A Frame Euler angles
N(A) $\perp$	2.40 (0.02)	0.34 (0.03)	[10 100 60]	2.20 (0.05)	0.40 (0.02)	[10 110 60]
N(B) $\perp$	2.00 (0.05)	0.30 (0.05)	[0 94 20]	1.90 (0.05)	0.35 (0.05)	[120 70 0]
N(A) $\parallel$	2.40 (0.02)	0.34 (0.05)	[10 100 60]	2.25 (0.05)	0.40 (0.05)	[10 110 60]
N(B) $\parallel$	2.00 (0.05)	0.30 (0.05)	[0 94 20]	1.85 (0.05)	0.35 (0.05)	[120 60 0]

Determining accurate values for the electric field gradient at the ring  $^{14}\text{N}$ -nitrogen atoms requires the accurate evaluation of both the quadrupole and hyperfine tensor values. This is normally difficult, however as described above, simulation of the  $^{15}\text{N}$  HYSCORE spectra in the absence of substrate gave  $a_{iso}$  values of  $2.40 \pm 0.02$  and  $2.00 \pm 0.05$  MHz and dipolar coupling values (T) of



0.34±0.03 and 0.30±0.05 MHz. Upon substrate addition, the  $a_{iso}$  values reduce by 5-10% and the T values increase by ~15% (Table 3, Figure 4). These accurate values for coupling could then be used in the  $^{14}\text{N}$  HYSCORE simulations (Figure S6), from which it was then possible to determine the quadrupole tensor elements ( $e^2qQ/h = K$  and  $\eta$ ), where the asymmetry parameter  $\eta$  is related to the strength of the hydrogen bond interaction (56). In particular, a value of  $\eta$  close to 1 is associated with the N-H participating in a strong H-bond with an outside H-bond acceptor, while values between 0.45 and 0.75 are associated with weak H-bonds. The nuclear quadrupole parameters for  $^{63}\text{Cu-B/LPMO10A}$  before and after addition of  $\beta$ -chitin (Table 4) show that the H-bonding environment of one of the nitrogen atoms is not perturbed by the binding of chitin. Conversely, the  $\eta$  value of the other nitrogen atom increases from 0.7 to 0.9, showing that the substrate drives a significant change in the H-bond network around this N-H group, possibly through a direct H-bonding interaction with the substrate. On the basis of the DFT calculations (Table S8), it is possible to tentatively assign the former (N(B) in Tables 3 and 4) as the remote nitrogen on His32 and the latter (N(A) in Tables 3 and 4) as the remote nitrogen on His121, in accordance with modelling studies carried out on another chitin-active LPMO (32). This change in hydrogen-bonding pattern around the histidine is likely an important contributor to the observed changes in Cu-N(His) covalency brought about by substrate addition.



**Figure 4.**  $^{15}\text{N}$  HYSORE spectra and simulations of  $^{63}\text{Cu}$ - $^{15}\text{N}$ -*B/LPMO10A* (**A, B**) and  $^{63}\text{Cu}$ - $^{15}\text{N}$ -*B/LPMO10A* with squid pen  $\beta$ -chitin (**C, D**). Numerical simulations (in pink) were obtained with the values reported in Table 3. **A**,  $^{63}\text{Cu}$ - $^{15}\text{N}$ -*B/LPMO10A* near  $g_{\perp}$  with  $\tau = 136$  ns at 3395 G; **B**,  $^{63}\text{Cu}$ - $^{15}\text{N}$ -*B/LPMO10A* near  $g_{\parallel}$  with  $\tau = 136$  ns at 3060 G; **C**,  $^{63}\text{Cu}$ - $^{15}\text{N}$ -*B/LPMO10A* with squid pen  $\beta$ -chitin near  $g_{\perp}$  with  $\tau = 200$  ns at 3390 G; **D**,  $^{63}\text{Cu}$ - $^{15}\text{N}$ -*B/LPMO10A* near  $g_{\parallel}$  with  $\tau = 200$  ns at 3090 G.

**Table 4.**  $^{14}\text{N}$  HYSORE simulation parameters for  $^{63}\text{Cu}(\text{II})$ -*B/LPMO10A* with and without squid pen  $\beta$ -chitin. The spectra were collected near  $g_{\perp}$  (3385 G and 3390 G for the sample without and with  $\beta$ -chitin, respectively, with  $\tau = 200$  ns or 136 ns). The numbers in brackets represent the error on the measurement estimated from the quality of simulated fits. The Euler angles define the  $zy'z''$  rotations with respect to the  $g$  matrix. Spectra are shown in Figure S6.

$^{63}\text{Cu}$ - <i>B/LPMO10A</i>						
$a_{iso}$	T	A Frame	K	$\eta$	Q Frame	

	Euler angles			Euler angles		
N(A)	1.6	0.3	[10 100 60]	1.75 (0.05)	0.7 (0.05)	[60 10 95]
N(B)	1.3	0.25	[0 94 20]	1.40 (0.05)	0.85 (0.05)	[20 0 -95]
<b><sup>63</sup>Cu-B/LPMO10A + <math>\beta</math>-chitin</b>						
N(A)	1.55	0.35	[10 100 60]	1.35 (0.05)	0.9 (0.05)	[60 30 95]
N(B)	1.25	0.30	[120 70 0]	1.40 (0.05)	0.8 (0.05)	[10 -10 -80]

### *Substrate-induced changes in d-orbital energies*

To probe further the changes at the copper brought about by substrate binding, a calculation was made to determine changes in the value of  $g_{\text{iso}}$ . Unlike  $A_{\text{iso}}$ , this value reports on the nature of both the ground state SOMO and its associated excited state SOMOs (generated by Cu based d-d excitations), including the energy separations of the d orbitals and the notional metal-ligand  $\pi$ -covalency of the excited states. Accordingly, an analysis of  $g_{\text{iso}}$  is complicated and care must be exercised in its interpretation. Notwithstanding this caveat however, the  $g_{\text{iso}}$  value is seen to reduce ( $\Delta \sim 0.03$ ) upon substrate addition (Table 1). This reduction is *counter* to that expected from a decrease in either the covalency and/or  $d(z^2)$  mixing of the ground state SOMO, but it is commensurate with an increase in the energy separation of the SOMO from excited d orbital states (which is also reflected in a decrease in the value of  $g_3$  by  $\sim 0.05$  upon substrate addition). Without access to electronic absorption data, which is precluded by the solid nature of the chitin substrate, it is not possible to be more definitive about the d-d transition energies. However, in an indication of the adoption of a more axial-like symmetry upon substrate binding, the difference in values of  $g_1$  and  $g_2$ ,  $\Delta g$ , reduces from  $\sim 0.06$  to  $\sim 0.01$ . A large  $\Delta g$  is associated with  $d(z^2)$  mixing, which would arise from the distorted coordination geometry before substrate addition. The greater degree of  $d(z^2)$  character in the SOMO before substrate addition is reproduced by the DFT calculations (2.7% compared to 0.5% following substrate addition, see below and Table S9). Therefore, the changes in  $g$  values are ones that would be expected when the copper coordination sphere rearranges from a distorted square pyramid (5 coordinate) to one which is near axial square planar (4 coordinate), a geometry change which would be accompanied by an increase in the relative energy (with respect to the other d orbitals) and orbital character of the  $d(x^2-y^2)$  SOMO.

Overall, the electronic changes to the copper SOMO in chitin-active AA10 LPMOs which occur upon addition of substrate are: a reduction in the asymmetry of the equatorial plane of the copper

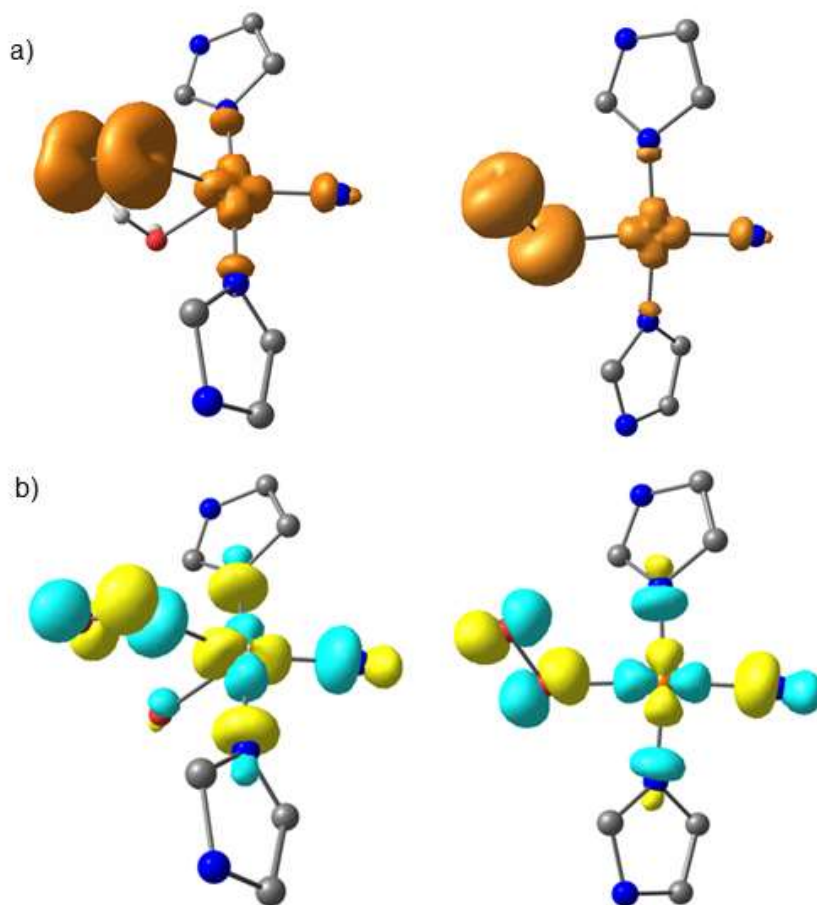
coordination sphere (likely associated with a smaller difference in the Cu–N(His) to Cu–N(amino) distances upon chitin addition), an increase in the relative energy of the SOMO, a reduction in ground state metal-ligand covalency to the histidine ligands, and an increase in the  $d(x^2-y^2)$  character of the SOMO through a process of reduced interactions with metal and ligand-based orbitals. A simpler view of these changes would be that, before substrate addition, the redox-active orbital on the copper is somewhat delocalized through mixing with other orbitals on both the metal and ligands. However, the addition of substrate allows for a spatially contracted and energetically well-defined  $d(x^2-y^2)$  orbital to ‘surface’ at the copper, one that is capable of forming a strong covalent interaction with an exogenous ligand in the equatorial plane of the copper coordination sphere. The mechanistic consequences are discussed further below, but such a switch in the character and energy of the frontier redox-active orbital clearly provides a basis for a potential coupling mechanism between the substrate and any exogenous ligands on the copper (*e.g.*  $O_2^-$ ).

### **DFT calculations**

To corroborate this analysis, DFT calculations were performed on the closely-related *BaAA10* LPMO, as described above for two different models of the active site (see Supplementary Information for details). The first, which emulates the enzyme in the absence of substrate, contained a superoxide and a water molecule in the copper’s coordination sphere in addition to the coordinating atoms of the histidine brace, and the second, emulating the enzyme in the presence of substrate, contained only superoxide as the exogenous ligand. From each calculation, spin population analysis (Figure 5, Table S9) reveal that the switch from five coordinate to four coordinate Cu(II) is accompanied by a large decrease in spin population on Cu(II) from 54% to 41%, which transfers almost completely to the distal oxygen atom of the superoxide (without a significant change in the O-O bond length,  $\Delta r(O-O) = 0.02 \text{ \AA}$ ). These changes in spin population therefore reflect the high degree of covalency between the copper and the superoxide ligand in the four coordinate state, *i.e.* when substrate is bound (Figure 5), corroborating the foregoing DELFT analysis of the changes that occur at the active site on the addition of substrate to *B/LPMO10A*.

It was further possible from these calculations to estimate the relative change in the strength of the Cu(II)-superoxide bonds upon substrate addition. This estimation is made by performing single point optimisations of the active sites in the presence and absence of superoxide (Figure S11), and

then calculating the difference in electronic energies between the two for both five-coordinate Cu ( $\Delta E_1$ ) and four coordinate Cu ( $\Delta E_2$ ) cases.  $\Delta\Delta E (= \Delta E_1 - E_2)$  can then be calculated in which all intrinsic errors in the calculated point electronic energies, save for a small basis set superposition error, are expected to cancel leaving only the difference in copper-superoxide bond strength as the principal contributor to the value of  $\Delta\Delta E$ . This value shows that the Cu(II)-superoxide bond is 34  $\text{kJ mol}^{-1}$  greater in the four coordinate substrate-bound case. Translated into equilibrium constant terms at 298 K, where it is assumed that  $\Delta\Delta G \sim \Delta\Delta E$ , this difference means that—in the presence of substrate—the copper-superoxide complex is  $\sim 10^6$  more stable to dissociative elimination than in the absence of substrate.



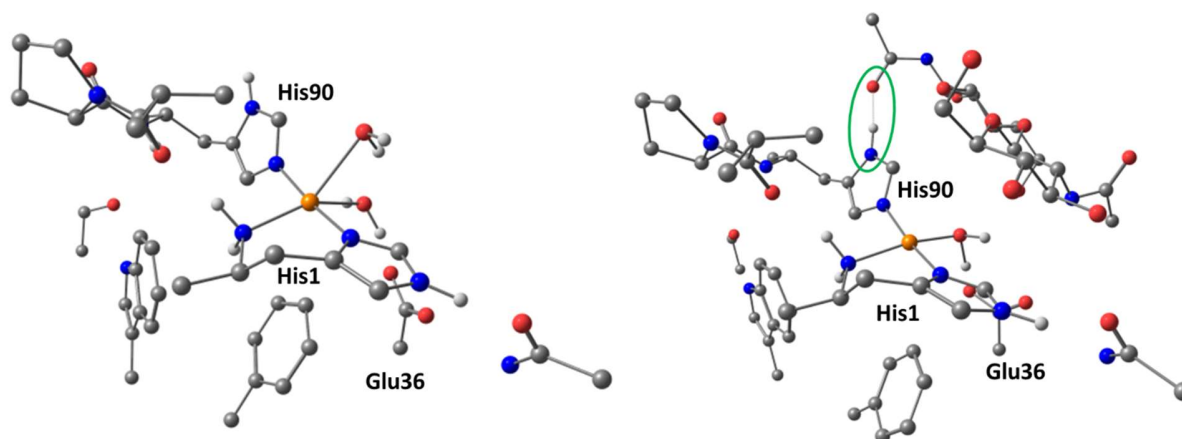
**Figure 5.** a) spin-density contour and b)  $\alpha$ -HOMO electron density (with wavefunction phase depicted in color) contour plots of the five coordinate (left) and four coordinate (right) copper-superoxide complexes within the active site of AA10 LPMOs.

In addition to the thermodynamic stabilization of the Cu(II)-superoxide intermediate, it is expected that these stabilizing effects are enhanced by the lack of a water molecule in the copper coordination sphere, which is only the case when the substrate is bound. Such a complex can be expected to have a longer lifetime than the equivalent one in the absence of substrate where both a water molecule and superoxide coordinate to the copper ion. In this case, as shown previously by Kjaergaard *et al.* for AA9 LPMOs from stopped flow experiments and DFT calculations, the superoxide can be expelled from the copper coordination sphere with an activation barrier of  $\sim 10$  kcal mol<sup>-1</sup> (14), although this value is quite a lot lower than that recently calculated by Caldararu *et al.* in QM/MM calculations for superoxide release from an AA10 LPMO at 19 kcal mol<sup>-1</sup> (57). Our own relaxed-surface-scan DFT calculations for *B/LPMO10A* indicate, in the absence of substrate, that a water-assisted superoxide dissociation from Cu(II) is indeed feasible with an activation barrier of only  $\sim 4$  kcal mol<sup>-1</sup> (Figure S10), similar to the experimental findings of Kjaergaard *et al.* Thus, the Cu(II)-superoxide complex in the absence of substrate appears to be unstable. A coupling mechanism between substrate binding and selective O<sub>2</sub> activation in AA10 LPMOs therefore emerges from this analysis, in which the Cu(II)-superoxide intermediate is kinetically unstable to dissociative elimination in the absence of substrate (possibly followed by reduction of the Cu(II) and formation of O<sub>2</sub>), but is thermodynamically stabilized in the presence of substrate.

### **Mechanism of substrate and O<sub>2</sub> coupling in chitin-active AA10 LPMOs**

The combined EPR data show that addition of chitin results in significant changes in the copper d-orbital electronics. This change in electronics can be explained by the formation of a more axial coordination geometry, and—in what is the most significant perturbation to the electronics of the copper brought about by chitin binding—a large increase in the relative energy/ $d(x^2-y^2)$  character of the SOMO, accompanied by a small reduction in metal-histidine covalency, the latter driven by formation of a hydrogen bond between substrate and a histidine and a reduction in the coordination number of the copper from five to four (Figure 6). Importantly, within the context of O<sub>2</sub> activation at the copper, any reduction in the covalency of the RAMO (redox-active MO, which is the doubly-occupied  $d(x^2-y^2)$  in the Cu(I) oxidation state) coupled to an increase in its relative energy increases the relative stability of a Cu(II)-superoxide by: i) reducing the energy gap between the SOMO of

Cu(II) and the  $\pi^*$  antibonding orbital of  $O_2^-$  (such a strong bond is needed to offset the negative reduction potential of  $O_2$  to  $O_2^-$ ), and ii) maximizing the stabilizing effects of the nephelauxetic expansion which occurs upon the formation of a covalent bond between  $O_2^-$  and Cu(II). This latter effect reveals the contributing role of electron-electron repulsions at the copper within a histidine brace coordination, which—upon formation of a Cu(II)-superoxide—results in a net transfer of spin density from the Cu(II) to the distal atom of the superoxide (table S9), while maintaining the superoxide character of the ligand,  $\Delta r(O-O) = -0.02 \text{ \AA}$ . Such an increase in the spin density at this oxygen atom would be in accord with it acting as the site of hydrogen atom transfer from the substrate, and would further contribute to any coupling mechanism induced by the substrate.



**Figure 6.** DFT-optimized structures of absence and presence of substrate in the active site of *BaAA10*, highlighting the change in Cu coordination geometry in going from five to four ligands. All hydrogen atoms apart from those on the N and O atoms of the metal ligands were hidden for clarity.

## Conclusions

The NMR structures of the *apo*- and Cu(I)-forms of *B/LPMO10A* were determined in order to provide structural information on LPMOs in solution. There are minimal differences between the Cu(I)- and *apo*-structures of *B/LPMO10A*. These differences are centered around the LPMO copper site and are likely related to the structural effects of copper binding. In addition, the PRE effect in NMR spectra of Cu(II)-*B/LPMO10A* was evaluated using parameters calculated from EPR data ( $g_{iso}$ ) and derived from heteronuclear relaxation data, and shown to be consistent with

PREs calculated from *B/LPMO10A* structures. Multi-frequency CW-EPR spectroscopy enabled the determination of accurate spin-Hamiltonian values, showing that the addition of chitin drives a rearrangement of the copper coordination sphere from five- to four-coordinate, accompanied by a reduction in metal-histidine covalency and the associated generation of a high energy SOMO with significant  $d(x^2-y^2)$  character. This orbital essentially emerges as a well-defined frontier orbital at the copper which can then form strong interactions with exogenous ligands such as  $O_2^-$ . These changes at the copper upon substrate binding provide a means by which the formation of a Cu(II)-superoxide can be stabilized. Overall, these results show that the mechanism of substrate- $O_2$  coupling can be effected through rearrangement of the copper coordination geometry and subsequent changes in the d-orbital electronics, with minimal change in the rest of the protein backbone structure. These results underline recent observations in other copper proteins (58) that minimal structural changes can be coupled to large electronic changes at the copper active site.

In a wider observation on the mechanisms of LPMOs, this work shows that substrate binding is coupled to the activation of the  $O_2$  co-substrate. For LPMOs, with their exposed copper sites, substrate-induced activation of the catalytic center is an attractive scenario since this will reduce off-pathway reactions that may lead to enzyme inactivation. As such, any studies on LPMOs must take into account the fact that specific LPMOs may be associated with specific substrates, an association through which the “on-pathway” coupled mechanism operates. As a caution, therefore, investigators need to be aware that any studies performed on LPMOs not correctly bound to their natural substrate may not have an on-pathway mechanism available to them, potentially leading to rapid enzyme inactivation via indiscriminate redox chemistry. Finally, the results presented in this current study demonstrate the power of an integrated NMR/EPR spectroscopic approach to studying LPMOs.

## **Materials and Methods**

Detailed information for all experimental procedures is provided in the Materials and Methods section of the SI.

**Sample Preparation.** Isotope-labeled and non-labeled *B/LPMO10A* were recombinantly produced in *E. coli* using isotope-enriched ( $^{13}C$ ,  $^{15}N$ ) minimal medium or LB, respectively, and purified by multiple chromatographic steps as described previously (38, 39, 59).



**Enzyme Activity.** Substrate degradation was performed using standard reaction conditions and product formation was analyzed using hydrophilic interaction chromatography (HILIC) as described by Loose *et al.* (60). Reactions were set up with 10 mg/mL  $\alpha$ - or  $\beta$ -chitin, 1  $\mu$ M of Cu(II)-loaded *B/LPMO10A* and 2 mM ascorbic acid in 20 mM Tris/HCl buffer (pH 8.0) in a shaking incubator at 40 °C.

**NMR spectroscopy.** NMR spectra of *B/LPMO10A* were recorded on Bruker Avance III 600 MHz and 800 MHz spectrometers. 3D NOESY-edited spectra were recorded, assigned and integrated. NOE cross-peak intensities were converted to distance restraints and were used together with dihedral torsion angles predicted by TALOS-N (61) as input for structure calculations in CYANA (62). The 20 conformers with lowest CYANA target function values were energy-minimized using YASARA (63). The *apo-B/LPMO10A* ensemble was deposited in the Protein DataBank under the accession ID 5LW4. The Cu(I)-*B/LPMO10A* ensemble was deposited in the Protein DataBank under the accession ID 6TWE. Cu(II) PREs were measured by analyzing  $^{15}\text{N}$ -HSQC signal intensities before and after addition of Cu(II) to *apo-B/LPMO10A*. PREs were calculated from the structural ensemble.

**EPR spectroscopy.** EPR experiments were recorded on *B/LPMO10A* and  $^{15}\text{N}$ -*B/LPMO10A* loaded with  $^{63}\text{Cu(II)}$ , with and without squid pen  $\beta$ -chitin. CW X-band spectra were acquired at 165 K on a Bruker EMX spectrometer operating at  $\sim 9.30$  GHz, with modulation amplitude of 4 G, modulation frequency of 100 kHz and microwave power of 10.02 mW. CW Q-band spectra were acquired at 113 K on a Jeol JES-X320 spectrometer operating at  $\sim 34.7$  GHz, with modulation width 0.8 mT and microwave power of 1.0 mW. Hyperfine sublevel correlation (HYSCORE) spectra were collected on a Bruker ElexSys E580 spectrometer equipped with an Oxford CF 935 helium flow cryostat. The rf pulse in Davies ENDOR spectra was generated by the Bruker DICE system and amplified by a 60 dB gain ENI A-500 W amplifier. Spectral simulations were carried out using EasySpin 5.2.6 (64) integrated into MATLAB R2017a software.

## Acknowledgements

This work was financed by SO-funds from NTNU, the Norwegian University of Science and Technology, the Novo Nordisk Foundation (grant number NNF18OC0032242), the OXYMOD project, and the Norwegian NMR Platform (grants and 269408 and 226244 from the Research

Council of Norway, respectively). Further support came from the European Union through funding of a Marie Curie international training network called “Oxytrain” (contract number 722390). Part of the work was conducted at the NMR laboratory at Aalborg University which is supported by the Obel Family, SparNord and Carlsberg Foundations. PHW and LC thank the UK Biotechnology and Biological Sciences Research Council (BB/L001926/1) for funding. AP thanks Mrs Lesley Wild for funding. We thank Prof Eric McInnes for useful discussions and Dr Amga Baldansuren, Dr Floriana Tuna, Dr Alena Sheveleva and Mr Adam Brookfield for collecting the HYSORE and ENDOR data.

## References

1. J. G. Fernandez; D. E. Ingber, Manufacturing of large-scale functional objects using biodegradable chitosan bioplastic. *Macromol. Mater. Eng.* **299**, 932-938 (2014).
2. S. Ifuku; H. Saimoto, Chitin nanofibers: preparations, modifications, and applications. *Nanoscale* **4**, 3308-3318 (2012).
3. M. Rinaudo, Chitin and chitosan: Properties and applications. *Prog. Polym. Sci.* **31**, 603-632 (2006).
4. V. Lombard; H. Golaconda Ramulu; E. Drula; P. M. Coutinho; B. Henrissat, The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490-D495 (2013).
5. A. Levasseur; E. Drula; V. Lombard; P. M. Coutinho; B. Henrissat, Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol. Biofuels* **6**, 41 (2013).
6. S. J. Horn; G. Vaaje-Kolstad; B. Westereng; V. Eijsink, Novel enzymes for the degradation of cellulose. *Biotechnol. Biofuels* **5**, 45 (2012).
7. Z. Forsberg *et al.*, Cleavage of cellulose by a CBM33 protein. *Protein Sci.* **20**, 1479-1483 (2011).
8. C. M. Phillips; W. T. Beeson; J. H. Cate; M. A. Marletta, Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem. Biol.* **6**, 1399-1406 (2011).
9. G. Vaaje-Kolstad *et al.*, An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* **330**, 219-222 (2010).
10. M. Eibinger *et al.*, Cellulose surface degradation by a lytic polysaccharide monooxygenase and its effect on cellulase hydrolytic efficiency. *J. Biol. Chem.* **289**, 35929-35938 (2014).
11. W. T. Beeson; V. V. Vu; E. A. Span; C. M. Phillips; M. A. Marletta, Cellulose degradation by polysaccharide monooxygenases. *Annu. Rev. Biochem.* **84**, 923-946 (2015).
12. A. G. Hamre; K. B. Eide; H. H. Wold; M. Sørli, Activation of enzymatic chitin degradation by a lytic polysaccharide monooxygenase. *Carbohydr. Res.* **407**, 166-169 (2015).
13. W. T. Beeson; C. M. Phillips; J. H. D. Cate; M. A. Marletta, Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J. Am. Chem. Soc.* **134**, 890-892 (2012).
14. C. H. Kjaergaard *et al.*, Spectroscopic and computational insight into the activation of O<sub>2</sub> by the mononuclear Cu center in polysaccharide monooxygenases. *Proc. Natl. Acad. Sci.* **111**, 8797-8802 (2014).
15. B. Bissaro *et al.*, Oxidative cleavage of polysaccharides by monocopper enzymes depends on H<sub>2</sub>O<sub>2</sub>. *Nat. Chem. Biol.* **13**, 1123-1128 (2017).
16. R. J. Quinlan *et al.*, Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc. Natl. Acad. Sci.* **108**, 15079-15084 (2011).
17. S. Kim; J. Ståhlberg; M. Sandgren; R. S. Paton; G. T. Beckham, Quantum mechanical calculations suggest that lytic polysaccharide monooxygenases use a copper-oxyl, oxygen-rebound mechanism. *Proc. Natl. Acad. Sci.* **111**, 149-154 (2014).
18. J. A. Hangasky; A. T. Iavarone; M. A. Marletta, Reactivity of O<sub>2</sub> versus H<sub>2</sub>O<sub>2</sub> with polysaccharide monooxygenases. *Proc. Natl. Acad. Sci.* **115**, 4915-4920 (2018).
19. P. H. Walton; G. J. Davies, On the catalytic mechanisms of lytic polysaccharide monooxygenases. *Curr. Opin. Chem. Biol.* **31**, 195-207 (2016).

20. K. K. Meier *et al.*, Oxygen activation by Cu LPMOs in recalcitrant carbohydrate polysaccharide conversion to monomer sugars. *Chem. Rev.* **118**, 2593-2635 (2018).
21. L. Bertini *et al.*, Catalytic mechanism of fungal lytic polysaccharide monooxygenases investigated by first-principles calculations. *Inorg. Chem.* **57**, 86-97 (2018).
22. K. E. H. Frandsen *et al.*, The molecular basis of polysaccharide cleavage by lytic polysaccharide monooxygenases. *Nat. Chem. Biol.* **12**, 298 (2016).
23. L. Ciano; G. J. Davies; W. B. Tolman; P. H. Walton, Bracing copper for the catalytic oxidation of C–H bonds. *Nat. Catal.* **1**, 571-577 (2018).
24. E. D. Hedegård; U. Ryde, Targeting the reactive intermediate in polysaccharide monooxygenases. *J. Biol. Inorg. Chem.* **22**, 1029-1037 (2017).
25. E. D. Hedegård; U. Ryde, Molecular mechanism of lytic polysaccharide monooxygenases. *Chem. Sci.* **9**, 3866-3880 (2018).
26. B. Wang *et al.*, QM/MM Studies into the H<sub>2</sub>O<sub>2</sub>-dependent activity of lytic polysaccharide monooxygenases: evidence for the formation of a caged hydroxyl radical intermediate. *ACS Catal.* **8**, 1346-1351 (2018).
27. B. Wang; P. H. Walton; C. Rovira, Molecular mechanisms of oxygen activation and hydrogen peroxide formation in lytic polysaccharide monooxygenases. *ACS Catal.* **9**, 4958-4969 (2019).
28. B. Bissaro *et al.*, Molecular mechanism of the chitinolytic peroxygenase reaction. *Proc. Natl. Acad. Sci.* **117**, 1504-1513 (2020).
29. A. S. Borisova *et al.*, Structural and functional characterization of a lytic polysaccharide monooxygenase with broad substrate specificity. *J. Biol. Chem.* **290**, 22955-22969 (2015).
30. D. Kracher; M. Andlar; P. G. Furtmüller; R. Ludwig, Active-site copper reduction promotes substrate binding of fungal lytic polysaccharide monooxygenase and reduces stability. *J. Biol. Chem.* (2017).
31. T. J. Simmons *et al.*, Structural and electronic determinants of lytic polysaccharide monooxygenase reactivity on polysaccharide substrates. *Nature Communications* **8**, 1064 (2017).
32. B. Bissaro; I. Isaksen; G. Vaaje-Kolstad; V. G. H. Eijsink; Å. K. Røhr, How a lytic polysaccharide monooxygenase binds crystalline chitin. *Biochemistry* **57**, 1893-1906 (2018).
33. A. Paradisi *et al.*, Formation of a copper(II)–tyrosyl complex at the active site of lytic polysaccharide monooxygenases following oxidation by H<sub>2</sub>O<sub>2</sub>. *J. Am. Chem. Soc.* **141**, 18585-18599 (2019).
34. F. L. Aachmann; M. Sørli; G. Skjåk-Bræk; V. G. H. Eijsink; G. Vaaje-Kolstad, NMR structure of a lytic polysaccharide monooxygenase provides insight into copper binding, protein dynamics, and substrate interactions. *Proc. Natl. Acad. Sci.* **109**, 18779-18784 (2012).
35. G. R. Hemsworth *et al.*, The copper active site of CBM33 polysaccharide oxygenases. *J. Am. Chem. Soc.* **135**, 6069-6077 (2013).
36. Z. Forsberg *et al.*, Structural determinants of bacterial lytic polysaccharide monooxygenase functionality. *J. Biol. Chem.* **293**, 1397-1412 (2018).
37. J. S. M. Loose *et al.*, Multipoint precision binding of substrate protects lytic polysaccharide monooxygenases from self-destructive off-pathway processes. *Biochemistry* **57**, 4114-4124 (2018).
38. Z. Forsberg *et al.*, Comparative study of two chitin-active and two cellulose-active AA10-type lytic polysaccharide monooxygenases. *Biochemistry* **53**, 1647-1656 (2014).

39. G. Courtade *et al.*,  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$  resonance assignment of the chitin-active lytic polysaccharide monooxygenase BILPMO10A from *Bacillus licheniformis*. *Biomol. NMR Assign.* **9**, 207-210 (2015).
40. G. R. Hemsworth; B. Henrissat; G. J. Davies; P. H. Walton, Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nat. Chem. Biol.* **10**, 122-126 (2014).
41. R. C. Gregory *et al.*, Activity, stability and 3-D structure of the Cu(II) form of a chitin-active lytic polysaccharide monooxygenase from *Bacillus amyloliquefaciens*. *Dalton Trans.* **45**, 16904-16912 (2016).
42. I. Bertini; R. Pierattelli, Copper (II) proteins are amenable for NMR investigations. *Pure Appl. Chem.* **76**, 321-333 (2004).
43. S. Yao; M. G. Hinds; R. S. Norton, Improved estimation of protein rotational correlation times from  $^{15}\text{N}$  relaxation measurements. *J. Magn. Reson.* **131**, 347-350 (1998).
44. T. Murakami; T. Takei; Y. Ishikawa, Spectroscopic properties and electronic states of five-coordinate copper(II) complexes with linear pentadentate ligands containing two amide groups. *Polyhedron* **16**, 89-93 (1997).
45. M. Gudmundsson *et al.*, Structural and electronic snapshots during the transition from a Cu(II) to Cu(I) metal center of a lytic polysaccharide monooxygenase by X-ray photoreduction. *J. Biol. Chem.* **289**, 18782-18792 (2014).
46. Z. Forsberg *et al.*, Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. *Proc. Natl. Acad. Sci.* **111**, 8446-8451 (2014).
47. Z. Forsberg *et al.*, Structural and functional analysis of a lytic polysaccharide monooxygenase important for efficient utilization of chitin in *Cellvibrio japonicus*. *J. Biol. Chem.* (2016).
48. I. V. Ovchinnikov; V. N. Konstantinov, Extra absorption peaks in EPR spectra of systems with anisotropic g-tensor and hyperfine structure in powders and glasses. *Journal of Magnetic Resonance (1969)* **32**, 179-190 (1978).
49. G. R. Hemsworth; L. Ciano; G. J. Davies; P. H. Walton, Chapter Three - Production and spectroscopic characterization of lytic polysaccharide monooxygenases. In *Methods Enzymol.*, (Armstrong, F., Ed. Academic Press, 2018) Vol. 613, pp 63-90.
50. F. Neese, A critical evaluation of DFT, including time-dependent DFT, applied to bioinorganic chemistry. *J. Biol. Inorg. Chem.* **11**, 702-711 (2006).
51. A. A. Gewirth; S. L. Cohen; H. J. Schugar; E. I. Solomon, Spectroscopic and theoretical studies of the unusual EPR parameters of distorted tetrahedral cupric sites: correlations to x-ray spectral features of core levels. *Inorg. Chem.* **26**, 1133-1146 (1987).
52. F. Neese, Sum-over-states based multireference ab initio calculation of EPR spin Hamiltonian parameters for transition metal complexes. A case study. *Magn. Reson. Chem.* **42**, S187-S198 (2004).
53. H. A. Kuska; M. T. Rogers; R. E. Drullinger, Effect of substituents on the anisotropic electron spin resonance parameters in copper acetylacetonates. *J. Phys. Chem.* **71**, 109-114 (1967).
54. R. E. Watson; A. J. Freeman, Origin of Effective Fields in Magnetic Materials. *Phys. Rev.* **123**, 2027-2047 (1961).
55. C. Remenyi; R. Reviakine; M. Kaupp, Density functional study of EPR parameters and spin-density distribution of azurin and other blue copper proteins. *J. Phys. Chem. B* **111**, 8290-8304 (2007).

56. F. Jiang; J. McCracken; J. Peisach, Nuclear quadrupole interactions in copper(II)-diethylenetriamine-substituted imidazole complexes and in copper(II) proteins. *J. Am. Chem. Soc.* **112**, 9035-9044 (1990).
57. O. Caldararu; E. Oksanen; U. Ryde; E. D. Hedegård, Mechanism of hydrogen peroxide formation by lytic polysaccharide monooxygenase. *Chem. Sci.* **10**, 576-586 (2019).
58. A. J. Leguto *et al.*, Dramatic Electronic Perturbations of CuA Centers via Subtle Geometric Changes. *J. Am. Chem. Soc.* **141**, 1373-1381 (2019).
59. G. Courtade; S. B. Le; G. I. Sætrum; T. Brautaset; F. L. Aachmann, A novel expression system for lytic polysaccharide monooxygenases. *Carbohydr. Res.* **448**, 212-219 (2017).
60. J. S. M. Loose; Z. Forsberg; M. W. Fraaije; V. G. H. Eijssink; G. Vaaje-Kolstad, A rapid quantitative activity assay shows that the *Vibrio cholerae* colonization factor GbpA is an active lytic polysaccharide monooxygenase. *FEBS Lett.* **588**, 3435-3440 (2014).
61. Y. Shen; A. Bax, Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227-241 (2013).
62. P. Güntert, Automated NMR Structure Calculation With CYANA. In *Protein NMR Techniques*, (Downing, A. K., Ed. Humana Press, 2004) pp 353-378.
63. E. Krieger; G. Koraimann; G. Vriend, Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins: Struct., Funct., Bioinf.* **47**, 393-402 (2002).
64. S. Stoll; A. Schweiger, EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *J. Magn. Reson.* **178**, 42-55 (2006).

## Supporting Information for

### Mechanistic basis of substrate-O<sub>2</sub> coupling within a chitin-active lytic polysaccharide monooxygenase: an integrated NMR/EPR study

Gaston Courtade<sup>a,1</sup>, Luisa Ciano<sup>b,c,1,2</sup>, Alessandro Paradisi<sup>b</sup>, Peter Lindley<sup>b</sup>, Zarah Forsberg<sup>d</sup>, Morten Sørli<sup>d</sup>, Reinhard Wimmer<sup>e</sup>, Gideon J. Davies<sup>b</sup>, Vincent G. H. Eijssink<sup>d</sup>, Paul H. Walton<sup>b,3</sup>, Finn L. Aachmann<sup>a,3</sup>

<sup>1</sup> these authors contributed equally to this work.

#### Author affiliation

<sup>a</sup> NOBIPOL, Department of Biotechnology and Food Science, NTNU Norwegian University of Science and Technology, Sem Sælands vei 6/8, N-7491 Trondheim, Norway

<sup>b</sup> Department of Chemistry, University of York, Heslington, York YO10 5DD, UK

<sup>c</sup> School of Chemistry and Photon Science Institute, University of Manchester, Oxford Road, Manchester, M13 9PL, UK

<sup>d</sup> Faculty of Chemistry, Biotechnology and Food Science, NMBU Norwegian University of Life Sciences, N-1432 Ås, Norway

<sup>e</sup> Department of Chemistry and Bioscience, Aalborg University, Frederik Bajers vej 7H, 9220 Aalborg Ø, Denmark

<sup>2</sup>Current address: School of Chemistry, University of Nottingham, University Park, Nottingham, NG7 2RD, UK

#### Corresponding authors

<sup>3</sup> To whom correspondence should be addressed.

Finn L. Aachmann

NOBIPOL, Department of Biotechnology and Food Science, NTNU Norwegian University of Science and Technology, Sem Sælands vei 6/8, N-7491 Trondheim, Norway

Tel.: +47 73593317

Paul H. Walton

Department of Chemistry, University of York, Heslington, York YO10 5DD, UK

Tel: +44 1904 324457

E-mail: [finn.l.aachmann@ntnu.no](mailto:finn.l.aachmann@ntnu.no), [paul.walton@york.ac.uk](mailto:paul.walton@york.ac.uk)

## TABLE OF CONTENTS

### SI Materials and Methods

**Supplementary Discussion** – Calculations used in EPR analysis

**Supplementary Discussion** - DFT calculations in presence of substrate

**Table S1** – Structural statistics of *apo-B/LPMO10A*

**Table S2** – Davies  $^{14}\text{N}$  ENDOR simulation parameters for *B/LPMO10A*

**Table S3** – The functional and basis set schemes used to calculate the EPR properties of *B/LPMO10A*

**Table S4** – Calculated spin Hamiltonian parameters for the ‘resting state’ Model A of *B/LPMO10A*

**Table S5** – Calculated spin Hamiltonian parameters for ‘4-coordinate’ Model B of *B/LPMO10A*

**Table S6** – Calculated spin Hamiltonian parameters for ‘substrate bound’ Model C of *B/LPMO10A*

**Table S7** – EPR properties of *B/LPMO10A* with and without  $\beta$ -chitin

**Table S8** – EPR properties of the remote nitrogen atoms of the histidine rings of *B/LPMO10A* with and without  $\beta$ -chitin, determined experimentally and by DFT calculations

**Table S9** – Löwdin spin population analysis of 4-coordinate and 5-coordinate superoxide models (adding up to 200% for a total of 2 spins).

**Table S10** – Calculated superoxide binding energies using cluster models (D) and (E).

**Figure S1** – *B/LPMO10A* activity and binding towards chitin

**Figure S2** – Effect of Cu(II)-binding on *apo-B/LPMO10A*

**Figure S3** – Backbone dynamics of *apo-B/LPMO10A*

**Figure S4** – X and Q band CW-EPR spectra

**Figure S5** – Davies  $^{14}\text{N}$  ENDOR spectra of  $^{63}\text{Cu-B/LPMO10A}$

**Figure S6** –  $^{14}\text{N}$  HYSCORE of  $^{63}\text{Cu-B/LPMO10A}$

**Figure S7** – Starting structure for geometry optimization of Cu(II)-*B/LPMO10A*

**Figure S8** – Geometry optimized structure of Cu(II)-*B/LPMO10A*

**Figure S9** – Geometry optimized structure of Cu(II)-*B/LPMO10A* in the ‘substrate-bound’ case

**Figure S10** – Superoxide release from Cu(II)-AA10. DFT-optimized structures with Cu-O distances fixed at 2.20, 2.50 and 3.00 Å.



**Figure S11** – Geometry optimized structures of 4-coordinate and 5-coordinate Cu(II)-*B/LPMO10A*-superoxide cluster models and resulting geometries following superoxide release.

## References

## Materials and Methods

### Protein production and functional characterization

Production and purification of *B/LPMO10A* were performed as previously described (1). Copper saturation of purified protein was achieved by incubation with a 3-fold molar surplus of Cu(II)SO<sub>4</sub> at room temperature for 30 min followed by removal of excess copper using a PD MidiTrap G-25 desalting column (GE Healthcare; Uppsala, Sweden) equilibrated with 20 mM Tris/HCl buffer pH 8.0 (2). Analysis of oxidized chito-oligosaccharides released in reactions with ball-milled shrimp shell (*Pandalus borealis*)  $\alpha$ -chitin with a particle size of  $\sim\mu\text{m}$  (Chitinor AS; Senjahopen, Norway) or ball milled squid  $\beta$ -chitin with a particle size of  $< 0.85$  mm (France chitin; Orange, France) was performed using hydrophilic interaction chromatography (HILIC) as described by Loose *et al.* (2). The enzymatic reaction mixtures contained 1  $\mu\text{M}$  of Cu(II)-loaded *B/LPMO10A*, 10 mg/mL suspension of  $\alpha$ - or  $\beta$ -chitin, 20 mM Tris/HCl buffer, pH 8.0, and 2 mM ascorbic acid. Reactions were incubated at 40 °C in an Eppendorf Thermomixer set to 800 rpm for 24 hours.

Binding experiments were performed using 2  $\mu\text{M}$  of Cu(II)-loaded *B/LPMO10A* and 10 g/L  $\beta$ -chitin. The binding reactions was carried out in 50 mM sodium phosphate buffer pH 7.0 and incubated at 40 °C in an Eppendorf Thermomixer set to 800 rpm. At various time points (0, 5, 15, 30, 60, 120 and 240 min) a sample was taken and filtrated, using a 96-well filter plate (Millipore) operated by a Millipore vacuum manifold, to remove insoluble substrate and substrate-bound protein. The concentration of protein in the supernatant (e.g. non-substrate bound protein) was determined using the Bradford assay and measuring A<sub>595</sub> in an Eppendorf Biophotometer (Eppendorf, Hamburg).

### Sample preparation for NMR

Cloning, protein production using an LPMO expression cassette, and purification of the natural abundance and isotope-labeled (<sup>15</sup>N and <sup>13</sup>C or <sup>15</sup>N) *B/LPMO10A* from *Bacillus licheniformis* (UniProt entry Q62YN7, residues 32-203), as well as conditions for NMR spectroscopy measurements, have been described previously (3, 4). In order to obtain the *apo* form of

*B/LPMO10A*, protein samples were incubated in 25 mM sodium phosphate, pH 5.5, containing 10 mM NaCl and 8 mM EDTA at 4 °C overnight, prior to exchanging the buffer to either 25 mM sodium phosphate, pH 5.5, 10 mM NaCl, or 20 mM MES, pH 5.5, using Vivaspin 6 spin-columns (5 kDa cut-off, Sartorius).

NMR samples with Cu(II) were prepared and recorded in the following way. First, the buffer of a solution of <sup>13</sup>C- and <sup>15</sup>N-labeled *apo-B/LPMO10A* was changed to a 25 mM sodium acetate buffer, pH 5.5, 10 mM NaCl. The protein solution was then concentrated to 0.1 mM and ~450 μL, and reference spectra (1D-proton, <sup>15</sup>N-HSQC and aromatic <sup>13</sup>C-HSQC) were recorded. Then, CuSO<sub>4</sub> was added to concentrations of 0.15 mM, and spectra were recorded with CuSO<sub>4</sub> present.

NMR samples with Cu(I) were prepared and recorded in the following way. First, the buffer of a solution of <sup>13</sup>C- and <sup>15</sup>N-labeled *apo-B/LPMO10A* was changed to 20 mM MES, pH 5.5, followed by concentrating the protein solution to 0.1 mM and ~450 μL, after which reference spectra (1D-proton and <sup>15</sup>N-HSQC) were recorded. Then, the sample was incubated for 48 h in BBL GasPak Jar (Becton Dickinson, NJ, USA) chamber with a BBL CO<sub>2</sub> Generator bag (Becton Dickinson, NJ, USA) to remove oxygen gas. The chamber was transferred to a Whitley A45 Anaerobic Workstation (Don Whitley Scientific Limited, UK), where the protein sample was transferred to a 5 mm NMR tube and Cu(I) was added in the form of a pellet (appr. 1 mg) of Cu(I)Cl. Finally, the tube was sealed with a rubber septum and parafilm. NMR spectra (1D-proton, <sup>15</sup>N-HSQC, aliphatic <sup>13</sup>C-HSQC, aromatic <sup>13</sup>C-HSQC, <sup>15</sup>N-edited-NOESY-HSQC and <sup>13</sup>C-edited-NOESY-HSQC) were recorded immediately after.

### **NMR spectroscopy**

NMR spectra of *B/LPMO10A* were recorded at 25 °C on a Bruker Avance III 600 MHz or Avance III 800 MHz spectrometer, both equipped with a 5 mm Z-gradient CP-TCI (H/C/N) cryoprobe, at the NT-NMR-Center/ Norwegian NMR Platform in Trondheim, Norway, and on a Bruker Avance III 600 MHz equipped with a 5 mm Z-gradient Prodigy TCI (H/C/N) cryoprobe, at the Department of Chemistry and Biosciences, Aalborg University, Aalborg, Denmark. NMR data were processed using Bruker TopSpin version 3.5. NMR spectral analysis was performed using CARA version 1.5.5 (5). The NMR assignment of *B/LPMO10A* has been published elsewhere (3). For structure

determination, three-dimensional  $^{13}\text{C}$ -edited and  $^{15}\text{N}$ -edited NOESY-HSQC spectra, as well as  $^1\text{H}$ - $^1\text{H}$  NOESY spectra were recorded. NOE cross-peaks were manually identified, assigned, and integrated using the NEASY program within CARA version 1.5.5 (5). The  $\{^1\text{H}\}$ - $^{15}\text{N}$  heteronuclear NOEs were derived with Protein Dynamic Center software version 2.3.3 from Bruker BioSpin using two independently measured and integrated  $^{15}\text{N}$ - heteronuclear correlated spectra with and without  $^1\text{H}$  saturation (6). Nuclear magnetic relaxation time measurements of  $^{15}\text{N}$  nuclei ( $T_1$  and  $T_2$ ) were analyzed with Protein Dynamic Center software version 2.3.3 from Bruker BioSpin, where the data from  $^{15}\text{N}$ -HSQC-type spectra acquired with different relaxation delays was exponentially fitted (6, 7). The rotational correlation time was estimated using the  $T_1/T_2$  ratio, assuming overall isotropic tumbling of the protein. Secondary structure elements were analyzed using the web-based version of the TALOS-N software [spin.niddk.nih.gov/bax/nmrserver/talosn/](http://spin.niddk.nih.gov/bax/nmrserver/talosn/) (8) using selected chemical shifts (N,  $\text{H}^{\text{N}}$ ,  $\text{C}'$ ,  $\text{C}^{\alpha}$ ,  $\text{C}^{\beta}$ ,  $\text{H}^{\alpha}$  and  $\text{H}^{\beta}$ ).

### PRE calculation

The PRE effect was calculated using the atomic coordinates for the 20-conformer ensemble of Cu(I)-B/LPMO10A (PDB ID: 6TWE) as input. Transverse ( $R_{2,\text{para}}$ ) PRE rates are described by the Solomon-Bloembergen equation (9, 10):

$$R_{2,\text{para}} = \frac{1}{15} \left( \frac{\mu_0}{4\pi} \right)^2 \gamma_I^2 g^2 \mu_B^2 s_e (s_e + 1) [4J_{\text{SBMF}}(0) + 3J_{\text{SBMF}}(\omega_I)]$$

where  $\mu_0$  is the permeability of vacuum,  $\gamma_I$  is the proton gyromagnetic ratio,  $g$  is the electron g-factor,  $\mu_B$  is the free electron magnetic moment,  $s_e$  is the paramagnetic electron spin number ( $s_e = 1/2$  for Cu(II)),  $\frac{\omega_I}{2\pi}$  is the Larmor frequency of the proton, and  $J_{\text{SBMF}}(\omega)$  is the model-free (11) extension generalized spectral density function, defined by the following equation:

$$J_{\text{SBMF}}(\omega) = \langle r^{-6} \rangle \left[ \frac{S^2 \tau_c}{1 + \omega^2 \tau_c^2} + \frac{(1 - S^2) \tau_t}{1 + \omega^2 \tau_t^2} \right]$$

where  $r$  is the distance between the proton and the paramagnetic electron,  $S^2$  is the square of the generalized order parameter (see (12) for details regarding its calculation),  $\tau_c$  is the overall protein rotational correlation time, and  $\tau_t$  is the total correlation time defined as  $(\tau_c^{-1} + \tau_i^{-1})^{-1}$ , where  $\tau_i$  is the correlation time for the internal motion.

The calculated reduction in signal intensity  $\frac{I_{para}}{I_{dia}}$  can then be calculated using the following equation:

$$\frac{I_{para}}{I_{dia}} = \frac{R_{2,dia} e^{-R_{2,para} t}}{R_{2,dia} + R_{2,para}}$$

where  $R_{2,dia}$  is the transverse relaxation rate in the absence of Cu(II) and  $t$  is the length of the  $^1\text{H}$  transverse relaxation evolution during INEPT coherence transfers in  $^{15}\text{N}$ -HSQC.

For *B/LPMO10A*,  $g$  was set to 2.13, which is the average of the calculated  $g_{iso}$  (Table 1);  $R_{2,dia}$  was set to  $12.5 \text{ s}^{-1}$ , which is the average of  $1/T_2$  (Figure S3);  $\tau_c$  was set to 10.2 ns, calculated from the average  $T_1/T_2$  (Figure S3);  $\tau_i$  was set to 500 ps. The PRE calculation was performed using a script available at <https://github.com/gcourtade/BILPMO10A>. Part of the script uses code available from <https://github.com/KULL-Centre/DEERpredict>.

### NMR structure calculation

NOE cross-peak intensities were converted into distance restraints using the CALIBA (13) subroutine in CYANA 3.97 (13, 14). Dihedral torsion angles ( $\phi$ ,  $\psi$ ) predicted by TALOS-N (8) were included as conformational restraints, as was one disulfide bridge (Cys45–Cys56). Based on this input, the structure was calculated using the torsion angle dynamics program CYANA 3.97 (14). The structure calculation started by generating 200 conformers with random torsion angles, and the dihedral angles in each conformer were optimized using simulated annealing in 10,000 steps, to fit the restraints. The 20 conformers with the lowest CYANA target function values were energy-minimized using YASARA (15), first *in vacuo*, followed by using water as the explicit solvent and calculating electrostatics by applying the particle mesh Ewald method (16). In both these steps the YASARA force field (17) was applied. The coordinates of the minimized *apo-B/LPMO10A* conformers have been deposited in the Protein Data Bank under the ID 5LW4.

The Cu(I) structure of *B/LPMO10A* was generated in YASARA in the following way: (i)  $^{15}\text{N}$ -edited and  $^{13}\text{C}$ -edited NOESY-HSQC spectra that had been recorded in the presence of Cu(I) were analyzed, integrated and converted to distance constraints. (ii) A constraint for copper was introduced by modelling a copper atom that was placed arbitrarily equidistant from the active site histidine atoms (N-His32,  $\text{N}^{\delta 1}$ -His32,  $\text{N}^{\epsilon 1}$ -His121, ). Its geometry was constrained using partial

charges for the copper-site histidines and force constraints for the bonds and angles describing the copper-histidine brace interaction, obtained by Bissaro *et al.* (18) from a minimal model of the truncated histidine brace of *SmLPMO10A*. (iii) These new constraints were incorporated to each of the 20 conformers from the *apo-BILPMO10A* ensemble by running 100 ps of restrained molecular dynamics in YASARA to produce the Cu(I)-ensemble. The YASARA macro used for the calculation is available at <https://github.com/gcourtade/BILPMO10A>, The coordinates of the Cu(I)-*BILPMO10A* conformers have been deposited in the Protein Data Bank under the ID 6TWE.

## EPR spectroscopy

### *Preparation of $^{63}\text{Cu(II)}$ stock*

A 30% solution of  $\text{H}_2\text{O}_2$  (1 mL) was very slowly added to 300  $\mu\text{L}$  of conc.  $\text{H}_2\text{SO}_4$  cooled in ice. Pure  $^{63}\text{Cu}$  sheets (60 mg, purchased from Sigma Aldrich) were added to this solution, which turned progressively blue. After 36 h, the solution was decanted and the undissolved copper metal was added to a freshly prepared  $\text{H}_2\text{O}_2/\text{H}_2\text{SO}_4$  mixture. All the copper was dissolved after a further 18 h period. The two solutions were combined and bright blue crystals appeared within 48 h, which were then filtered, washed with EtOH and dried.

### *CW and HYSCORE EPR experiments*

The *apo* forms of *B/LPMO10A* and  $^{15}\text{N}$ -*B/LPMO10A* were copper loaded by addition of 0.9 equivalents of  $^{63}\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$  (prepared as described above) in water. Continuous wave (CW) X-band frozen solution EPR spectra of 0.29 mM and 0.17 mM solutions of  $^{63}\text{Cu(II)}$ -*B/LPMO10A* and  $^{63}\text{Cu(II)}$ - $^{15}\text{N}$ -*B/LPMO10A*, respectively, in 10% *v/v* glycerol at pH 5.5 (20 mM MES buffer) and 165 K with and without squid pen  $\beta$ -chitin were acquired on a Bruker EMX spectrometer operating at  $\sim 9.30$  GHz, with modulation amplitude of 4 G, modulation frequency of 100 kHz and microwave power of 10.02 mW (4 scans). (Note: absence of glycerol did not affect the EPR spectra). CW Q-band frozen solution spectra of a 1.3 mM solution of  $^{63}\text{Cu(II)}$ - $^{15}\text{N}$ -*B/LPMO10A* with and without squid pen  $\beta$ -chitin at pH 5.5 (20 mM MES buffer) and 113 K were acquired on a Jeol JES-X320 spectrometer operating at  $\sim 34.7$  GHz, with modulation width 0.8 mT and microwave power of 1.0 mW (8 scans).

Spectral simulations were carried out using EasySpin 5.2.6 (19) integrated into MATLAB R2017a software. Simulation parameters are given in Table 1.  $g_3$  and  $|A_3|$  values were determined accurately from the absorptions at low field. It was assumed that  $g$  and  $A$  tensors were axially coincident. Accurate determination of the  $g_1$ ,  $g_2$ ,  $|A_1|$  and  $|A_2|$  was obtained by simultaneous fitting of both X and Q band spectra. The spectra obtained upon addition of  $\beta$ -chitin presented a mixture of free enzyme and enzyme bound to the substrate. Spectra were collected up to 3 days after addition of chitin, but the free:bound ratio did not improve over time. The simulations of the chitin-bound form of the enzyme were performed after subtraction of the normalized *B/LPMO10A* or

$^{15}\text{N}$ -*B/LPMO10A* (40% and 30% of free enzyme present in the samples, respectively) from the corresponding spectra.

Hyperfine sublevel correlation (HYSCORE) spectra were collected on a 1.7 mM sample of  $^{63}\text{Cu}$ - $^{15}\text{N}$ -*B/LPMO10A* or a 1.8 mM sample of  $^{63}\text{Cu}$ -*B/LPMO10A* before and after addition of squid pen  $\beta$ -chitin in 20 mM sodium phosphate buffer pH 5.5 on a Bruker ElexSys E580 spectrometer equipped with an Oxford CF 935 helium flow cryostat. The  $^{15}\text{N}$  HYSCORE spectra were recorded near parallel (3060 or 3090 G) and near perpendicular (3395 G) directions employing the sequence  $\pi/2 - \tau - \pi/2 - t_1 - \pi - t_2 - \pi/2 - \tau - \text{echo}$  with  $\tau = 136$  ns for the resting state and  $\tau = 200$  ns for the substrate-bound spectra at 20 K, collecting 256 data points in both dimensions. The  $^{14}\text{N}$  HYSCORE spectra were recorded near parallel (3090 G) and near perpendicular (3390 G) directions employing the same sequence reported above with  $\tau = 136$  ns and  $\tau = 200$  ns at 5 K or 20 K, collecting 256 data points in both dimensions. The relaxation decay was subtracted by baseline corrections (fitting by polynomials of 3-4 degrees) in both time domains, subsequently applying apodization (Hamming window) and zero-filling to 1,024 data points in both dimensions. After 2D Fourier transformation, the spectra were simulated using EasySpin (19).

The Davies ENDOR spectra were obtained using the sequence  $\pi - T - \pi/2 - \tau - \pi - \tau - \text{echo}$  with mw pulses of length  $t_{\pi/2} = 32/128$  ns and  $t_{\pi} = 64/256$  ns. During time T a radio frequency (rf) pulse of length  $t_{\text{rf}} = 12$   $\mu\text{s}$  was generated by the Bruker DICE system and amplified by a 60 dB gain ENI A-500 W amplifier.

Raw EPR data are available on request through the Research Data York (DOI: 10.15124/969dd5ce-c1fa-47f4-ba56-e0b026050ed0).

## Computational Details

### *Geometry optimization*

Atomic coordinates of *BaAA10* were obtained from the crystal structure (PDB:5IJU, resolution 1.7 Å) from the Carbohydrate-Active enZymes (CAZy) database) (20). Included in the truncated models were the central Cu(II) ion and 9 residues (His28, Glu68, Gln92, Ala123, Pro124, His125,

Thr127, Trp187 and Phe196 (numbering starting at the first histidine is position 28; His 28 and His 125 are analogous to His32 and His121 in *B/LPMO10A*). Hydrogens were added to appropriate positions and the following modifications were made to decrease the computational cost of the calculations: His28 and His125 were truncated at the carbonyl carbon, which was replaced by a methyl group, Glu68 and Gln92 were truncated with methyl substitution of the C $\gamma$ , the nitrogen of the amide bond between Ala123 and Thr122 was replaced by methyl groups, Trp187 and Phe196 were truncated with methyl substitution of the C $\beta$ , Thr127 was truncated with methyl substitution of the C $\alpha$  and the methyl group of C $\beta$  was removed. For the ‘resting state’ model, three water molecules were retained from the crystallographic coordinates, including the two coordinating water molecules, and the nearby ‘distal’ water molecule (O498, O454 and O526 respectively). To mimic substrate binding, a second model was made, where the water molecules from the crystal structure were replaced by one single equatorially coordinating water molecule. In order to account for the structural constraints imposed by the protein, multiple atoms were kept frozen throughout the optimization; these atoms are denoted by asterisks in Fig. S7.

Geometry optimizations were performed using the Gaussian 09 software package. Optimizations of both models were carried out using the generalized gradient approximation (GGA) functional uBP86. Ahlrichs’s Def-2-TZVP basis set was used to treat the copper, the first coordination sphere nitrogen atoms and the oxygen atoms of the copper-ligating water molecules. On all remaining atoms a Def2-SVP basis set was used. Solvation effects were accounted for using the polarizable continuum model with a dielectric constant of 80.0, as implemented by Gaussian 09. Empirical dispersion corrections were accounted for using Grimme’s empirical dispersion (GD3). The resulting coordinates from the geometry optimizations of the ‘resting state’ (A) and the ‘4-coordinate’ (B) models are shown by the pink wires in Fig S8 and compared to the crystallographic coordinates (blue sticks).

An additional model was constructed (based on model B) to include a short chitin oligosaccharide (NAG<sub>2</sub>) bound to the enzyme. The substrate molecule was positioned in accord with the findings from Bizarro *et al.* (21), with particular focus on the distance of the remote nitrogen on His125 to the carbonyl of an *N*-acetyl group. Only two sugar units were included in this calculation to minimize computational expense. The substrate molecule was held in place by keeping the methyl



carbon atoms of the N-acetyl groups and the terminal oxygen atoms (in the 1 and 4 positions) frozen. The resulting structure is shown in Fig. S9.

### ***EPR Property Calculations***

EPR property calculations were performed on the optimized geometries using the ORCA 4.1.0 program at the DFT level of theory. The cartesian reference system was oriented as such that the NH<sub>2</sub>-Cu(II)-O axis was aligned with the y-axis and the N-Cu(II)-N axis was oriented along x. The integration grid size was kept large (AngularGrid = 7 for all atoms and IntAcc = 7 for the Cu(II) ion) to ensure that the core density was correctly described. Solvation effects were accounted for in the property calculations by implementing the conductor-like polarized continuum model (CPCM) with a dielectric constant of 80.0 and a refractive index of 1.33 (water). The hyperfine coupling calculations included the Fermi-contact, spin dipolar and spin orbit contributions.

A variety of basis set combinations and functionals were employed to ensure extracted trends were not functional specific or basis set limited (Table S3). The fraction of Hartree-Fock exchange was also altered in the hybrid functional, following several studies showing improvements in the EPR property calculations (22). Calculations were performed using non-relativistic and scalar relativistic (zeroth order regular approximation, ZORA) approximations.

The overall accuracy of the spin-Hamiltonian parameters determined by DFT showed to be limited (Tables S4-S6). However, the trends predicted between Models A and B showed to be consistent. All the calculations show a change in both *g* and Cu(II) hyperfine tensors from rhombic to axial, going from the ‘resting state’ to the ‘4-coordinate’ model. The best agreement of the experimentally derived spin-Hamiltonian parameters was achieved using scheme 2 (Table S3). This scheme employed the hybrid functional, B3LYP, with an adjusted fraction of Hartree-Fock exchange (38%). This scheme also utilises the IGLO-III basis set which has additional flexibility in the core region, making it more suited for EPR properties calculations with respect to the Def2-TZVP basis set.

### ***Superoxide model geometry optimizations***

DFT calculations were performed to compare the 5-coordinate and the 4-coordinate Cu-superoxide binding energy. Cu-superoxide cluster models were constructed based off models (A) and (B) to compare the Cu-superoxide bond strength in the presence and absence of substrate. A five-coordinate superoxide model (D) was constructed, based off the resting state 5-coordinate model (A), replacing a water molecule with an end-on superoxide (retaining the water molecule H-bonded to the glutamate). The 4-coordinate superoxide model (E) was based off the 4-coordinate ‘substrate-bound’ model (B), replacing the water molecule with an end-on superoxide. Cluster models were also constructed for the resulting geometries following superoxide release. For the 5-coordinate superoxide model (D), this is simply the 4-coordinate water model (B) and superoxide. For the 4-coordinate superoxide model (E), an additional 3-coordinate Cu<sup>II</sup> model (F) was generated to represent the geometry following superoxide release. A single superoxide molecule, model (G), was also produced as the remaining product following superoxide dissociation. All other amino acid residues were retained in keeping with the previous models. Geometry optimizations for models (D) and (E) were performed on the triplet potential energy surface. The individual models following superoxide dissociation (B), (F) and (G) were all optimized on the doublet potential energy surface. Both oxygen atoms on the superoxide molecules were treated using Ahlrich’s Def2-TZVP basis set. All geometry optimizations were performed using the same basis set and functional schemes as outlined in the *Geometry optimization* section.

The atomic coordinates of all models used in the calculations are provided in the Appendix.

### ***Superoxide bond strength calculations***

To evaluate the superoxide binding energy, single point calculations were performed (as implemented by ORCA 4.2.0) on the optimized geometries: models (B), (D), (E), (F) and (G). These were completed using the hybrid B3LYP functional and the Def2-TZVP basis set across all atoms for improved accuracy. The RIJCOSX approximation and def2/J auxiliary basis set were implemented to help reduce the computation expense of these calculations. Grimme’s dispersion correction with Becke-Johnson damping (D3BJ) was included, together with the conductor-like polarizable continuum model (CPCM) was implemented with a dielectric constant of 80.4 and refractive index of 1.33 to account for solvation effects in water. The relative binding energy of the superoxide in models (D) and (E) was examined by subtracting the electronic energies of

superoxide (G) and the resulting geometry (models (B) and (F), respectively). The results are collected in **table S10**. The strength of the Cu(II)-superoxide bond is shown to be 8.2 kcal mol<sup>-1</sup> stronger in the 4-coordinate superoxide model (E) when compared to the 5-coordinate superoxide model (D). This is a 32% increase in binding energy when the superoxide is oriented in plane with the d(x<sup>2</sup>-y<sup>2</sup>) SOMO.

In addition, differences in the Cu(II)-superoxide bonding are evident when assessing the Löwdin spin population analyses of the two superoxide models (**Table S9**). These values were obtained from the same single point frequency calculation described above. The spin population on the copper ion is shown to decrease by 12.5% upon changing from 5 to 4 coordinate (mimicking substrate binding). This decrease in spin population on the metal is in accord with an analogous increase of the spin population located on the two oxygen atoms. Therefore, this analysis shows that the decrease in coordination number from 5 to 4 leads to a large increase in covalency of the Cu(II)-superoxide bond. A reduction in spin population of d(z<sup>2</sup>) character (ca. 1.3%) is also seen following a reduction in coordination number.

## Supplementary Discussion – Calculations used in EPR analysis

The Cu d based molecular orbitals are written as:

$$\psi_{x^2-y^2} = \alpha_{GS}(a d_{x^2-y^2} - b d_{z^2}) - \sqrt{1 - \alpha_{GS}^2} \psi_L$$

$$\psi_{z^2} = \alpha_{GS}(b d_{x^2-y^2} + a d_{z^2}) - \sqrt{1 - \alpha_{GS}^2} \psi_L$$

$$\psi_{xy} = \alpha_{xy} d_{xy} - \sqrt{1 - \alpha_{xy}^2} \psi_L$$

$$\psi_{xz} = \alpha_{xz} d_{xz} - \sqrt{1 - \alpha_{xz}^2} \psi_L$$

$$\psi_{yz} = \alpha_{yz} d_{yz} - \sqrt{1 - \alpha_{yz}^2} \psi_L$$

The  $\psi_{x^2-y^2}$  orbital is the SOMO; the  $\alpha_i$  represents the metal d-orbital contribution to the molecular orbital, while  $a$  and  $b$  are the coefficient for the  $d_{x^2-y^2}$  and  $d_{z^2}$  orbitals in the ground state orbital (GS), with  $a^2 + b^2 = 1$ .

The  $g$  values can be expressed as:

$$\Delta g_z \approx \frac{8\zeta_{Cu} \alpha_{GS}^2 \gamma_{xy}^2 a^2}{\Delta E_{xy \rightarrow x^2-y^2}}$$

$$\Delta g_y \approx \frac{2\zeta_{Cu} \alpha_{GS}^2 \alpha_{xz}^2 (a + \sqrt{3}b)^2}{\Delta E_{xz \rightarrow x^2-y^2}}$$

$$\Delta g_x \approx \frac{2\zeta_{Cu} \alpha_{GS}^2 \alpha_{yz}^2 (a - \sqrt{3}b)^2}{\Delta E_{yz \rightarrow x^2-y^2}}$$

$\zeta_{Cu}$  represent the one-electron quasi-atomic copper spin-orbit coupling constant (usually taken as  $-830 \text{ cm}^{-1}$ ), and the  $\Delta E$  values are excitation energies of the ligand field transitions.

Similarly, the Cu hyperfine coupling can be written as:

$$A_z = P_d \left[ -K - \frac{4}{7} \alpha_{GS}^2 (a^2 - b^2) + \Delta g_z + \frac{\Delta g_y (3a - \sqrt{3}b)}{14(a - \sqrt{3}b)} + \frac{\Delta g_x (3a + \sqrt{3}b)}{14(a - \sqrt{3}b)} \right]$$

$$A_y = P_d \left[ -K + \frac{2}{7} \alpha_{GS}^2 (a^2 - b^2) - \frac{4\sqrt{3}}{7} \alpha_{GS}^2 (ab) + \Delta g_y - \frac{\Delta g_x (3a + \sqrt{3}b)}{14(a - \sqrt{3}b)} \right]$$

$$A_x = P_d \left[ -K + \frac{2}{7} \alpha_{GS}^2 (a^2 - b^2) + \frac{4\sqrt{3}}{7} \alpha_{GS}^2 (ab) + \Delta g_x - \frac{\Delta g_y (3a - \sqrt{3}b)}{14(a - \sqrt{3}b)} \right]$$

$P_d = g_e g_{Cu} \mu_e \mu_{Cu}$  is the quasi atomic parameter usually taken as 1180 MHz, the term  $-P_d K$  (in blue) represents the isotropic Fermi contact ( $A^{\text{Fermi}}$ ), indicated in blue; the Spin-Dipolar ( $A^{\text{SD}}$ ) contribution is in green and the Spin-Orbit contribution is indicated in red ( $A^{\text{SO}}$ ).

The different Cu hyperfine contributions, the %  $d_{z^2}$  in the SOMO and the  $\alpha_{GS}^2$  were determined as follows:

- 1) The ratio between  $a/b$  (and therefore the %  $d_{z^2}$  in the ground state) was estimated from the rhombicity parameter ( $R_g$ ) and the  $g$  values equations following Gewirth *et al.*:(23)

$$R_g = \frac{2(\Delta g_2 - \Delta g_1)}{\Delta g_2 + \Delta g_1} \approx 2 \frac{(a + \sqrt{3}b)^2 - (a - \sqrt{3}b)^2}{(a + \sqrt{3}b)^2 + (a - \sqrt{3}b)^2}$$

Assuming  $\alpha_{yz}^2 \approx \alpha_{xz}^2$  and  $\Delta E_{yz \rightarrow x^2 - y^2} \approx \Delta E_{xz \rightarrow x^2 - y^2}$ .

- 2)  $A^{\text{SO}}$  was obtained from the experimental  $\Delta g$  and the  $a$  and  $b$  values obtained in 1)
- 3)  $A^{\text{SO}}$  was subtracted from  $A^{\text{Total}}$  to get  $A^{\text{Fermi}} + A^{\text{SD}}$
- 4) The  $x$ ,  $y$  and  $z$  components of  $A^{\text{Fermi}} + A^{\text{SD}}$  were averaged to obtain  $A^{\text{Fermi}}$
- 5)  $A^{\text{Fermi}} + A^{\text{SO}}$  was subtracted from  $A^{\text{Total}}$  to get  $A^{\text{SD}}$
- 6)  $\alpha_{GS}^2$  was calculated from  $A^{\text{SD}}$

### Supplementary Discussion - DFT calculations in presence of substrate

In searching for the origin of the change in EPR parameters and hydrogen-bonding patterns around the histidine brace upon substrate addition, we undertook a series of DFT calculations in which

the effects of adding a chitin substrate to the Cu(II) active site were evaluated. To this end, three different cluster models of the enzyme Cu site were built, starting from the crystallographic coordinates of a spectroscopically similar AA10, *BaAA10*, PDB:5IJU (20). In these models the protein backbone was truncated to only include important residue side chains near the metal, in order to reduce computational cost while still representing the major interactions imposed by the protein on the Cu(II) ion. All geometries were optimized with the uBP86 functional (see Methods for details). Model A represents the Cu(II) resting state of the enzyme, where the metal first coordination sphere consisted of the His-brace and two water molecules in a distorted square-pyramidal geometric structure (Figure S7-A). The optimized geometry showed minimal distortion from the crystallographic coordinates (Figure S8), showing that the model faithfully represented the Cu(II) active site.

The spin-Hamiltonian parameters obtained from the EPR of *B/LPMO10A* after addition of substrate suggested that the copper coordination sphere is near axial and, in accordance with previous studies, it is reasonable to think that water molecule in the nominal axial position is displaced by binding of substrate. Hence, a second model (Model B) was produced with a square planar-like symmetry about the copper center, achieved by the removal of the pseudo-axial water molecule from the optimized geometry structure of Model A (Figure S7-B and S8-B). Additionally, a third model was constructed based on Model B, but also including a short chitin oligosaccharide ( $\text{NAG}_2$ ) bound to the active site (Figure S9). The positioning of this substrate molecule was informed from a recent combined molecular mechanics/quantum mechanics (QM/MM) study by Bissaro *et al.* (21). Again, the calculated optimized geometry for these models showed good agreement with the crystallographic coordinates of the *BaAA10* resting state.

The optimized geometries were then used to calculate the EPR spin-Hamiltonian parameters for the models at the DFT level of theory, testing various functionals and basis sets to monitor the reproducibility of the trends (Tables S3-S6). Notwithstanding some deviation from experimental values, the results reproduced the experimental trends (Table S7) upon substrate addition. Most notably, the ‘resting state’ Model A predicts the rhombic  $g$  and  $A$  values seen experimentally, including the large  $A_1$  value, that is a characteristic feature of chitin active AA10s. As expected, Model B afforded an almost axial set of spin-Hamiltonian parameters, commensurate with the values of *B/LPMO10A* upon addition of substrate. These data confirm the conclusions from the

LFT analysis above that most changes to the spectral envelope of *B/LPMO10A* upon substrate binding are associated with a change in coordination number from 5 to 4, generating an almost square-planar coordination geometry for the Cu(II) ion.

The change of coordination number from 5 to 4 (between models A and B) emulates the changes seen in copper hyperfine values upon substrate binding. However, the decrease of  $a_{iso}$  observed in the HYSORE spectra, assigned to the remote nitrogen of His121, is not reproduced in the differences between Models A and B. In Model C, the short substrate molecule introduces a hydrogen bond between the substrate and the N $\delta$ -H group of His121. The calculated spin-Hamiltonian parameters for this model show (Table S8) a decrease in the  $a_{iso}$  value from 1.92 to 1.48 MHz for the remote N of His121, with a simultaneous increase of the value of the electric field gradient asymmetry parameter,  $\eta$ , from 0.2 to 0.9, wholly in accord with the experimental data. In other words, the formation of a hydrogen bond between the N-H group of His121 and the substrate is required to fully reproduce the experimental trend.

**Table S1.** Input data for calculation of the structure of *apo-B/LPMO10A* and structural assessment statistics for *apo-B/LPMO10A* and Cu(I)-bound *B/LPMO10A*.

	<i>apo-B/LPMO10A</i>	Cu(I)- <i>B/LPMO10A</i>
Total number of NOE distance constraints	1,623	1,209
Intraresidue	724	440
Sequential	496	421
Medium-range	101	98
Long-range	302	250
Torsion angle restraints <sup>a</sup>	264	264
Structure statistics (20 conformers) <sup>b</sup>		
CYANA target function value (Å <sup>2</sup> )	3.69 ± 0.47	-
Maximum residual distance constraint violation (Å)	0.43 ± 0.15	-
Maximum torsion angle constraint violation (°)	4.20 ± 2.10	-
Ramachandran plot analysis <sup>c</sup>		
Residues in favored regions (%)	89.5	83.9
Residues in additionally allowed regions (%)	8.7	14.5
Residues in generously allowed regions (%)	0.6	0.7
Residues in forbidden regions (%)	1.2	0.9
rmsd to the lowest target energy conformer (Å)		
N, C <sup>α</sup> , C'	2.41 ± 0.36	3.96 ± 1.15
Heavy atoms	2.71 ± 0.33	4.34 ± 1.15
N, C <sup>α</sup> , C' (Secondary structure) <sup>d</sup>	1.48 ± 0.23	2.52 ± 0.71
Heavy atoms (Secondary structure) <sup>d</sup>	1.82 ± 0.25	2.97 ± 0.79

<sup>a</sup> Calculated from secondary chemical shifts using the TALOS-N software (8).

<sup>b</sup> The values are the average and standard deviation over the 20 lowest CYANA target function values before energy minimization.

<sup>c</sup> Calculated using PROCHECK-NMR (24) and the PSVS server (25).

<sup>d</sup> Residues 33-36, 37-40, 41-47, 57-61, 82-84, 89-94, 103-107, 110-117, 125-132, 150-154, 159-161, 164-168, 175-184, 185-190 and 191-201.



**Table S2.** Davies  $^{14}\text{N}$  ENDOR simulation parameters for *B/LPMO10A*. The spectra and simulations are shown in Figure S5. The numbers in brackets represent the error on the measurement estimated from the quality of simulated fits. The Euler angles define the  $zy'z''$  rotations with respect to the  $g$  matrix.

$^{63}\text{Cu-B/LPMO10A}$		
	$A$ (MHz)	A Frame Euler angles
N(Im)	33 33 40; 33 33 40 ( $\pm 2$ )	[0 90 0]
NH <sub>2</sub>	18 18 23 ( $\pm 1$ )	[0 90 90]

**Table S3.** The functional and basis set schemes used to calculate the EPR properties of *B/LPMO10A*.

Scheme	Functional	Cu(II)	Ligands <sup>a</sup>	Remaining atoms
1	B3LYP	CP(PPP)	IGLO III	Def2-SVP
2	B3LYP 38%HFX <sup>b</sup>	CP(PPP)	IGLO III	Def2-SVP
3	B3LYP 38%HFX <sup>b</sup> ZORA	CP(PPP)	ZORA-Def2-TZVP	ZORA-Def2-SVP
4	PBE0 ZORA	CP(PPP)	ZORA-Def2-TZVP	ZORA-Def2-SVP

*a* – ‘Ligands’ refers to the amino terminus nitrogen, all histidine ring atoms and the coordinating oxygen atoms.

*b* – ‘38%HFX’ refers to an adjusted degree of Hartree-Fock exchange to 38%, from the default 20% in the B3LYP functional.

**Table S4.** Calculated spin Hamiltonian parameters for the ‘resting state’ Model A of *B/LPMO10A* under different functional and basis set schemes.

Scheme <sup>a</sup>	$g$ -values			Cu(II) Hyperfine (MHz)			Principal N Super-Hyperfine (MHz) <sup>b</sup>		
	$g_1$	$g_2$	$g_3$	$A_1$	$A_2$	$A_3$	N1 <sup>c</sup>	N2 <sup>c</sup>	N3 <sup>c</sup>
1	2.028	2.076	2.161	284	65.9	-374	56.9	44.0	44.0
2	2.036	2.104	2.220	309	50.7	-411	48.0	38.4	38.4
3	2.036	2.104	2.221	291	33.9	-432	46.6	35.0	35.7
4	2.032	2.086	2.183	251	21.8	-429	51.8	38.6	39.2

*a* – The scheme refers to the basis set and functional used in the calculation, as outlined in Table S3.

*b* – ‘Principal N Superhyperfine’ values refer to the largest individual value calculated in the tensor. The principal SHF value is  $a_{yy}$  N1 and  $a_{xx}$  for N2 and N3.

*c* – N1 refers to the amino terminus nitrogen, N2 refers to the coordinating ring nitrogen of His1 and N3 refers to the coordinating ring nitrogen of His121.

**Table S5.** Calculated spin Hamiltonian parameters for the ‘4-coordinate’ Model B model of B/LPMO10A under different functional and basis set schemes.

Scheme <sup>a</sup>	g-values			Cu(II) Hyperfine (MHz)			Principal N Super-Hyperfine (MHz) <sup>b</sup>		
	<i>g</i> <sub>1</sub>	<i>g</i> <sub>2</sub>	<i>g</i> <sub>3</sub>	<i>A</i> <sub>1</sub>	<i>A</i> <sub>2</sub>	<i>A</i> <sub>3</sub>	N1 <sup>c</sup>	N2 <sup>c</sup>	N3 <sup>c</sup>
1	2.041	2.043	2.154	3.8	-4.7	-564	48.5	45.2	47.1
2	2.055	2.060	2.213	22.7	-11.2	-603	42.4	39.8	41.9
3	2.057	2.060	2.214	-24.1	-55.1	-650	41.1	37.7	39.9
4	2.048	2.049	2.176	-53	-74.1	-638	44.4	41.3	43.2

*a* – The scheme refers to the basis set and functional used in the calculation, as outlined in Table S3.

*b* – ‘Principal N Superhyperfine’ values refer to the largest individual value calculated in the tensor. The principal SHF value is *a*<sub>yy</sub> N1 and *a*<sub>xx</sub> for N2 and N3.

*c* – N1 refers to the amino terminus nitrogen, N2 refers to the coordinating ring nitrogen of His1 and N3 refers to the coordinating ring nitrogen of His121.

**Table S6.** Calculated spin Hamiltonian parameters for ‘substrate bound’ Model C of B/LPMO10A.

Scheme <sup>a</sup>	g-values			Cu(II) Hyperfine (MHz)			Principal N Super-Hyperfine (MHz) <sup>b</sup>		
	<i>g</i> <sub>1</sub>	<i>g</i> <sub>2</sub>	<i>g</i> <sub>3</sub>	<i>A</i> <sub>1</sub>	<i>A</i> <sub>2</sub>	<i>A</i> <sub>3</sub>	N1 <sup>c</sup>	N2 <sup>c</sup>	N3 <sup>c</sup>
2 <sup>d</sup>	2.054	2.057	2.203	-8.97	-28.9	-638	43.1	42.7	45.4

*a* – The scheme refers to the basis set and functional used in the calculation, as outlined in Table S3.

*b* – ‘Principal N Superhyperfine’ values refer to the largest individual value calculated in the tensor. The principal SHF value is *a*<sub>yy</sub> N1 and *a*<sub>xx</sub> for N2 and N3.

*c* – N1 refers to the amino terminus nitrogen, N2 refers to the coordinating ring nitrogen of His1 and N3 refers to the coordinating ring nitrogen of His121.

*d* – Only scheme 2 was tested, due to size limitations of the model.

**Table S7.** EPR properties of *B/LPMO10A* with and without  $\beta$ -chitin, determined experimentally and by DFT calculations (using B3LYP with 38% Hartree-Fock exchange). The signs  $-/+$  refer to absence or presence of substrate, respectively.

	$g_1$	$g_2$	$g_3$	$A_1$ (MHz)	$A_2$ (MHz)	$A_3$ (MHz)
<b>EPR</b> <sup>-a</sup>	2.027	2.095	2.261	255	110	336 <sup>a</sup>
<b>DFT</b> <sup>-b</sup>	2.036	2.104	2.220	309	50.7	-411
<b>EPR</b> <sup>+a</sup>	2.042	2.053	2.205	80	85	620 <sup>a</sup>
<b>DFT</b> <sup>+c</sup>	2.055	2.060	2.213	22.7	-11.2	-603
<b>DFT</b> <sup>+d</sup>	2.054	2.057	2.203	-8.97	-28.9	-638

*a* – Signs of Cu hyperfine could not be determined from simulation of the CW EPR spectra.

*b* – Refers to the DFT ‘resting state’ Model A.

*c* – Refers to the DFT ‘4-coordinate’ Model B.

*d* – Refers to the DFT ‘substrate bound’ Model C.

**Table S8.** EPR properties of the remote nitrogen atoms of the histidine rings of *B/LPMO10A* with and without  $\beta$ -chitin, determined experimentally and by DFT calculations. The signs  $-/+$  refer to absence or presence of substrate, respectively.

<b>System</b>	$a_{iso}$ His32 (MHz)	T His32 (MHz)	$\eta$ His32	$a_{iso}$ His121 (MHz)	T His121 (MHz)	$\eta$ His121
<b>EPR</b> <sup>-</sup>	1.3	0.25	0.85	1.6	0.3	0.7
<b>DFT</b> <sup>-a</sup>	1.21	0.24	0.5	1.75	0.21	0.2
<b>EPR</b> <sup>+</sup>	1.25	0.30	0.8	1.55	0.35	0.9
<b>DFT</b> <sup>+b</sup>	1.13	0.21	0.5	1.92	0.21	0.2
<b>DFT</b> <sup>+c</sup>	1.38	0.31	0.7	1.48	0.23	0.9

*a* – Refers to the DFT ‘resting state’ Model A.

*b* – Refers to the DFT ‘4-coordinate’ Model B.

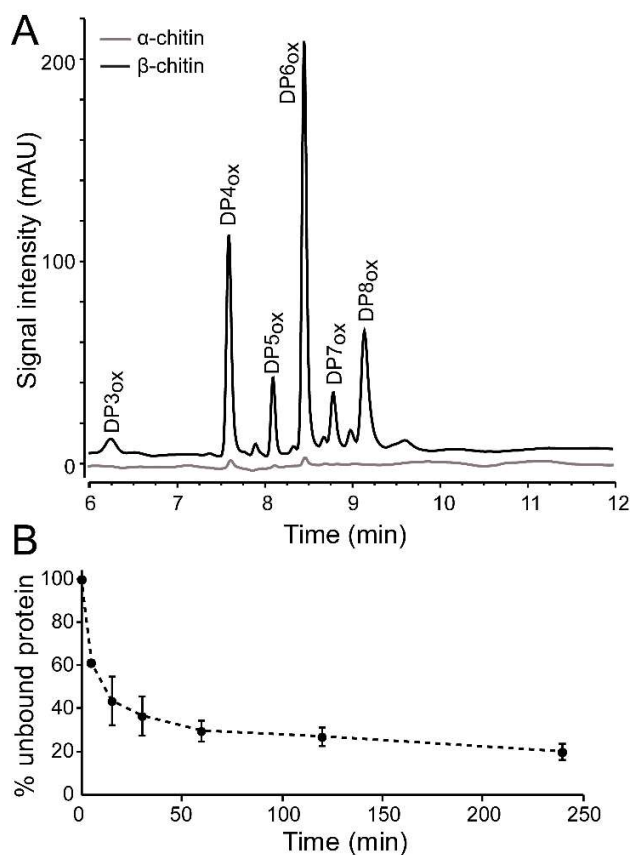
*c* – Refers to the DFT ‘substrate bound’ Model C.

**Table S9.** Löwdin spin population analysis of 4-coordinate and 5-coordinate superoxide models (adding up to 200% for a total of 2 spins).

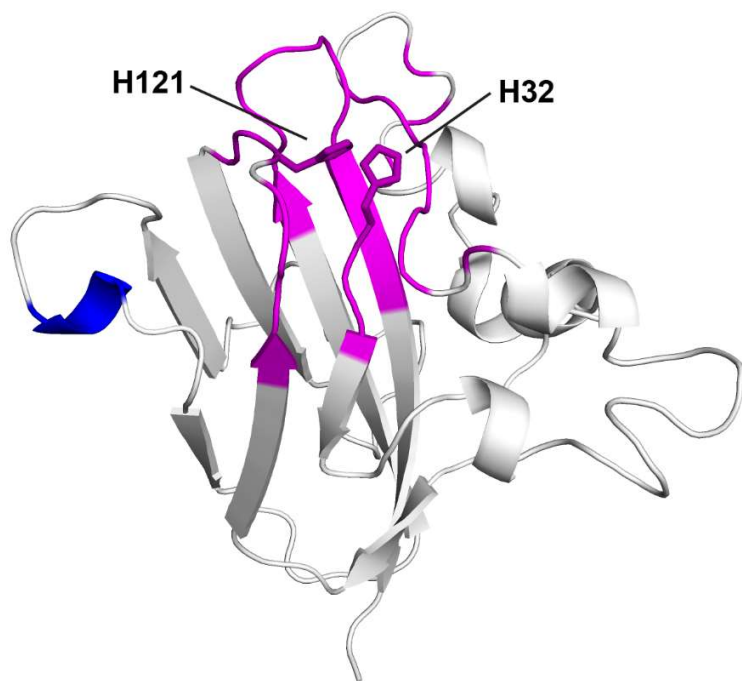
Model	N (NH <sub>2</sub> )	N (His1)	N (His2)	Cu <sup>a</sup>	O (O <sub>2</sub> proximal)	O (O <sub>2</sub> distal)	O (Water)	O-O distance / Å
<b>4 - coordinate (E)</b>	5.5	3.8	3.3	Total: <b>41.0</b> s = -0.9 p = -1.3 d = 43.2 (z <sub>2</sub> = 0.2) (xz = 0.2) (yz = 0.8) (xy = 0.7) (x <sub>2</sub> -y <sub>2</sub> = 41.2)	71.6	72.5	-	1.28
<b>5 - coordinate (D)</b>	6.8	5.4	4.8	Total: <b>53.5</b> s = -0.5 p = -1.4 d = 55.4 (z <sub>2</sub> = 1.5) (xz = 0.6) (yz = 1.3) (xy = 0.3) (x <sub>2</sub> -y <sub>2</sub> = 51.7)	64.9	60.9	0.5	1.30

**Table S10.** Calculated energies of superoxide cluster models (D) and (E) and the resulting geometries following superoxide dissociation (B), (F) and (G). Cu(II)-superoxide binding energy calculated by the difference in electronic energy.

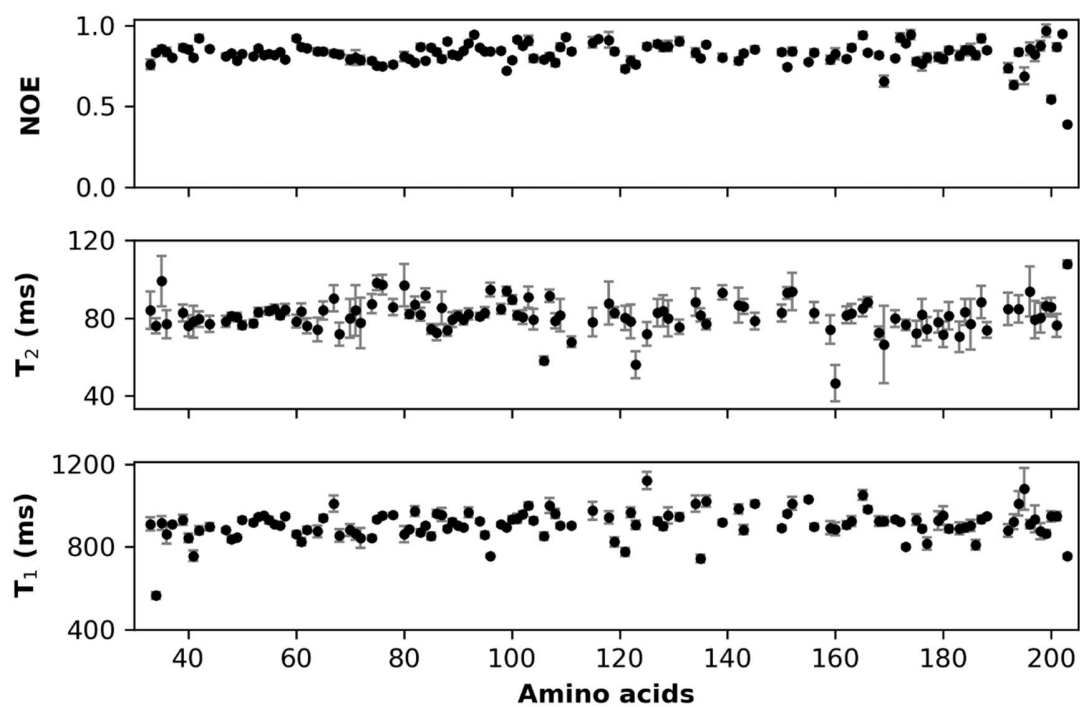
Model	5-coordinate superoxide (D)	4-coordinate superoxide (E)	4-coordinate water (B)	3-coordinate Hisbrace (F)	Superoxide (G)
Single point energy / Hartree	-4489.6	-4413.2	-4339.1	-4262.7	-150.5
Cu(II)-O <sub>2</sub> binding energy / Hartree	-0.0417	-0.0549	-	-	-
Cu(II)-O <sub>2</sub> binding energy / kcal mol <sup>-1</sup>	-26.2	-34.4	-	-	-



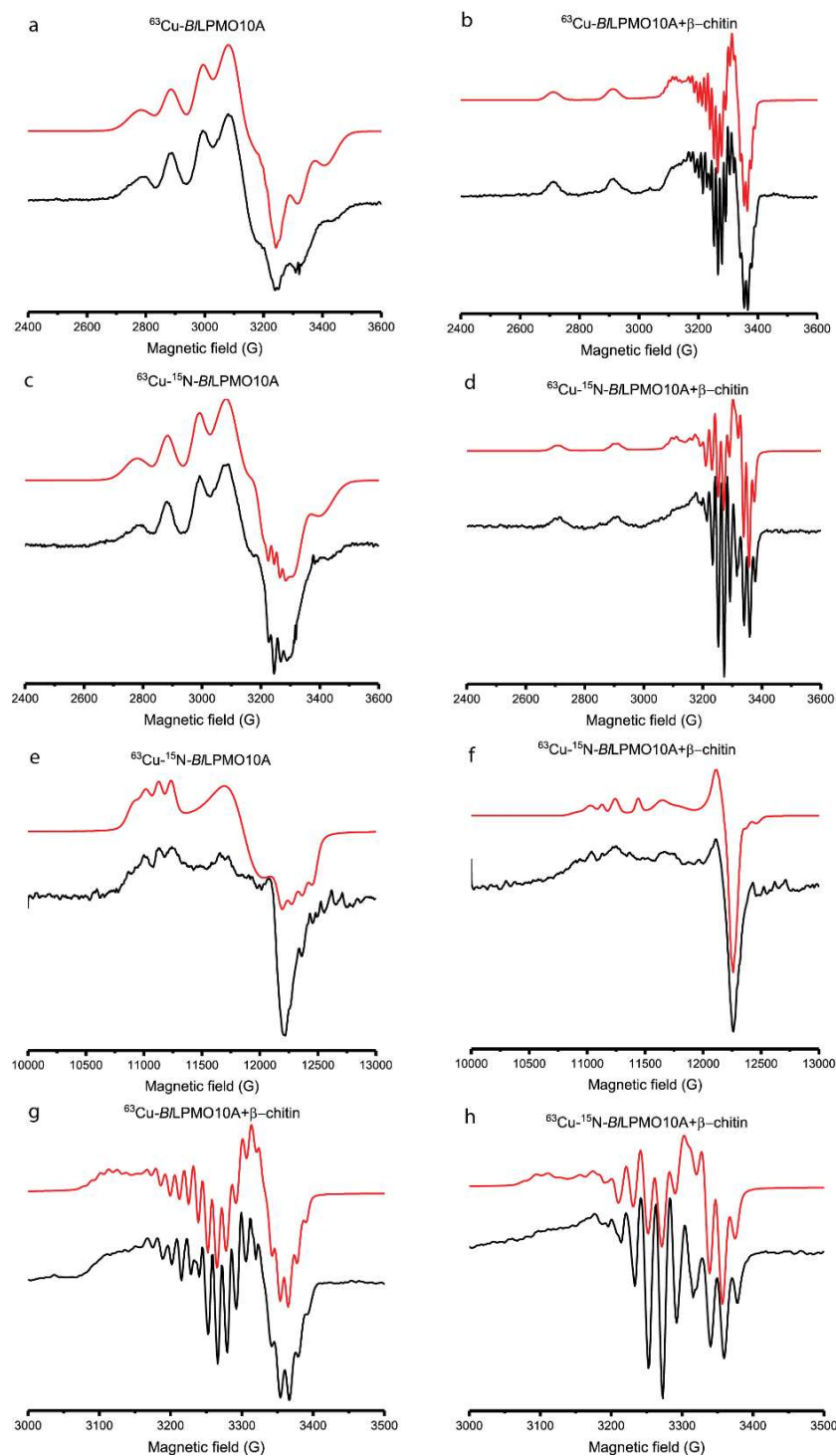
**Figure S1. *B/LPMO10A* activity and binding towards chitin.** A) Chromatographic analysis of C1-oxidized chito-oligosaccharides obtained from degradation reactions containing 10 mg/mL  $\alpha$ -chitin (grey) or  $\beta$ -chitin (black), 1  $\mu$ M Cu(II)-loaded *B/LPMO10A*, 50 mM Tris/HCl pH 8.0 and 2 mM ascorbic acid. Reactions were incubated for 24 h in an Eppendorf Thermomixer set to 40 °C and 800 rpm. B) Binding of *B/LPMO10A* to 10 g/L  $\beta$ -chitin. The percentage of free protein was determined by measuring the reduction in concentration of soluble protein over time. The binding experiment was carried out at 40 °C using in 50 mM sodium phosphate buffer pH 7.0. The error bars show  $\pm$  S.D. (n = 3).



**Figure S2. Effect of Cu(II)-binding on *apo-BILPMO10A*.** Residues with less than 30% remaining intensity (see Figure 1E) and within a 12 Å radius from the expected Cu(II) coordination site are colored pink, whereas residues with less than 30% remaining intensity and further than 12 Å from the Cu(II) site (Ala160-Arg162) are colored blue. The side-chains of His32 and His121 are shown as sticks.



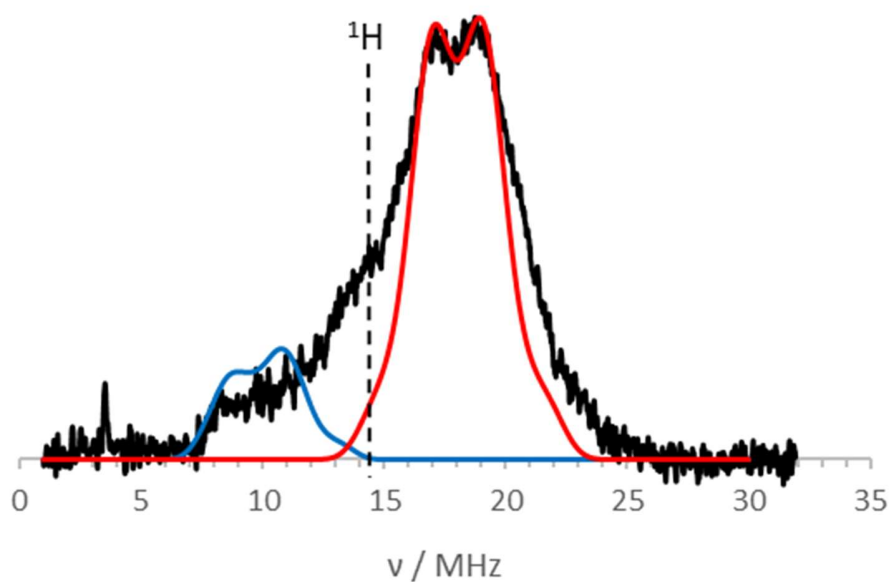
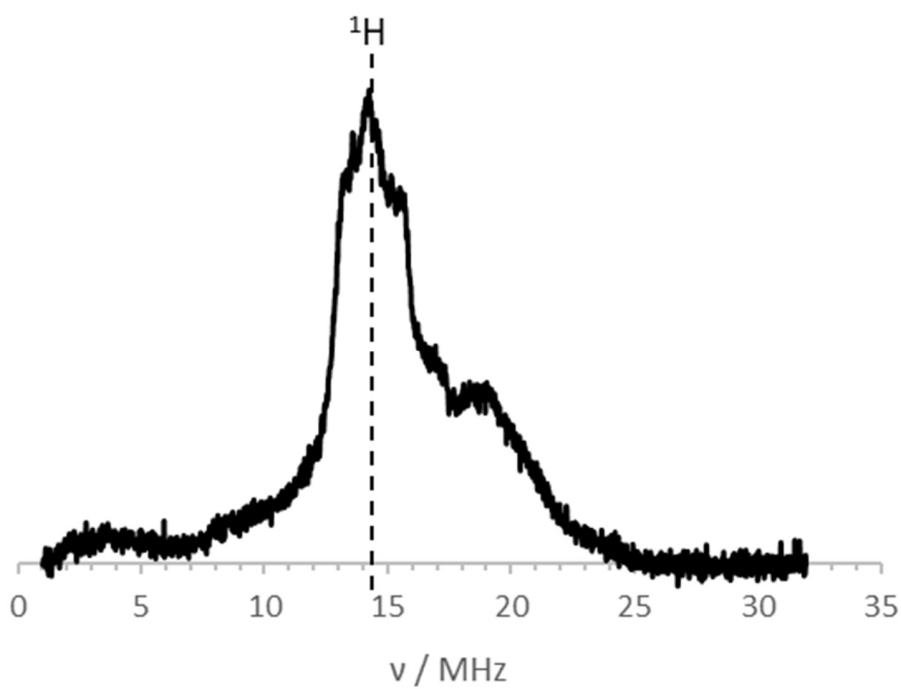
**Figure S3. Backbone dynamics of *apo-BILPMO10A*.** Decreased  $\{^1\text{H}\}\text{-}^{15}\text{N}$  NOE and increased  $^{15}\text{N}\text{-T}_2$  values are an indication of conformational flexibility. The data show that *apo-BILPMO10A* has an overall rigid backbone, with some flexibility in loops (e.g. between  $\alpha 2$  and  $\beta 2$ ), a flexible N-terminus and a flexible C-terminus. The rotational correlation time, calculated from the average  $T_1/T_2$  ratio, is  $\tau_c = 10.2 \pm 0.9$  ns.



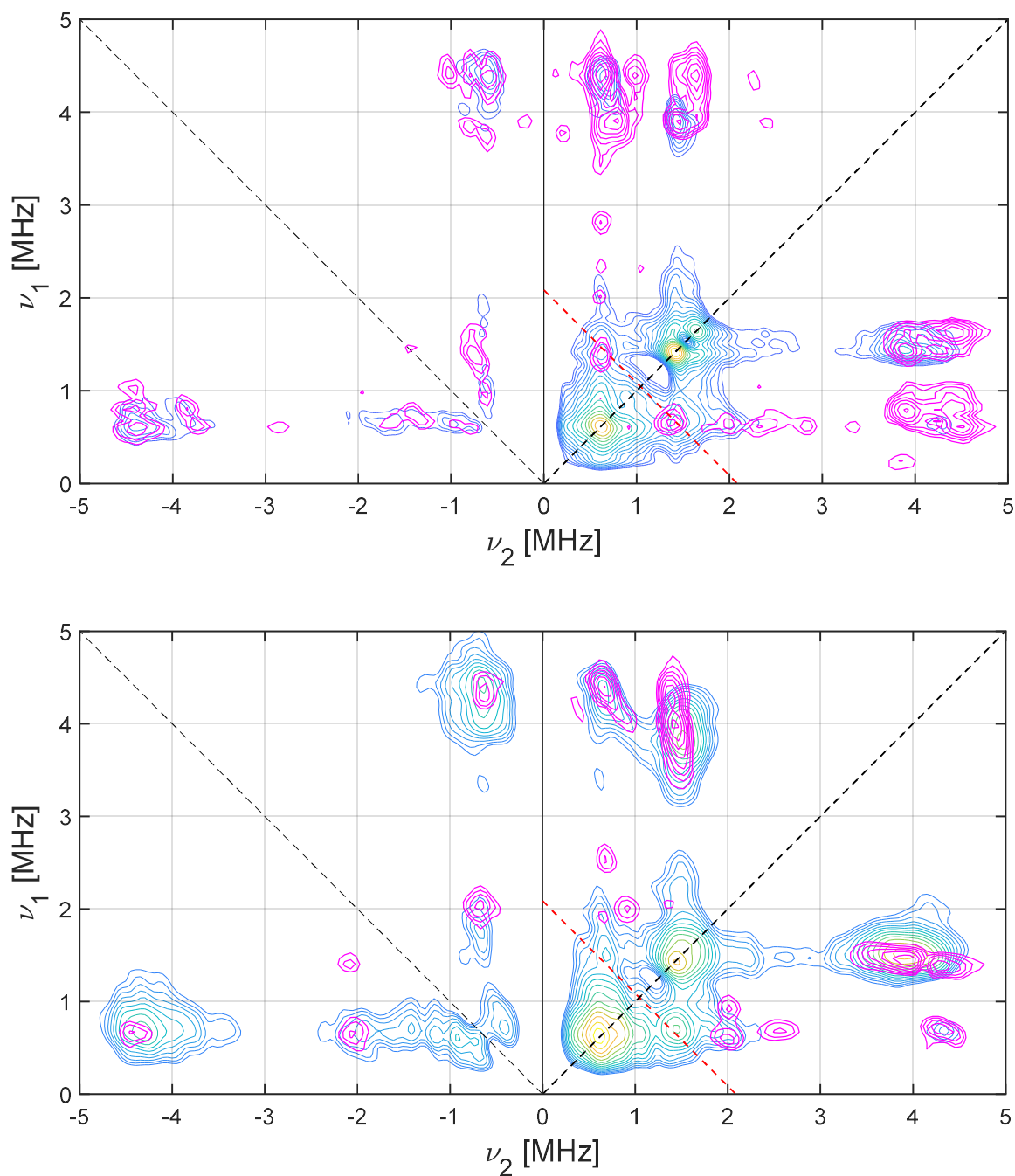
**Figure S4. X and Q band CW-EPR spectra.** The panels show X-band (**a-d, g-h**) and Q-band (**e-f**) CW-EPR spectra (in black, bottom lines) and corresponding simulations (in red, top lines) for  $^{63}\text{Cu}$ -*B/LPMO10A* (**a**),  $^{63}\text{Cu}$ - $^{15}\text{N}$ -*B/LPMO10A* (**c, e**)  $^{63}\text{Cu}$ -*B/LPMO10A* with squid pen  $\beta$ -chitin (**b**) and  $^{63}\text{Cu}$ - $^{15}\text{N}$ -*B/LPMO10A* with squid pen  $\beta$ -chitin (**d, f**). **g** and **h**: detail of the superhyperfine (SHF) coupling with



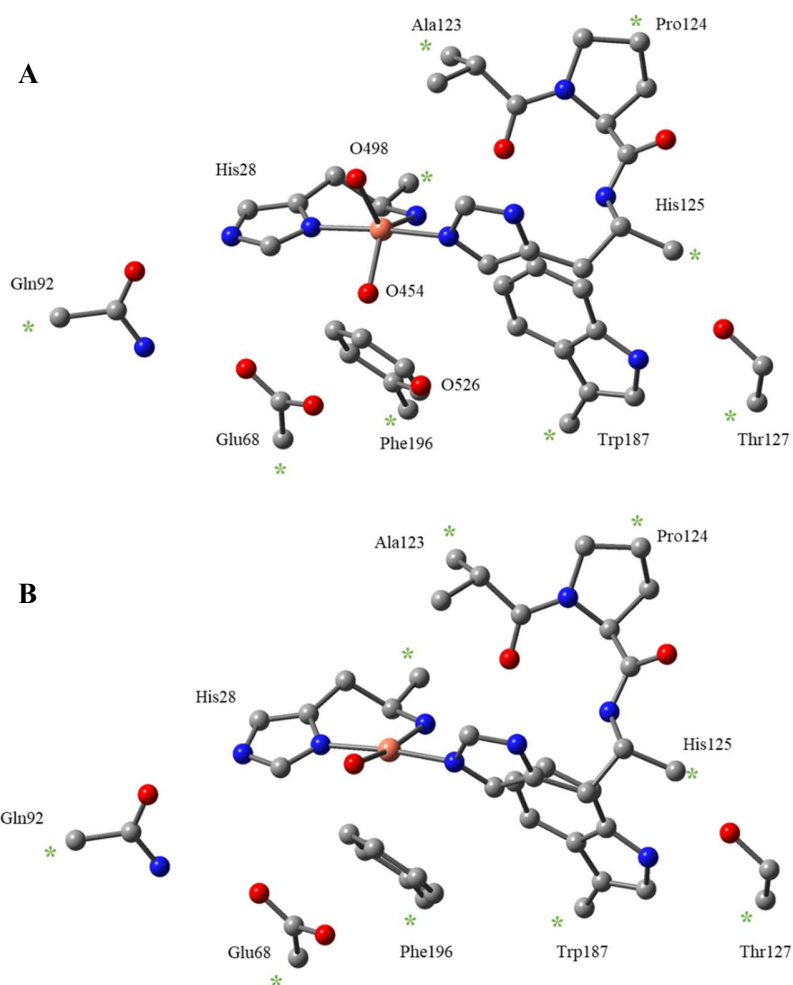
simulations for  $^{63}\text{Cu}$ -*B/LPMO10A* and  $^{63}\text{Cu}$ - $^{15}\text{N}$ -*B/LPMO10A*, respectively, after addition of  $\beta$ -chitin. The spectra shown in panels **b** and **d** were obtained by subtraction of the spectrum of  $^{63}\text{Cu}$ -*B/LPMO10A* and  $^{63}\text{Cu}$ - $^{15}\text{N}$ -*B/LPMO10A*, respectively, to correct for the free enzyme not bound to the substrate.



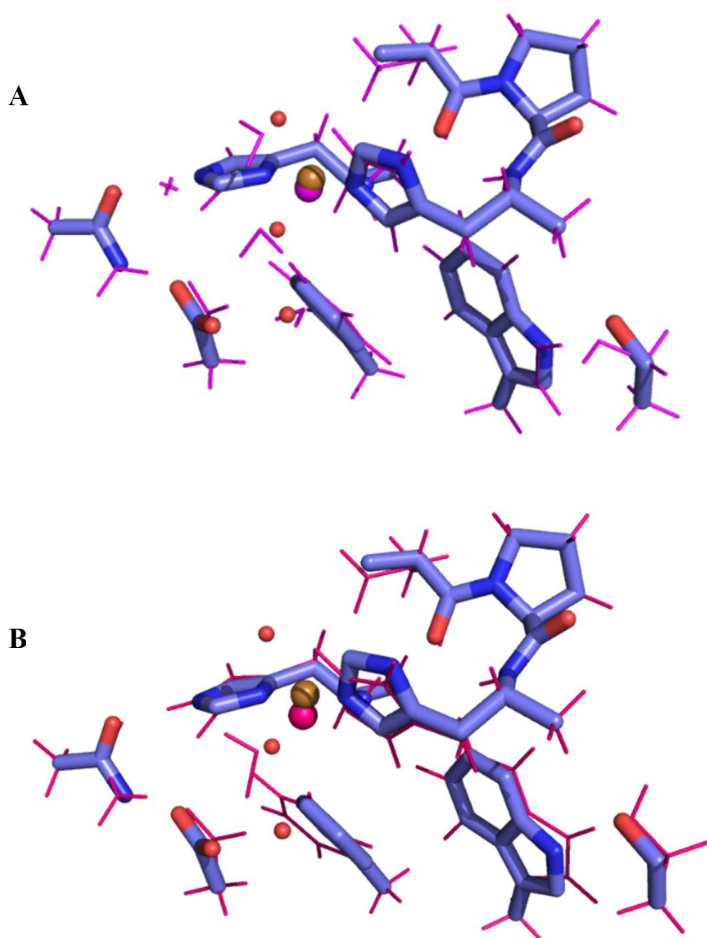
**Figure S5.** Davies  $^{14}\text{N}$  ENDOR spectra of  $^{63}\text{Cu}$ -BLPMO10A recorded at 3400 G near  $g_1$ . **Top:** spectrum recorded with soft pulses,  $\pi/2 = 128$  ns. **Bottom:** spectrum recorded with hard pulses,  $\pi/2 = 32$  ns, with simulations for the coordinated nitrogen atoms in red (His N) and blue (N-terminus). The dashed line indicates the position of the  $^1\text{H}$  Larmor frequency at this field.



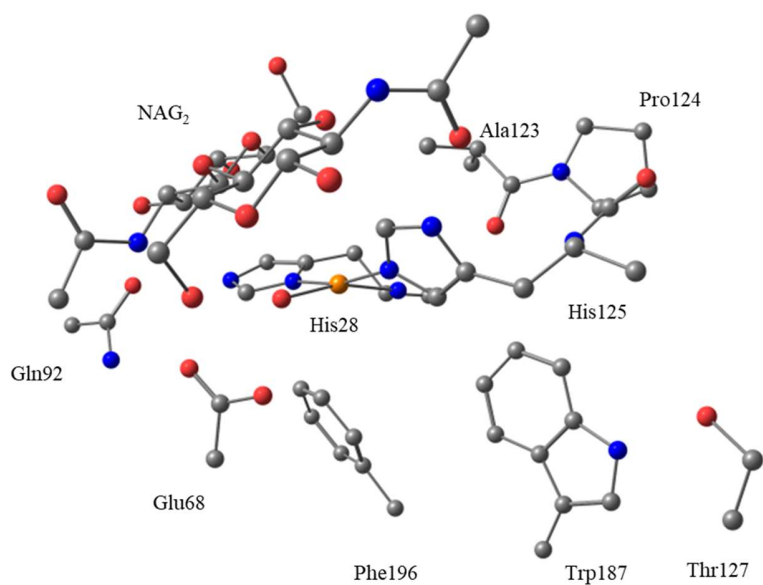
**Figure S6.**  $^{14}\text{N}$  HSCORE spectra and simulations (in pink) of  $^{63}\text{Cu}$ -*B/LPMO10A* (top) and  $^{63}\text{Cu}$ -*B/LPMO10A* with squid pen  $\beta$ -chitin (bottom). **Top**,  $^{63}\text{Cu}$ -*B/LPMO10A* near  $g_{\perp}$  with  $\tau = 200$  ns at 3385 G and 32 ns increments. **Bottom**,  $^{63}\text{Cu}$ -*B/LPMO10A* with squid pen  $\beta$ -chitin near  $g_{\perp}$  with  $\tau = 200$  ns at 3390 G and 32 ns increments.



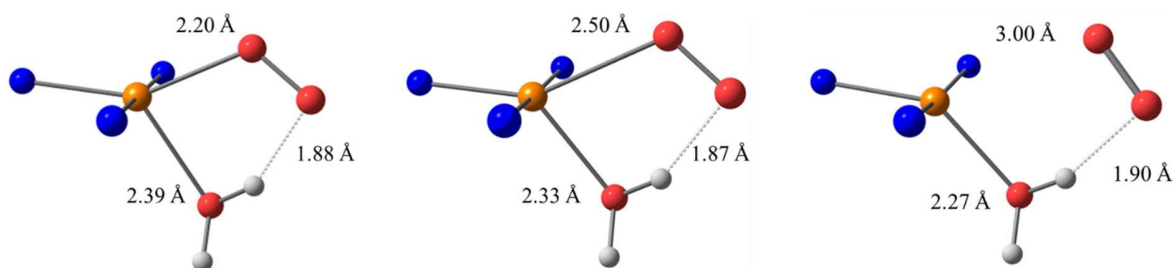
**Figure S7. Starting structure for geometry optimization of Cu(II)-BILPMO10A.** The panels show the structure in the ‘resting state’ (Model A) and in the ‘4-coordinate’ state (Model B). Hydrogen atoms are omitted for clarity. Residue labels are given as found in the crystallographic coordinates of *BaAA10* (PDB:5IJU). Atoms that were kept frozen throughout the geometry optimizations are denoted with a green asterisk.



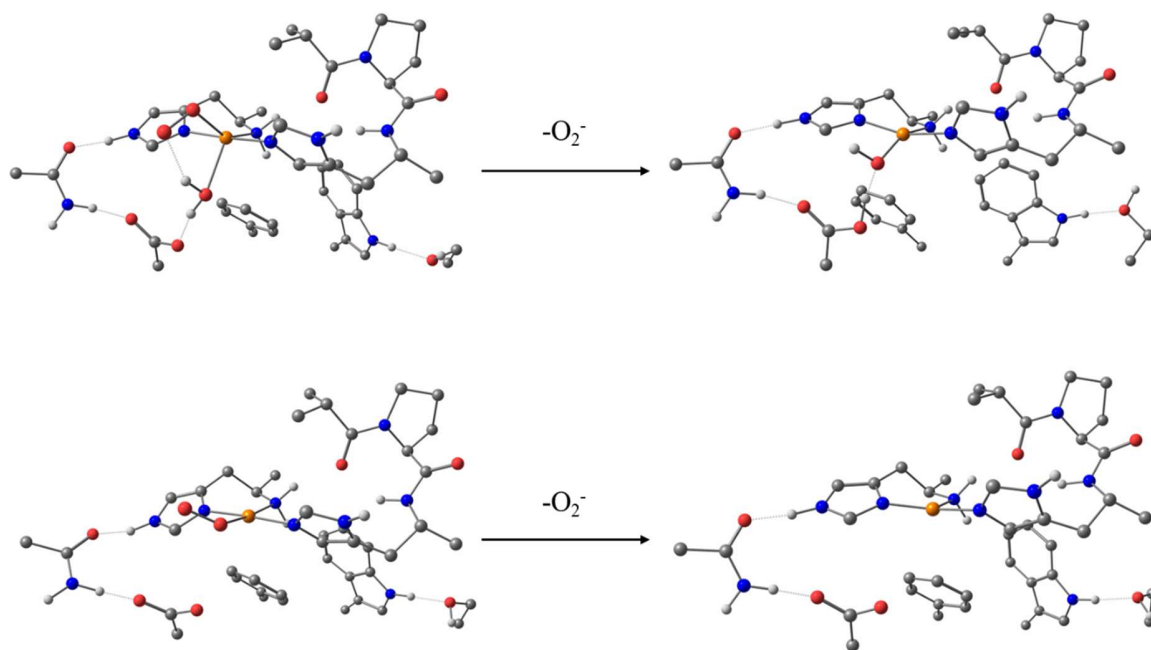
**Figure S8. Geometry optimized structure (pink wires) of Cu(II)-BLPMO10A.** The panels show the structure in the 'resting state' (Model A) and the '4-coordinate' state (Model B), plotted against the crystallographic coordinates (PDB:5IJU) (blue sticks). The water molecules present in the crystal structures are shown as red spheres.



**Figure S9. Geometry optimized structure of Cu(II)-*B/LPMO10A* in the ‘substrate-bound’ case (Model C).** The figure shows the structure with a chitin oligosaccharide bound to the active site, with hydrogen atoms omitted for clarity. Residue labels are given as found in the crystallographic coordinates of *BaAA10* (PDB:5IJU).



**Figure S10. Superoxide release from Cu(II)-AA10. DFT-optimized structures with Cu-O distances fixed at 2.20, 2.50 and 3.00 Å;** at Cu-O = 2.20: O-O = 1.30, at Cu...O = 2.50: O-O = 1.29, at Cu...O = 3.00: O-O = 1.28 Å. The figure shows the DFT-optimized geometries of an AA10 active site containing a superoxide in the coordination sphere. A relaxed surface scan was performed along the Cu-O coordinate between 2.20 and 3.00 Å and a barrier of 3.83 kcal mol<sup>-1</sup> was determined. DFT optimizations were performed using the uBP86 functional and Ahlrichs def2-SVP basis set. The basis set size was increased to Ahlrichs def2-TZVP on the copper ion and all ligating atoms for improved accuracy.



**Figure S11.** Top: DFT optimized geometries of 5-coordinate superoxide model (D) (left) and resulting structure following superoxide removal (model (B)) (right). Bottom: DFT optimized geometries of 4-coordinate superoxide model (E) (left) and resulting structure following superoxide removal (model (F)) (right). Hydrogen atoms attached to carbon atoms were omitted for clarity.



## References

1. Z. Forsberg *et al.*, Comparative study of two chitin-active and two cellulose-active AA10-type lytic polysaccharide monooxygenases. *Biochemistry* **53**, 1647-1656 (2014).
2. J. S. M. Loose; Z. Forsberg; M. W. Fraaije; V. G. H. Eijsink; G. Vaaje-Kolstad, A rapid quantitative activity assay shows that the *Vibrio cholerae* colonization factor GbpA is an active lytic polysaccharide monooxygenase. *FEBS Lett.* **588**, 3435-3440 (2014).
3. G. Courtade *et al.*,  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$  resonance assignment of the chitin-active lytic polysaccharide monooxygenase BILPMO10A from *Bacillus licheniformis*. *Biomol. NMR Assign.* **9**, 207-210 (2015).
4. G. Courtade; S. B. Le; G. I. Sætrom; T. Brautaset; F. L. Aachmann, A novel expression system for lytic polysaccharide monooxygenases. *Carbohydr. Res.* **448**, 212-219 (2017).
5. R. Keller, The computer aided resonance assignment tutorial. CANTINA verlag Goldau: (2004).
6. N. A. Farrow *et al.*, Backbone Dynamics of a Free and a Phosphopeptide-Complexed Src Homology 2 Domain Studied by  $^{15}\text{N}$  NMR Relaxation. *Biochemistry* **33**, 5984-6003 (1994).
7. L. E. Kay; D. A. Torchia; A. Bax, Backbone dynamics of proteins as studied by  $^{15}\text{N}$  inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972-8979 (1989).
8. Y. Shen; A. Bax, Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227-241 (2013).
9. I. Solomon, Relaxation Processes in a System of Two Spins. *Phys. Rev.* **99**, 559-565 (1955).
10. N. Bloembergen; L. O. Morgan, Proton Relaxation Times in Paramagnetic Solutions. Effects of Electron Spin Relaxation. *J. Chem. Phys.* **34**, 842-850 (1961).
11. G. Lipari; A. Szabo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* **104**, 4546-4559 (1982).
12. G. M. Clore; J. Iwahara, Theory, Practice, and Applications of Paramagnetic Relaxation Enhancement for the Characterization of Transient Low-Population States of Biological Macromolecules and Their Complexes. *Chem. Rev.* **109**, 4108-4139 (2009).
13. P. Güntert; W. Braun; K. Wüthrich, Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J. Mol. Biol.* **217**, 517-530 (1991).
14. P. Güntert, Automated NMR Structure Calculation With CYANA. In *Protein NMR Techniques*, (Downing, A. K., Ed. Humana Press, 2004) pp 353-378.
15. E. Krieger; G. Koraimann; G. Vriend, Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins: Struct., Funct., Bioinf.* **47**, 393-402 (2002).
16. U. Essmann *et al.*, A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577-8593 (1995).
17. E. Krieger *et al.*, Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins: Struct., Funct., Bioinf.* **77**, 114-122 (2009).
18. B. Bissaro; I. Isaksen; G. Vaaje-Kolstad; V. G. H. Eijsink; Å. K. Røhr, How a lytic polysaccharide monooxygenase binds crystalline chitin. *Biochemistry* **57**, 1893-1906 (2018).

19. S. Stoll; A. Schweiger, EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *J. Magn. Reson.* **178**, 42-55 (2006).
20. R. C. Gregory *et al.*, Activity, stability and 3-D structure of the Cu(II) form of a chitin-active lytic polysaccharide monooxygenase from *Bacillus amyloliquefaciens*. *Dalton Trans.* **45**, 16904-16912 (2016).
21. B. Bissaro *et al.*, Molecular mechanism of the chitinolytic monocopper peroxygenase reaction. *bioRxiv* 541292 (2019).
22. F. Neese, Sum-over-states based multireference ab initio calculation of EPR spin Hamiltonian parameters for transition metal complexes. A case study. *Magn. Reson. Chem.* **42**, S187-S198 (2004).
23. A. A. Gewirth; S. L. Cohen; H. J. Schugar; E. I. Solomon, Spectroscopic and theoretical studies of the unusual EPR parameters of distorted tetrahedral cupric sites: correlations to x-ray spectral features of core levels. *Inorg. Chem.* **26**, 1133-1146 (1987).
24. Y. J. Huang; R. Powers; G. T. Montelione, Protein NMR Recall, Precision, and F-measure Scores (RPF Scores): Structure Quality Assessment Measures Based on Information Retrieval Statistics. *J. Am. Chem. Soc.* **127**, 1665-1674 (2005).
25. A. Bhattacharya; R. Tejero; G. T. Montelione, Evaluating protein structures determined by structural genomics consortia. *Proteins: Struct., Funct., Bioinf.* **66**, 778-795 (2007).

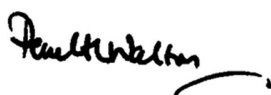
# 5 On the Mechanism of Substrate-O<sub>2</sub> Coupling in Lytic Polysaccharide Monooxygenases: EPR, X-Ray Crystallography, and DFT Studies of Azide Binding to a Chitin-Active LPMO

## Declaration

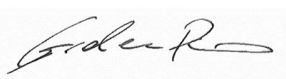
The following research paper draft is in its final stages and is soon to be submitted for peer-review. It builds on the previously mentioned publication, outlining the substrate induce perturbations that occur at the active sites of AA10 LPMOS. The manuscript details the use of a small molecule surrogate, azide, to mimic the electronic behaviour of superoxide bound to the active site of an AA10 LPMO, *BaAA10*. This study uses a combined approach of EPR spectroscopy, single crystal X-ray crystallography, UV-visible spectroscopy, and DFT/TD-DFT calculations to precisely characterise the structural and electronic characteristics of a Cu(II)-LPMO-N<sub>3</sub> complex and by analogy, the Cu(II)-LPMO-O<sub>2</sub> complex. The findings highlight the important implications that occur from substrate binding, including the generation of more potent oxidizing intermediates.

For this work Paul Walton and I conceived the original research ideas; I produced and purified enzyme samples for EPR/UV-visible studies; I performed the EPR experiments and simulated the spectra, I executed and analysed the UV-Vis studies; I carried out and analysed accordant DFT/TD-DFT calculations; I performed the O<sub>2</sub> binding experiments; Rebecca Gregory prepared, purified and crystallised *BaAA10* samples for X-ray studies; Glyn Hemsworth devised the original *BaAA10* preparation/purification protocol; Esther Johnson and Luisa Ciano prepared the *BaAA10*-Cu(II)-N<sub>3</sub> sample for K-band EPR. The EPSRC National Service for Electron Paramagnetic Resonance Spectroscopy centre in Manchester collected the K- and Q-band EPR spectra. Paul Walton and I wrote the manuscript; all co-authors contributed to the final manuscript.

Signed Prof Paul H. Walton



Signed Prof Gideon J. Davies



Signed Dr Alison Parkin



# On the Mechanism of Substrate-O<sub>2</sub> Coupling in Lytic Polysaccharide Monooxygenases: EPR, X-Ray Crystallography, and DFT Studies of Azide Binding to a Chitin-Active LPMO

Peter J Lindley<sup>1</sup>, Rebecca Gregory<sup>1</sup>, Glyn Hemsworth<sup>2</sup>, Alison Parkin<sup>1</sup>, Gideon J Davies<sup>1</sup> and Paul H Walton\*<sup>1</sup>

<sup>1</sup> Department of Chemistry, University of York, Heslington, York, UK. <sup>2</sup> Astbury Centre for Structural Molecular Biology and School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, UK

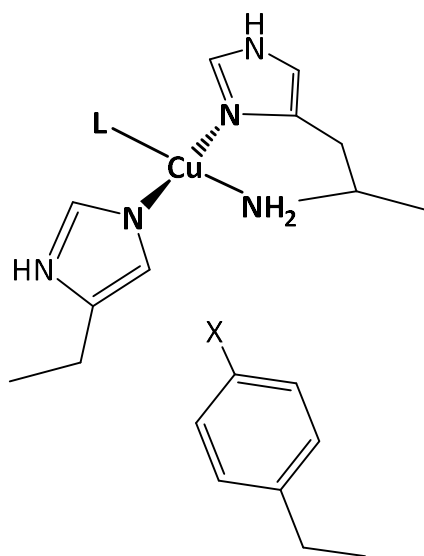
**ABSTRACT:** Lytic polysaccharide monooxygenases (LPMOs) are copper-containing enzymes that catalyze the oxidation of polysaccharide chains. The mechanism of action is proposed to proceed *via* a reactive copper-oxygen intermediate, such as [Cu-O<sub>2</sub>]<sup>+</sup> and/or [Cu-O]<sup>+</sup>. Here, we employ the fact that the frontier orbitals involved in the Cu(II)-azide interaction mimic that of the Cu(II)-superoxide interaction. On this basis we prepared Cu(II)-azide complexes at the active site of a chitin-active Cu(II)-AA10 LPMO (*BaAA10*), the EPR spectroscopic and structural features of which, by analogy, give insight into the nature of bonding within the equivalent Cu(II)-superoxide complexes. Using a combination of multi-frequency CW-EPR spectroscopy, UV/vis spectroscopy and DFT calculations, our results show in the absence of substrate ( $\beta$ -chitin) a five coordinate Cu(II)-azide complex is formed, in which the azide weakly coordinates to the basal plane of a Cu(II) in an overall trigonal bipyramidal coordination geometry. Following the addition of  $\beta$ -chitin to the Cu(II)-N<sub>3</sub> LPMO, the copper coordination number reduces from five to four, accompanied by an increase in covalency of the metal-ligand bond and a redistribution of spin density from the Cu(II) to the distal nitrogen atom of the azide. DFT calculations reveal that LUMO energies of the superoxide in Cu(II)-O<sub>2</sub> LPMO are lowered by the addition of substrate by 15 kJ mol<sup>-1</sup>. This lowering increases the reactivity of [Cu-O<sub>2</sub>]<sup>+</sup> to HAT from either the substrate or from a further reducing agent. The switch in frontier orbital character and energy upon substrate binding in AA10 LPMOs thus forms the basis of a coupling mechanism between substrate binding and oxygen activation.

## INTRODUCTION:

The at-scale generation of carbon-based fuels from sustainable feedstocks is a pressing global imperative. Of the technologies which offer promise in this regard, the production of cellulosic bioethanol is one that continues to attract interest and investment, not least because several second-generation biorefineries

around the world have already established operating viability. In this context, the controlled breakdown of recalcitrant biomass to fermentable sugars is one of the main technological challenges therein, requiring the rather delicate balance of enough chemical/mechanical power to degrade the biomass while, at the same time, causing no significant degradation of the sugar product and/or any catalyst which is used to affect the degradation. Such a combination of power and precision is one at which enzymes are paradigmatic, and it is of no surprise therefore that the enzymatic degradation of biomass, particularly polysaccharides, is the focus of both industrial and academic research programmes.

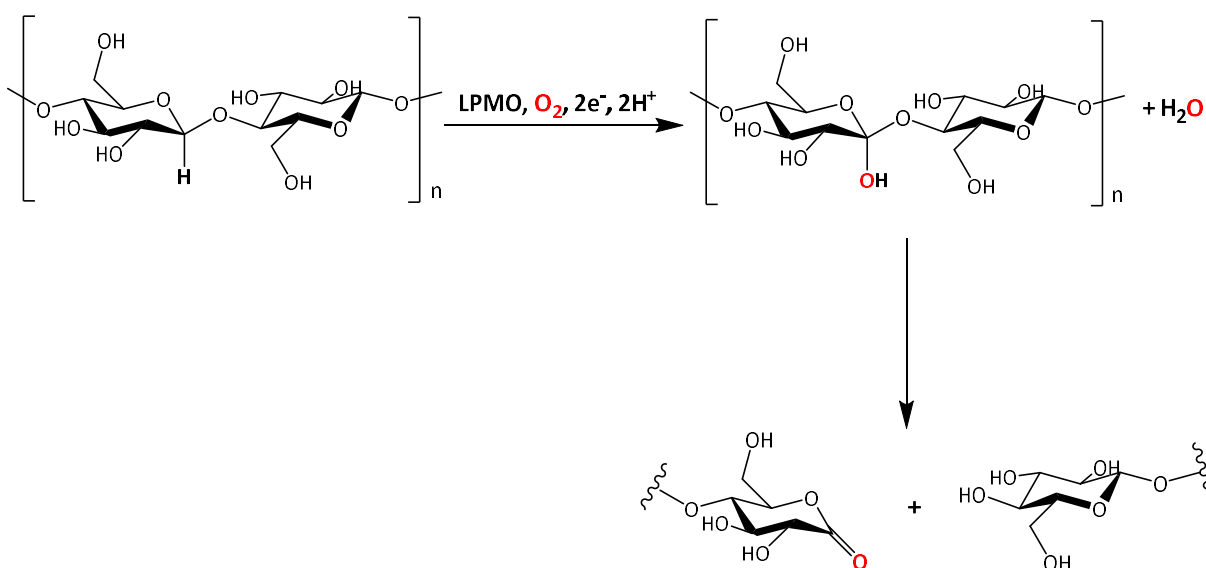
At the forefront of both commercial and laboratory research into biomass-degrading proteins are the enzymes lytic polysaccharide monooxygenases (LPMOs, also called PMOs). These enzymes are widespread across the Kingdom of Life, as shown by the CAZy database which lists eight sequence-distinct LPMO families: AA9-11 and AA13-17 that exist in the genomes of organisms as wide-ranging as bacteria, fungi, insects, molluscs and arthropods.<sup>1-7</sup> While the varied biochemical roles of LPMOs have yet to be fully revealed, what is known is that they boost several-fold the action of other polysaccharide-active enzymes on a range of recalcitrant polysaccharides including cellulose, starch and chitin.<sup>8</sup> In fact, due to the ability of LPMOs to operate synergistically with hydrolytic enzymes, enzyme cocktails containing mixtures of LPMOs are now routinely employed in biorefineries for biomass saccharification, and—it is argued—have contributed to the current renaissance in cellulosic bioethanol production.<sup>9</sup>



**Figure 1. Canonical LPMO active site with histidine brace motif. L denotes exogenous ligand(s) (either H<sub>2</sub>O, OH<sup>-</sup>, or Cl<sup>-</sup>) and X can be either H, or OH.**

Unusually for polysaccharide-degrading enzymes, the mechanism of action of LPMOs is oxidative as opposed to hydrolytic modification of the polysaccharide. This oxidation is performed by a mononuclear

copper active site, at which a single copper ion is coordinated by two conserved histidine residues in a “histidine brace” configuration (**Figure 1**).<sup>10</sup> Utilising either O<sub>2</sub> or H<sub>2</sub>O<sub>2</sub> as an oxidant and O atom source for incorporation into the product, LPMOs catalyse the conversion of C-H to C-OH at a saturated C atom on the polysaccharide (**Figure 2**), where it is estimated that the C-H bond dissociation enthalpy is 90-100 kcal mol<sup>-1</sup>. Breaking such a C-H bond necessitates that a species with high reduction potential and/or basicity is generated at the active site, for which several proposals exist ranging from hydroxyl radicals to reactive copper species (e.g. [Cu-O<sub>2</sub>]<sup>+</sup> and [Cu-O]<sup>+</sup>), the latter of which have the merit of being able to perform site-specific attack on the polysaccharide, as is affirmed by DFT calculations.<sup>11, 12</sup>



**Figure 2. Oxygenation of cellulose catalyzed by LPMOs followed by spontaneous lysis of the glycosidic bond.**

Notwithstanding the reactivity of the intermediate however, if non-specific attack of the substrate, enzyme and/or product is to be minimised, its generation must be coupled with both binding and positioning of the substrate at the enzyme’s active site. Some evidence to this effect is seen from laboratory experiments in which LPMOs are shunted with peroxides, which can display degradation of the enzyme to such an extent that the overall product release, whilst initially rapid, can be less than that observed when O<sub>2</sub> is used as an oxidant in conjunction with reducing agents. Additionally, attempts to use peroxide in larger scale saccharifications of biomass by LPMOs and other enzymes only seem to be viable when peroxide addition to the biomass-enzyme slurry is tightly restricted—a situation unlikely to be replicated *in vivo*. Moreover, at a molecular level of detail, DFT calculations of the catalytic cycle of AA9 LPMOs show that once H<sub>2</sub>O<sub>2</sub> interacts with the active site of a LPMO-Cu(I), the energy pathway to the generation of any of the reactive intermediates (including the deleterious hydroxyl radical) is met with barriers no higher than 6.9 kcal mol<sup>-1</sup>.

In other words, the forward reaction pathway in the mechanism of AA9 LPMOs beyond the formation of LPMO-(Cu...H<sub>2</sub>O<sub>2</sub>)<sup>+</sup> proceeds rapidly with the potential to be indiscriminately deleterious in the absence of substrate, a fact seemingly borne out by mass spectrometric experiments of LPMOs which show the number of amino acid sites which become oxidised upon treatment with H<sub>2</sub>O<sub>2</sub>.

Thus, for LPMOs to embody the principle described above of high chemical reactivity directed towards a specific C-H bond within the substrate, which is simultaneously benign to both enzyme and products, it is evident that there must exist a substrate-O<sub>2</sub> coupling mechanism on the reaction pathway. Additionally, for this reasoning to be correct there must also be at least one high energetic barrier on the reaction pathway *before* the generation of any (Cu...H<sub>2</sub>O<sub>2</sub>)<sup>+</sup> intermediate in LPMOs, the height of which is modulated by substrate binding to the enzyme.

In this regard, strong experimental evidence exists which shows that the reaction of <sup>3</sup>O<sub>2</sub> directly with any genomic class of LPMO-Cu(I) to give a LPMO-[Cu(O<sub>2</sub>)]<sup>+</sup> species in the *absence* of substrate is uniformly slow, thus providing a natural energetic barrier to uncoupled turnover.<sup>13</sup> Accordingly, any substrate-O<sub>2</sub> coupling step is likely associated with substrate-induced reduction of this barrier's height and, potentially also to the energetic barrier of HAT from the substrate (or another reducing agent) by the LPMO-[Cu(O<sub>2</sub>)]<sup>+</sup> intermediate which forms immediately after O<sub>2</sub> addition to LPMO-Cu(I). On this basis, the rate of formation, stability and reactivity of LPMO-[Cu(O<sub>2</sub>)]<sup>+</sup> from the reaction of LPMO-Cu(I) and O<sub>2</sub> are therefore pivotal to determining the details of any O<sub>2</sub>-substrate coupling mechanism in all genomic classes LPMOs.

To this end, previous EPR studies demonstrated that substrate binding to AA10 LPMOs drives a change in the coordination number of the copper ion at the active site, and this was proposed at the time, using DFT calculations, to be the basis of a substrate-O<sub>2</sub> coupling mechanism.<sup>7</sup> This proposal remains to be tested experimentally however since the inherent reactivity of the LPMO-[Cu(O<sub>2</sub>)]<sup>+</sup> intermediate precluded its direct study. As such, we now extend this earlier study by taking advantage of the fact that LPMO-[Cu(N<sub>3</sub>)]<sup>+</sup> complexes exhibit similar frontier orbitals as their LPMO-[Cu(O<sub>2</sub>)]<sup>+</sup> counterparts, yet unlike the latter the LPMO-[Cu(N<sub>3</sub>)]<sup>+</sup> complexes are amenable to study due to their chemically stable nature and the fact that they exist as a spin doublets and are thus accessible to standard CW-EPR studies.<sup>14-18</sup> Using this approach we prepared and characterised azide complexes of a chitin-active AA10 LPMOs, *Bacillus amyloliquefaciens* (BaAA10), using a combination of X-ray crystallography, multi-frequency EPR spectroscopy and DFT calculations. From these studies we affirm the substrate-driven change in coordination number of the copper at the active site of LPMOs, and—further—demonstrate a change in the covalency of the metal-ligand bond in [Cu(N<sub>3</sub>)]<sup>+</sup>, *ergo* [Cu(O<sub>2</sub>)]<sup>+</sup>, which is accompanied by large shifts in

the energy of the ligand LUMO and the spin density of the ligand distal atom. Both of these substrate-induced changes to the frontier orbitals of azide provide a conceptual framework for understanding the O<sub>2</sub>-substrate coupling mechanism in AA10 LPMOs.

## Results and Discussion:

### Binding of Azide to AA10 LPMO-Cu(II)

We first determined the extent of azide (N<sub>3</sub><sup>-</sup>) binding to the Cu(II)-LPMO in solution at pH 6. Using the well-known LMCT band at *ca* 385 nm as a marker of Cu(II)-azide bond formation, we titrated aqueous NaN<sub>3</sub> into a 0.15 mM solution of *BaAA10*-Cu(II), until the absorbance of the LMCT band reached a maximum. Full binding was achieved at azide to LPMO ratios of *ca.* 500:1 for showing that coordination of azide to the resting state of the enzyme is weak, with an approximate value of  $K = 110 \text{ dm}^3 \text{ mol}^{-1}$  (after appropriate fitting to the variation of azide/enzyme concentration to the intensity of the 385 nm band, Figure S1). The rather weak association constants of the azide to the enzyme are qualitatively in accord with the observed slow rate of Cu(II)-superoxide formation in the reaction of Cu(I)-AA10 LPMOs with O<sub>2</sub> and the reported crystal structure of a putative superoxide complex of an LPMO which exhibits a Cu...superoxide contact of 1.90 Å.<sup>13, 19, 20</sup>

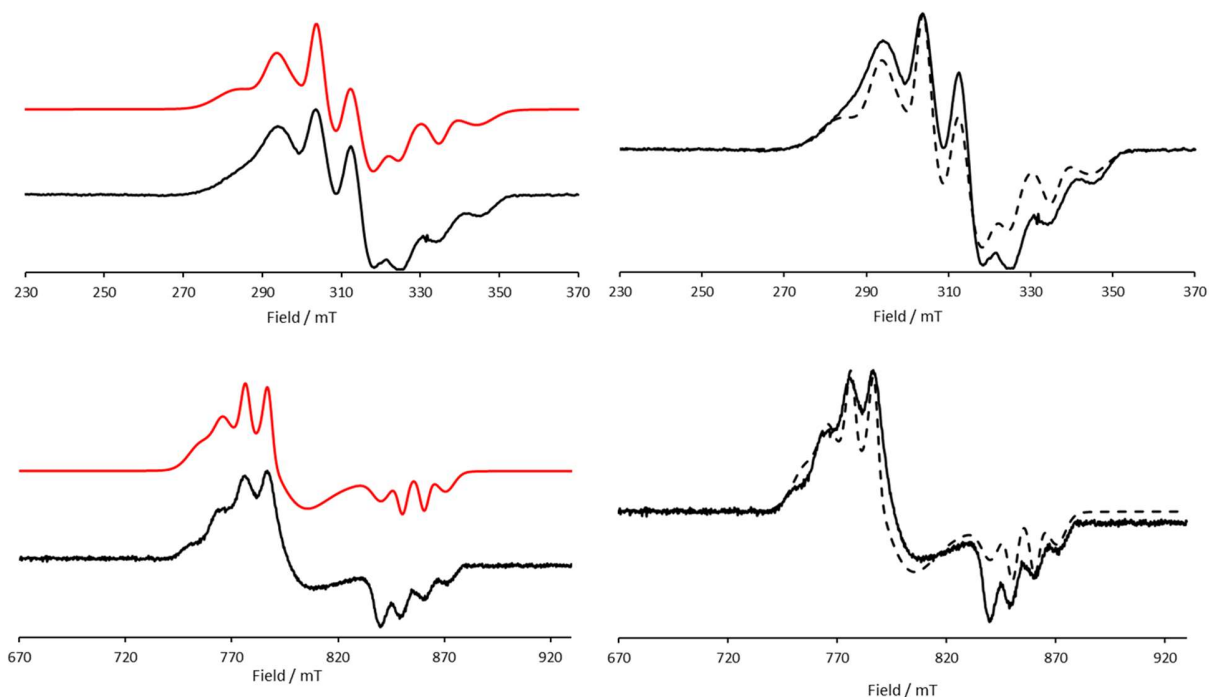
### Reoxidation of LPMO-Cu(I) with O<sub>2</sub>

We then determined the rate of *BaAA10*-Cu(II)-superoxide formation by performing stopped-flow spectrophotometry mixing an oxygen-free solution of *BaAA10*-Cu(I) (0.3 mM) with aerated MES buffer (20 mM pH 6) monitoring the growth of the characteristic d-d absorption band. The reoxidation was found to have a pseudo-first order rate constant of  $k = 2.3 \times 10^{-3} \text{ s}^{-1}$ , taking a period of minutes to fully reoxidize the sample (Figure S2); akin to the findings of Bissaro *et al.* on an analogous AA10 LPMO, *SmAA10*.<sup>19</sup> These data elude to a slow *outer*-sphere electron transfer (OSET) process between LPMO-Cu(I) and O<sub>2</sub> in AA10 LPMOs when the enzyme is unbound to polysaccharide substrate (see theoretical OSET discussion in the Supporting Information). This is markedly different to the *inner*-sphere electron transfer reoxidation mechanism observed for AA9 LPMOs which have pseudo-first order rate constants of  $>0.15 \text{ s}^{-1}$  (showing full reoxidation in *ca.* 5 s).<sup>13</sup> AA10 LPMOs therefore show a different mechanism of O<sub>2</sub> activation which has shown to be uniformly slow in the absence of polysaccharide substrate.



## X and K-band cw-Electron Paramagnetic Resonance Spectroscopy

EPR spectroscopy has been pivotal in building understanding about the active site electronic structure in LPMOs.<sup>21</sup> In the context of work reported herein, the EPR spectra of the *BaAA10* ‘resting state’ and spectroscopically similar AA10 LPMOs have been previously determined, in which a high degree of rhombicity ( $g_z \neq g_y \neq g_x$  and  $A_z \neq A_y \neq A_x$ ) is seen in their spin Hamiltonian parameters.<sup>7, 22, 23</sup> Additionally, the  $A_z$  value is also significantly reduced (ca. 350 MHz) compared to the other families, such as the AA9s (ca. 490 MHz). Since EPR has proven a robust technique for studying LPMOs electronic structure, we sought to characterise the LPMO-Cu(II)-N<sub>3</sub> complex, which by analogy, is informative of a LPMO-Cu(II)-O<sub>2</sub> complex. Given the overall weak binding constant of azide to the enzyme, all EPR studies were performed with a 1000× molar excess of azide, to ensure saturation of the Cu(II) active site. Under such conditions, the overall spectral envelope of the CW-EPR spectra (X-band and K-band) of *BaAA10* retains the overall highly rhombic envelope seen in the resting state of the enzyme, albeit with some significant changes in spin Hamiltonian parameters (**Figure 3**, Table 1). The values show even greater rhombicity than that of the resting state spectrum.



**Figure 3. Top: X-band CW-EPR spectra of *BaAA10* (0.55 mM in 20 mM MES buffer pH 6.0) with 1000 eq. of NaN<sub>3</sub> (black) and simulation shown in red (left) and dashed (right). Bottom: K-band CW-**

EPR spectra of *BaAA10* (1.2 mM in 20 mM MES buffer pH 6.0) with 1000 eq. of  $\text{NaN}_3$  (black) and simulation shown in red (left) and dashed (right).

**Table 1. Spin-Hamiltonian parameters obtained from simulation of multi-frequency EPR spectra.**

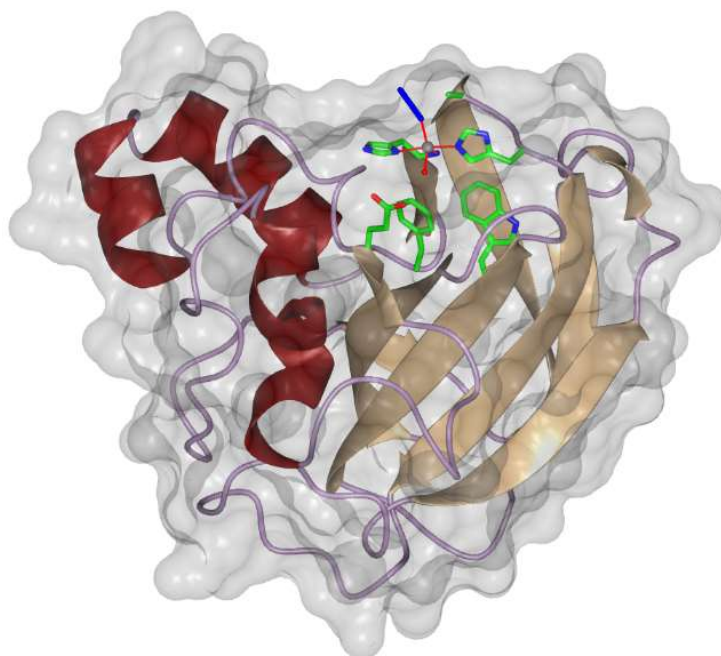
Spectrometer Frequency	<i>g</i> - values			Cu Hyperfine / MHz		
	$g_x$	$g_y$	$g_z$	$A_x$	$A_y$	$A_z$
<b>X - band</b>	2.004	2.167	2.220	280	125	320
<b>K - band</b>	2.009	2.167	2.220	270	125	320

The  $g_x$  value is notably low (*ca.* 2.007) when compared to typical values for Cu(II) EPR spectra with an  $d(x^2-y^2)$ -dominated ground state. (Note here that some previous determinations of *g* values of AA10 LPMOs have confused a high  $A_z$  coupling with anomalously low  $g_x$  values).<sup>24</sup> Herein, however, this potential confusion is removed by the use of two different frequencies to determine a common set of spin Hamiltonian parameters. A combination of X-band (*ca.* 9 GHz), K-band (*ca.* 24 GHz) EPR spectroscopy experiments allowed for the more accurate determination of the *g*- and *A*-values of the spin Hamiltonian (**Figure 3**). Additionally, a Q-band (*ca.* 33 MHz) EPR spectrum was obtained for this species. Though the resolution of this spectrum was poor, it shows promising agreement with the same set of SH parameters (Figure S3). The  $g_x$  value shows little deviation from the free-electron *g*-factor (2.0023), more in keeping with EPR spectra of Cu(II) with a  $d(z^2)$  ground state,<sup>25</sup> revealing that the excited *d*-electron states associated with the  $g_x$  parameter in Cu(II)- $\text{N}_3$ -*Ba*LPMO are well-separated energetically from the ground state, thus admitting only a small degree of orbital angular momentum in this direction. Conversely, the  $g_y$  value = 2.167, which is large for that commonly seen in Cu(II) species. The SH parameters altogether are informative of a 5-coordinate Cu(II) in which azide is bound, where the final electronic structure presents a large degree of  $d(z^2)$  mixing into the ground state.

Thus, at a qualitative level, the large difference in  $g_x$  and  $g_y$  values are commensurate with significant levels of rhombic distortion in the magnetic frame and thus the ligand field surrounding the copper in *BaAA10*-Cu(II)- $\text{N}_3$ . Adding to this notion, the further reduction in the value of  $A_z$  (*ca.* 40 MHz) is commensurate with increased delocalisation of the SOMO in the  $g_x$  value direction when compared to the resting state, as might be expected from the coordination of azide to Cu with its high nephelauxetic parameter.<sup>26</sup> While the EPR is indicative both a coordination geometry at the Cu in which the ligand field strength is unequal in all three principal directions of the *g* matrix and an increased covalency over the resting state, the exact geometrical details cannot be determined.

### Single Crystal X-ray Diffraction (XRD) of *BaAA10*-Cu(II)-N<sub>3</sub>

To this end, we successfully crystallised a solution of *BaAA10* in the presence of excess azide anion and obtained the resulting structure of the protein by single crystal X-ray diffraction (**Figure 4**). The resulting structure shows fantastic agreement with the spectroscopic data – incorporating a single azide ion in the coordination sphere (confirmed by Hill plot analysis, Figure S1) and adopting a highly rhombic coordination geometry (confirmed by the EPR spin-Hamiltonian parameters). The copper coordination geometry is best described as trigonal bipyramidal, with the His1(-NH<sub>2</sub>) and the two exogenous ligands making up the trigonal plane separated by approximately 120 degrees. The remainder of the protein structure appears unchanged from the resting state coordinates previously characterised.<sup>27</sup>



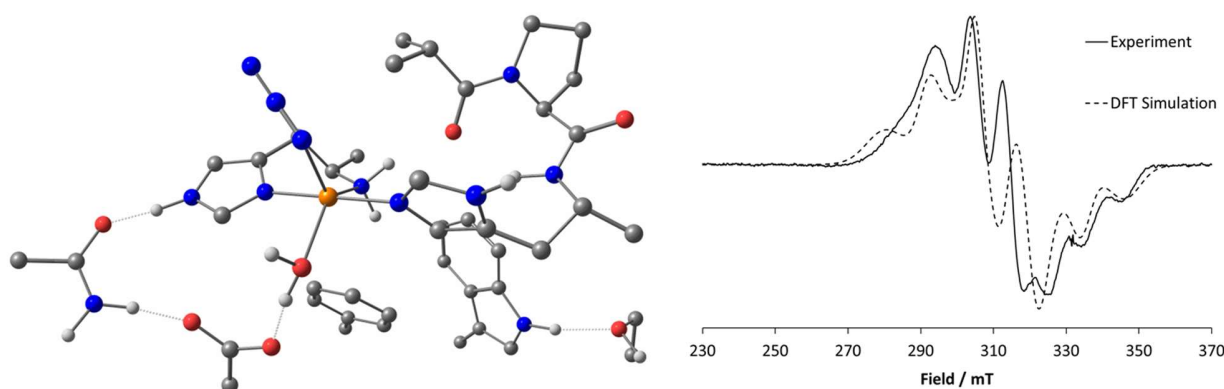
**Figure 4.** X-ray crystal structure of *BaAA10* incorporating an azide anion in the coordinate sphere (PDB:1TBD). Active site residues shown as sticks and colored by atom type,  $\alpha$ -helices are shown in crimson,  $\beta$ -sheets are shown in brown, and chains are shown in purple.

### Density Functional Theory (DFT) – Geometry Optimization and EPR Calculations

To corroborate the spectroscopy with theory, the crystallographic coordinates of the *BaAA10* resting state (PDB:5IJU, with a resolution of 1.70 Å) provided the basis for DFT calculations. These coordinates were chosen due to the comparatively improved resolution compared to the crystal structure of the *BaAA10*-N<sub>3</sub> complex (*ca.* 2 Å). From these initial coordinates, a cluster model was constructed in which a single azide

anion was added to the metal coordination sphere in keeping with the crystallographic, EPR and UV-visible results (model A, full computational details found in the Supplementary Information).<sup>23</sup>

The geometry was optimised using Becke's and Perdew's generalised gradient approximation (GGA) uBP86 DFT functional (left, **Figure 5**). The optimised geometry for this model showed high rhombicity, with the  $\angle(\text{NH}_2)\text{Cu}(\text{N}_3)$  retaining a distorted angle of  $131.5^\circ$  (wrt.  $\angle(\text{NH}_2)\text{Cu}(\text{H}_2\text{O}) = 131.9^\circ$  from crystallographic coordinates).<sup>23</sup> The DFT optimized coordinates show close agreement with those determined from the *BaAA10*-Cu(II)-N<sub>3</sub> crystal structure and are in accord with the expected electronic properties observed by EPR spectroscopy. This chemical LPMO-Cu(II)-N<sub>3</sub> species, resembling a LPMO-Cu(II)-O<sub>2</sub>, shows fantastic agreement between spectroscopy, crystallography, and theory; a level of characterization that would be tremendously difficult with the real oxidizing intermediate. The calculated spin-Hamiltonian parameters used in the simulation of this species derived from both simulation and DFT calculations are shown in **Table 2**.



**Figure 5.** Left: Optimized geometry of model A (hydrogen atoms attached to carbon atoms omitted for clarity). Right: EPR spectrum of *BaAA10* (0.55 mM in 20 mM MES buffer pH 6.0) with 1000 eq. of NaN<sub>3</sub> (black) and simulation using DFT derived spin-Hamiltonian parameters (dashed).

**Table 2.** Spin-Hamiltonian parameters used to simulate *BaAA10* with 1000 eq. NaN<sub>3</sub> determined by simulation and DFT calculations.

S-H parameters origin	<i>g</i> -values			Cu Hyperfine / MHz		
	<i>g<sub>x</sub></i>	<i>g<sub>y</sub></i>	<i>g<sub>z</sub></i>	<i>A<sub>x</sub></i>	<i>A<sub>y</sub></i>	<i>A<sub>z</sub></i>
Simulation <sup>a</sup>	2.004	2.167	2.220	280	125	320
DFT	2.018	2.140	2.217	333	-88.0 <sup>b</sup>	-392 <sup>b</sup>

<sup>a</sup> Spin-Hamiltonian parameters taken from the X-band simulation

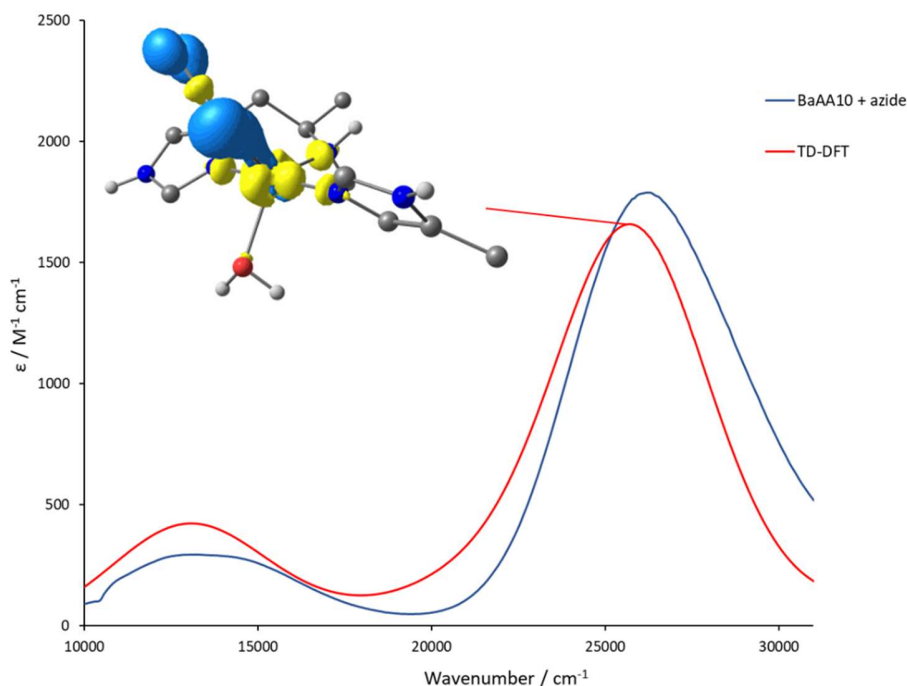
<sup>b</sup> Signs of HFC is obtainable via DFT methods but could not be ascertained by simulating the cw-EPR spectra

Some variation exists between the SH parameters determined by simulation compared with those obtained via computational methods. However, the overall trends are reproduced remarkably well. The rhombic  $g$ -values and  $A$  values are in keeping with those calculated for a low symmetry, 5-coordinate azide complex (model A), suggesting that this structure shows good resemblance with what we are observing in the spectroscopy. The overall fit from the simulation using the DFT-derived SH parameters appears to be mostly in accord with the experimental spectrum, with the biggest notable difference being the peak at *ca.* 310 mT, arising from an underprediction of the  $g_y$  parameter at 2.140 (wrt. 2.167 from simulation). The DFT calculations help provide an understanding of directionality in the electronic structure of the LPMO active site (Figure S8).

### Time-Dependent DFT (TD-DFT) – UV-Visible Calculations

In addition to the EPR properties of the LPMO-Cu(II)-N<sub>3</sub> species using DFT, we were able to reproduce the UV-visible spectra *via* TD-DFT calculations. The addition of sodium azide to solutions of *BaAA10* afforded a colour change from colorless/pale blue to an intense light green owing to the generation of a semi-intense ( $\epsilon \approx 2000 \text{ M}^{-1} \text{ cm}^{-1}$ ) UV-visible band centered at around  $26'200 \text{ cm}^{-1}$  (385 nm) (**Figure 6**). The positioning and intensities of the band are consistent with a ligand to metal charge transfer (LMCT), *ie.* where the transition is permitted by both spin and Laporte selection rules. The charge transfer from the non-bonding (nb) orbital of azide to the hole in the copper-based SOMO has a comparable transition energy to that of the Cu-O<sub>2</sub><sup>+</sup> moiety, which is characterized in the literature at around 386-395 nm.<sup>28, 29</sup> This indicates that the azide orbitals have a comparable energy and binding mode as end-on superoxide, confirming its suitability as surrogate for this study.

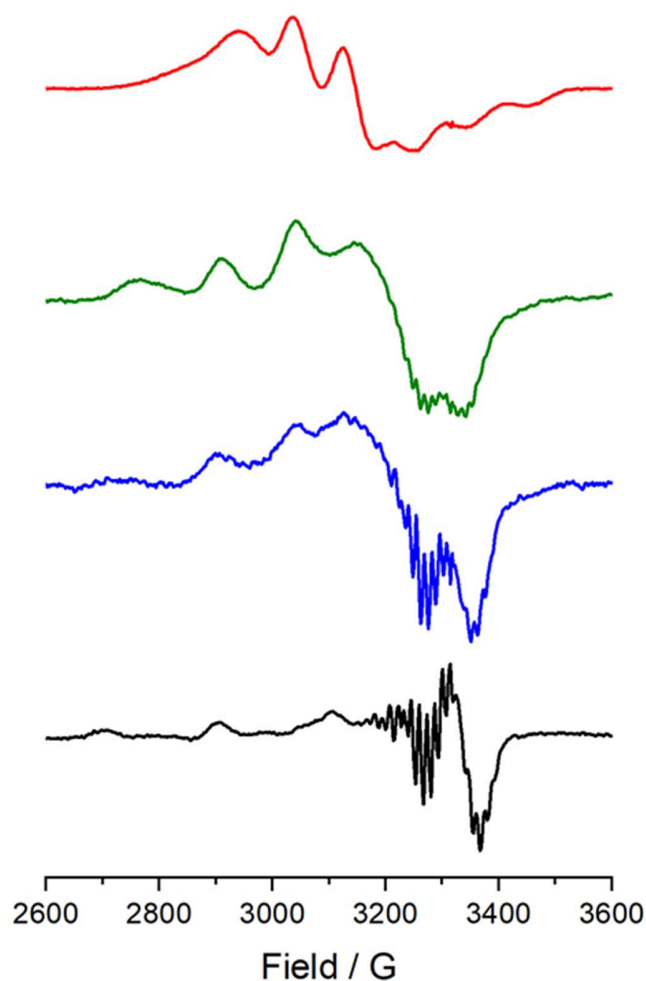
Using the same DFT-geometry optimized coordinates of *BaAA10*-Cu(II)-N<sub>3</sub> (model A, as used in the EPR calculations), TD-DFT calculations were performed to corroborate this experimental transition with theory (computational details outlined in the Supplementary Information). The generated spectrum shows good agreement with the experiment albeit with a slight redshift of *ca.*  $400 \text{ cm}^{-1}$  (**Figure 6**). Plotting the difference density from major band shows negative density localized on the nb azide orbital (blue) and positive density (yellow) situated on the  $d(x^2-y^2)$  orbital confirming this feature belongs to a LMCT transition. This level of characterization showing agreement between spectroscopy, crystallography and theory highlights the power of this approach.



**Figure 6. UV-Visible spectra of 0.5 mM *BaAA10* (blue) in MES (20 mM, pH 6.0), NaCl (200 mM) and NaN<sub>3</sub> (1000 equivalents) and spectrum generated by TD-DFT calculations (red). TD-DFT transition density from the 26'000 cm<sup>-1</sup> absorption band (negative density shown in blue, positive density shown in yellow).**

### **Substrate-induced changes to Cu(II)-N<sub>3</sub> / Cu(II)-O<sub>2</sub> electronic structure**

Following characterization of the 5-coordinate LPMO-Cu(II)-N<sub>3</sub> complex, we sought evaluate the effect of substrate binding to the overall electronic structure. As such, we performed CW-EPR experiments on samples of *BaAA10*-Cu(II)-N<sub>3</sub> after the addition of chitinous substrate. As previously reported with chitin-active AA10s, only partial binding of the protein to the substrate was seen giving rise to a mixture of species in the EPR spectra.<sup>22</sup> As such, we produced LPMO-Cu(II)-N<sub>3</sub>-substrate complexes where the substrate was added both before and after azide addition to properly characterize the substrate-bound species. The resulting EPR spectra revealed an inflation of the  $g_z$  and  $A_z$  parameters for the *BaAA10*-N<sub>3</sub> complex in keeping with the experiments of Bissaro *et al* and Courtade *et al* (green spectrum, **Figure 7**).<sup>7, 22</sup>



**Figure 7. Frozen solution, X-band EPR spectra of *BaAA10* +  $N_3$  (red), *BaAA10* +  $N_3$  + chitin\* (green), *BaAA10* + chitin +  $N_3$ \* (blue) and *BaAA10* + chitin (blue). All samples were collected at 150 K and were in a solution of 20 mM MES buffer at pH 6.0. \* Components added in the order shown.**

A notable difference after substrate addition is the disappearance of resolved features at high field, owing to the unusually low  $g_x$  and high  $A_x$  parameters of the previous species. These low  $g_x$  and large  $A_x$  parameters were indicative of an asymmetric, five-coordinate copper center that introduced considerable  $d(z^2)$  mixing into the SOMO. Therefore, the absence of these features is suggestive of an active site geometry that possesses heightened symmetry (wrt. *BaAA10*-Cu(II)- $N_3$  in the absence of chitin) that is now most likely 4-coordinate. However, the SH parameters required to simulate this species still demonstrate considerable rhombicity (figure S4 and table S1), indicating that the that ligand field is not uniform in an axially symmetric manner. Such rhombic distortions to the EPR spectra have also been observed for four-coordinate Cu-azide species in haemocyanins.<sup>30</sup>

What the data shows, is that the addition of  $\beta$ -chitin (substrate) to the *BaAA10-N<sub>3</sub>* complex leads to a considerable rearrangement of the ligand field and reduction in coordination number from five to four resulting in a significantly altered electronic structure. Substrate binding was also found to give rise to the appearance of SHFC in the x/y region of the spectrum. Such resolution of SHFC is a common feature following substrate binding in LPMOs, whereby the active site becomes structurally well-defined and has reduced translational freedom (decreasing strain and line broadening effects). The extent of SHFC resolved when bound to the substrate is particularly striking (upwards of 20 peaks) with the chitin-active AA10's when compared with the other LPMO families.<sup>7, 22</sup>

### Implications of Coordination Number on Cu(II)-O<sub>2</sub> Reactivity

In reference to the EPR data demonstrating a reduction in CN from 5-4, additional computational models were generated to compare the electronic structure of both the 5/4-coordinate LPMO-azide and 5/4-coordinate LPMO-superoxide species (models A-D; computational details in the Supplementary Information). The redistribution of ligand field after substrate binding causes an apparent increase in covalency of the copper-oxygen bond highlighted by the variation in spin population of the 5 and 4-coordinate states (**Table 3**). Löwdin spin population analyses reveal that for the LPMO-O<sub>2</sub><sup>+</sup> species the spin population localized on the copper ion falls from 54% to 41% following the reduction in coordination number.

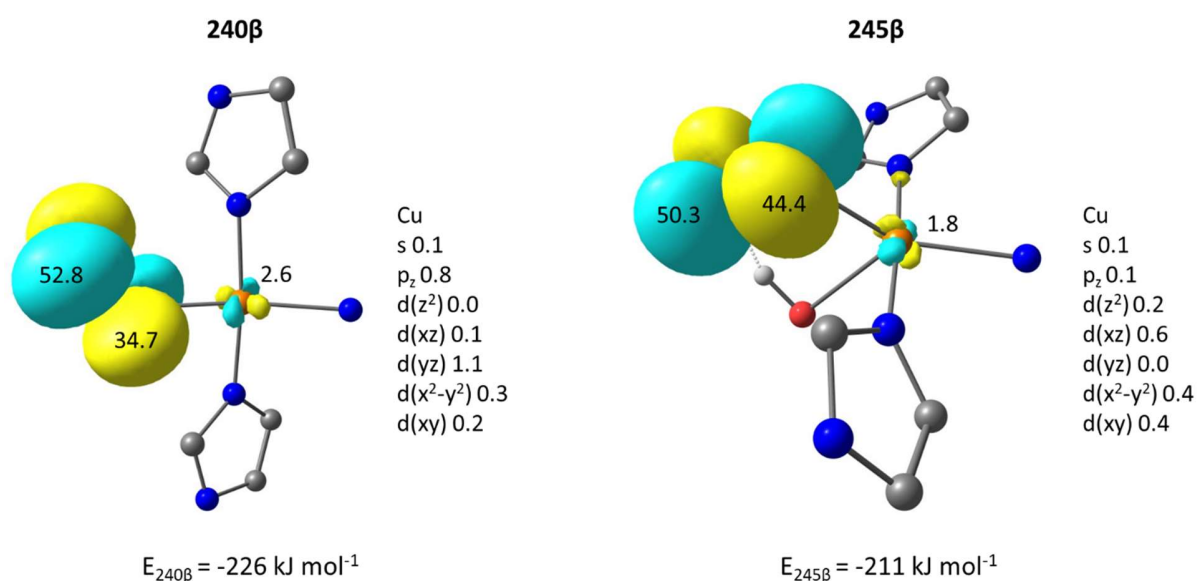
**Table 3. Copper and ligand Löwdin spin population analyses from DFT calculations on 4/5-coordinate LPMO-O<sub>2</sub> and LPMO-N<sub>3</sub> complexes**

DFT Model	Spin population* / %		
	Cu	X <sub>prox</sub>	X <sub>dist</sub>
<b>(A) 5-coordinate superoxide</b>	54	65	61
<b>(B) 4-coordinate superoxide</b>	41	72	73
<b>(C) 5-coordinate azide</b>	65	2.7	4.6
<b>(D) 4-coordinate azide</b>	60	5.4	9.4

\*spin population adding to a maximum of 200% on superoxide models (S=1) and 100% on the azide models (S=1/2)



Single point calculations were performed on the optimized coordinates of models A-D (computational details outlined in the Supplementary Information) to obtain energies, spin populations and orbital coefficients. This spin population is transferred near completely onto the oxygen atoms with a disproportionately high increase on the distal atom in the 4-coordinate models. This increase in covalency is a consequence of the exogenous ligand being oriented in the x/y plane allowing for better overlap of the ligand orbitals and the  $d(x^2-y^2)$ -dominated Cu SOMO. The mechanistically relevant orbital for carbohydrate oxidation was determined to be the low-lying  $\pi^*$  orbital on the superoxide ion due to its energy and proximity to the substrate binding surface (**Figure 8**).



**Figure 8. Superoxide  $\pi^*$  LUMO for 4-coordinate LPMO-O<sub>2</sub><sup>+</sup> (left) and 5-coordinate LPMO-O<sub>2</sub><sup>+</sup> species (right) calculated with DFT. The quoted values represent the percentage of atomic orbital parentage (per atom) for the respective molecular orbital. Further breakdown of atomic orbitals is provided for the copper ion.**

The calculations reveal crucial differences between the 5 and 4-coordinate models, not least of which is the lowering of redox active orbital energy by 15 kJ mol<sup>-1</sup> for the 4-coordinate structure. The lowering of this unoccupied orbital increases the oxidative potency of the intermediate and reducing the energetic barrier for hydrogen atom abstraction (HAA). Moreover, such a lowering of the superoxide RAMO energy is expected to ease a secondary reduction step of the O<sub>2</sub> co-substrate leading to the generation of the more potent Cu(II)-oxyl species outlined by Wang *et al.*<sup>12</sup> This provides further evidence of a coupling mechanism existing between substrate binding and oxygen activation, as originally discussed by Courtade *et al.*<sup>7</sup>

As well as lowering of the RAMO energy, there is also an attendant increase in the atomic orbital composition of this MO stemming from the distal oxygen atom upon expulsion of the water ligand (from 50.3 to 52.8%). This could be an important consequence of the forced 4-coordinate geometry by localizing the redox active orbital primarily on the atom that receives the hydrogen atom from the substrate. Upon forming the 4-coordinate species, the  $240\beta$  MO contains an inflated atomic orbital contribution from copper, primarily from the  $d(yz)$  and  $p(z)$  orbitals. The former of which is well positioned to  $\pi$  back-bond into the  $\pi^*$   $O_2$  SOMO, further strengthening the Cu-O bond and simultaneously activating the O-O bond. The latter group transition metals in period 4 are not renowned for their efficacy in  $\pi$ -back-bonding, owing to their high electronegativity. However, there is considerable evidence of synergic bonding for copper-oxygen complexes, in some cases to the extent of causing full scission of the O-O bond.<sup>31</sup> The latter component,  $p(z)$ , has a potentially critical effect of introducing 4p character into the superoxide SOMO, lowering the overall energy of the orbital. Therefore, the forced, 4-coordinate LPMO-Cu(II)- $O_2$  geometry expected after substrate binding gives rise to a low lying  $\pi^*$  orbital with very high distal oxygen character. The energy and composition of such orbital makes it primed for hydrogen atom abstraction or further reduction, akin to that seen in PHM chemistry.<sup>32</sup>

## Conclusions:

Herein, we have demonstrated that using the azide anion to mimic superoxide allows for the detailed exploration of electronic structures that are mechanistically relevant and ordinarily evade detection. The LPMO-Cu(II)- $N_3$  moiety served as a powerful spectroscopic probe to mimic an elusive ROS, LPMO-Cu(II)- $O_2$ , in a chitin-active chitin AA10 LPMO. The incorporation of azide into the LPMO coordination sphere gives rise to two novel electronic structures, showing unusually high degrees of covalency for a type-II copper site. The highly covalent bond formed between copper and azide is attributed to the high nephelauxetic parameter of  $N_3^-$  (and by extension,  $O_2^-$ ). *BaAA10* develops a highly rhombic electronic structure upon azide addition showing a considerable degree of  $d(z^2)$  mixing into the SOMO (*ca.* 10%). The EPR spectrum still displays an overall  $d(x^2-y^2)$  ground state but shows parallels to a Cu(II) spectrum with a  $d(z^2)$  ground state. The spin Hamiltonian parameters reveal the extent of the rhombicity with  $g_x$ ,  $g_y$  and  $g_z$  values of 2.009, 2.167 and 2.220, respectively, indicating a highly asymmetric, 5-coordinate copper center with a high degree of covalency. The active site was shown to adopt a 4-coordinate LPMO-Cu(II)- $N_3$  geometry following the binding of the substrate, chitin, *via* the expulsion of the  $H_2O/OH$  ligand. The resulting 4-coordinate geometry shows reduced rhombicity and a corresponding increase in covalency of

the Cu-N<sub>3</sub> bond. This increased covalency is reproduced in DFT calculations highlighting a drop in copper spin population of 5% in the azide models and 13% in superoxide models. Substrate addition was also shown to drive a 15 kJ mol<sup>-1</sup> depression in superoxide LUMO energy, reducing the energetic barrier for HAA of the substrate or production of more potent LPMO-Cu(II)-O species. We have therefore uncovered an oxygen activation pathway that is coupled to the substrate binding; such a coupling mechanism allows AA10 LPMOs to perform challenging site-specific oxidations whilst minimizing the risk of unselective or deleterious self-oxidations.

## Experimental methods:

### LPMO Preparation

*BaAA10* enzyme was obtained produced and purified using following the procedure previously published in 2016 by Gregory *et al.*<sup>23</sup> Briefly, the *BaAA10* gene was cloned into a Champion pET-SUMO vector and transformed into competent BL21 *E. coli* cells. Cell cultures were grown in baffled flasks containing Luria Broth (LB) media supplemented with kanamycin (30 µg ml<sup>-1</sup>) at 37 °C and 180 rpm. Once the culture reached OD<sub>600</sub> of 0.6-0.8, *BaAA10* production was induced by the addition of IPTG (1 mM final concentration) and incubated overnight at 16 °C and 180 rpm. The cells were then harvested by centrifugation (10,000 xg, 30 mins) before resuspending in 5x v:w buffer A (50 mM Tris pH 8, 200 mM NaCl, 30 mM imidazole). The cell suspensions were then lysed by sonication and the debris was removed by centrifugation (38,000 xg, 20 mins). The clarified supernatants were loaded onto a 5 mL His-trap (GE Healthcare) pre-equilibrated with buffer A before eluting with a 0 to 100% gradient of buffer B (50 mM Tris pH 8, 200 mM NaCl, 300 mM imidazole) over 20 CVs. SUMO-*BaAA10* containing fractions were pooled together treated overnight with SUMO protease (1:100 SUMO protease to SUMO-*BaAA10* w:w). The cleaved *BaAA10* enzyme mixture was loaded onto a second 5 ml His-trap (GE Healthcare) pre-equilibrated with buffer A and the flow through was collected. The protein solution was concentrated to < 2 ml and loaded onto a 16/60 Superdex 75 (GE Healthcare) column pre-equilibrated with MES buffer (20 mM pH 6). The protein was eluted from the column and 1 ml fractions were collected. Pure *BaAA10* fractions were collected, pooled, and concentrated for the various experiments. Protein purity was assessed by SDS-PAGE (Figure S6).

### UV-Visible Spectroscopy

A 0.15 mM sample of Cu(II)-*BaAA10* in MES buffer (20 mM, pH 6) was titrated against an aqueous solution of NaN<sub>3</sub> (1 M stock solution in MES buffer; 20 mM, pH 6). The azide concentration titrated stepwise so that 50 molar equivalents were introduced after each addition. After which the UV-visible spectrum was recorded on a benchtop spectrophotometer scanning between 200-1000 nm. The growth in LMCT maxima (*ca.* 385 nm) was used to determine the degree of azide coordination to the LPMO (further details in the Supplementary Information). The UV-vis spectrum of second sample of *BaAA10-N<sub>3</sub>* was recorded at greater protein concentration (0.5 mM) with 1000x molar excess of azide to generate the less noisy spectrum shown in **Figure 6**.

### **Electron Paramagnetic Resonance (EPR) Spectroscopy**

EPR samples were prepared in X-band, Quartz EPR tubes with a sample size of 100  $\mu$ L. The *BaAA10* enzyme (0.5 mM) was in a MES buffer solution (20 mM, pH 6.0) and NaCl (200 mM). For the azide sample preparation, a 1000-fold excess of sodium azide was added to the protein solution. Washed, crystalline, squid-pen chitin was then packed into the EPR tubes until the solution level was saturated with solid substrate. The protein-substrate samples were then incubated for 24 hours at 4 °C to allow time to bind. EPR spectroscopy was performed using a X-band BRUKER EMXmicro<sup>TM</sup> instrument with a frequency of *ca.* 9.3 GHz. And a magnetic field sweep range of 2300-3700 Gauss. A liquid nitrogen connection and thermocouple was used to maintain the cavity temperature at 150 K. Spectra were recorded using a 5.02 mW microwave power and a 4 G modulation amplitude. 1.2 mM and 0.5 mM samples of *BaAA10-Cu(II)-N<sub>3</sub>* were prepared in MES buffer (20 mM, pH 6) with 1000 molar equivalents of NaN<sub>3</sub> and their K- and Q-band EPR spectra were recorded by the EPSRC EPR National Service, respectively. Collected spectra were the intensity averaged over 3 scans. Simulations were carried out using the EasySpin 5.2.28 plugin for MATLAB.<sup>33</sup>

### **Acknowledgments:**

We thank Dr. Amga Baldansuren, Dr. Floriana Tuna, and Mr. Adam Brookfield for collecting the K-band EPR data at the EPSRC National Service for Electron Paramagnetic Resonance Spectroscopy. We thank Dr Luisa Ciano and Dr Esther Johnson for *BaAA10*-azide K-band sample preparation.

## References:

1. Raguz, S.; Yagüe, E.; Wood, D. A.; Thurston, C. F., Isolation and characterization of a cellulose-growth-specific gene from *Agaricus bisporus*. *Gene* **1992**, *119* (2), 183-90.
2. Schnellmann, J.; Zeltins, A.; Blaak, H.; Schrempf, H., The novel lectin-like protein CHB1 is encoded by a chitin-inducible *Streptomyces olivaceoviridis* gene and binds specifically to crystalline  $\alpha$ -chitin of fungi and other organisms. *Molecular Microbiology* **1994**, *13* (5), 807-819.
3. Hemsworth, G. R.; Henrissat, B.; Davies, G. J.; Walton, P. H., Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nat Chem Biol* **2014**, *10* (2), 122-126.
4. Vu, V. V.; Beeson, W. T.; Span, E. A.; Farquhar, E. R.; Marletta, M. A., A family of starch-active polysaccharide monooxygenases. *Proceedings of the National Academy of Sciences* **2014**, *111* (38), 13822.
5. Sabbadin, F.; Hemsworth, G. R.; Ciano, L.; Henrissat, B.; Dupree, P.; Tryfona, T.; Marques, R. D. S.; Sweeney, S. T.; Besser, K.; Elias, L.; Pesante, G.; Li, Y.; Dowle, A. A.; Bates, R.; Gomez, L. D.; Simister, R.; Davies, G. J.; Walton, P. H.; Bruce, N. C.; McQueen-Mason, S. J., An ancient family of lytic polysaccharide monooxygenases with roles in arthropod development and biomass digestion. *Nat Commun* **2018**, *9* (1), 756.
6. Filiatrault-Chastel, C.; Navarro, D.; Haon, M.; Grisel, S.; Herpoël-Gimbert, I.; Chevret, D.; Fanuel, M.; Henrissat, B.; Heiss-Blanquet, S.; Margeot, A.; Berrin, J.-G., AA16, a new lytic polysaccharide monooxygenase family identified in fungal secretomes. *Biotechnology for Biofuels* **2019**, *12* (1), 55.
7. Courtade, G.; Ciano, L.; Paradisi, A.; Lindley, P. J.; Forsberg, Z.; Sørli, M.; Wimmer, R.; Davies, G. J.; Eijsink, V. G. H.; Walton, P. H.; Aachmann, F. L., Mechanistic basis of substrate–O<sub>2</sub> coupling within a chitin-active lytic polysaccharide monooxygenase: An integrated NMR/EPR study. *Proceedings of the National Academy of Sciences* **2020**, *117* (32), 19178-19189.
8. Walton, P. H.; Davies, G. J., On the catalytic mechanisms of lytic polysaccharide monooxygenases. *Current Opinion in Chemical Biology* **2016**, *31*, 195-207.
9. Müller, G.; Várnai, A.; Johansen, K. S.; Eijsink, V. G. H.; Horn, S. J., Harnessing the potential of LPMO-containing cellulase cocktails poses new demands on processing conditions. *Biotechnology for Biofuels* **2015**, *8* (1), 187.
10. Quinlan, R. J.; Sweeney, M. D.; Lo Leggio, L.; Otten, H.; Poulsen, J.-C. N.; Johansen, K. S.; Krogh, K. B. R. M.; Jørgensen, C. I.; Tovborg, M.; Anthonsen, A.; Tryfona, T.; Walter, C. P.; Dupree, P.; Xu, F.; Davies, G. J.; Walton, P. H., Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proceedings of the National Academy of Sciences* **2011**, *108* (37), 15079.
11. Wang, B.; Johnston, E. M.; Li, P.; Shaik, S.; Davies, G. J.; Walton, P. H.; Rovira, C., QM/MM Studies into the H<sub>2</sub>O<sub>2</sub>-Dependent Activity of Lytic Polysaccharide Monooxygenases: Evidence for the Formation of a Caged Hydroxyl Radical Intermediate. *ACS Catalysis* **2018**, *8* (2), 1346-1351.
12. Wang, B.; Walton, P. H.; Rovira, C., Molecular Mechanisms of Oxygen Activation and Hydrogen Peroxide Formation in Lytic Polysaccharide Monooxygenases. *ACS Catalysis* **2019**, *9* (6), 4958-4969.
13. Kjaergaard, C. H.; Qayyum, M. F.; Wong, S. D.; Xu, F.; Hemsworth, G. R.; Walton, P. H.; Young, N. A.; Davies, G. J.; Walton, P. H.; Johansen, K. S.; Hodgson, K. O.; Hedman,

- B.; Solomon, E. I., Spectroscopic and computational insight into the activation of O<sub>2</sub> by the mononuclear Cu center in polysaccharide monooxygenases. *Proceedings of the National Academy of Sciences* **2014**, *111* (24), 8797-8802.
14. Casella, L.; Gullotti, M.; Pallanza, G.; Buga, M., Binding of azide and thiocyanate ligands to copper(II) model complexes. *Biology of Metals* **1990**, *3* (2), 137-140.
  15. Juda, G. A.; Shepard, E. M.; Elmore, B. O.; Dooley, D. M., A comparative study of the binding and inhibition of four copper-containing amine oxidases by azide: implications for the role of copper during the oxidative half-reaction. *Biochemistry* **2006**, *45* (29), 8788-800.
  16. Whittaker, M. M.; Whittaker, J. W., Ligand interactions with galactose oxidase: mechanistic insights. *Biophys J* **1993**, *64* (3), 762-772.
  17. Himmelwright, R. S.; Eickman, N. C.; LuBien, C. D.; Solomon, E. I.; Lerch, K., Chemical and spectroscopic studies of the binuclear copper active site of *Neurospora* tyrosinase: comparison to hemocyanins. *Journal of the American Chemical Society* **1980**, *102* (24), 7339-7344.
  18. Johannes, C.; Majcherczyk, A., Laccase activity tests and laccase inhibitors. *Journal of Biotechnology* **2000**, *78* (2), 193-199.
  19. Bissaro, B.; Streit, B.; Isaksen, I.; Eijnsink, V. G. H.; Beckham, G. T.; DuBois, J. L.; Røhr Å, K., Molecular mechanism of the chitinolytic peroxygenase reaction. *Proc Natl Acad Sci USA* **2020**, *117* (3), 1504-1513.
  20. O'Dell, W. B.; Agarwal, P. K.; Meilleur, F., Oxygen Activation at the Active Site of a Fungal Lytic Polysaccharide Monooxygenase. *Angewandte Chemie International Edition* **2017**, *56* (3), 767-770.
  21. Hemsworth, G. R.; Ciano, L.; Davies, G. J.; Walton, P. H., Production and spectroscopic characterization of lytic polysaccharide monooxygenases. In *Methods Enzymol*, 2018; pp 63-90.
  22. Bissaro, B.; Isaksen, I.; Vaaje-Kolstad, G.; Eijnsink, V. G. H.; Røhr, Å. K., How a Lytic Polysaccharide Monooxygenase Binds Crystalline Chitin. *Biochemistry* **2018**, *57* (12), 1893-1906.
  23. Gregory, R. C.; Hemsworth, G. R.; Turkenburg, J. P.; Hart, S. J.; Walton, P. H.; Davies, G. J., Activity, stability and 3-D structure of the Cu(ii) form of a chitin-active lytic polysaccharide monooxygenase from *Bacillus amyloliquefaciens*. *Dalton Trans* **2016**, *45* (42), 16904-16912.
  24. Chaplin, A. K.; Wilson, M. T.; Hough, M. A.; Svistunenko, D. A.; Hemsworth, G. R.; Walton, P. H.; Vijgenboom, E.; Worrall, J. A. R., Heterogeneity in the Histidine-brace Copper Coordination Sphere in Auxiliary Activity Family 10 (AA10) Lytic Polysaccharide Monooxygenases \*. *Journal of Biological Chemistry* **2016**, *291* (24), 12838-12850.
  25. Bennett, B.; Kowalski, J. M., EPR Methods for Biological Cu(II): L-Band CW and NARS. *Methods Enzymol* **2015**, *563*, 341-361.
  26. Schmidtke, H.-H.; Garthoff, D., The Electronic Spectra of Some Noble Metal Azide Complexes. *Journal of the American Chemical Society* **1967**, *89* (6), 1317-1321.
  27. Hemsworth, G. R.; Taylor, E. J.; Kim, R. Q.; Gregory, R. C.; Lewis, S. J.; Turkenburg, J. P.; Parkin, A.; Davies, G. J.; Walton, P. H., The Copper Active Site of CBM33 Polysaccharide Oxygenases. *Journal of the American Chemical Society* **2013**, *135* (16), 6069-6077.
  28. Abe, T.; Hori, Y.; Shiota, Y.; Ohta, T.; Morimoto, Y.; Sugimoto, H.; Ogura, T.; Yoshizawa, K.; Itoh, S., Cupric-superoxide complex that induces a catalytic aldol reaction-type C-C bond formation. *Communications Chemistry* **2019**, *2* (1), 12.

29. Tano, T.; Okubo, Y.; Kunishita, A.; Kubo, M.; Sugimoto, H.; Fujieda, N.; Ogura, T.; Itoh, S., Redox Properties of a Mononuclear Copper(II)-Superoxide Complex. *Inorganic Chemistry* **2013**, *52* (18), 10431-10437.
30. Alzueti, G.; Bubacco, L.; Casellai, L.; Rocco, G. P.; Salvato, B.; Beltramini, M., The Binding of Azide to Copper-Containing and Cobalt-Containing Forms of Hemocyanin from the Mediterranean Crab *Carcinus Aestuarii*. *European Journal of Biochemistry* **1997**, *247* (2), 688-694.
31. Elwell, C. E.; Gagnon, N. L.; Neisen, B. D.; Dhar, D.; Spaeth, A. D.; Yee, G. M.; Tolman, W. B., Copper-Oxygen Complexes Revisited: Structures, Spectroscopy, and Reactivity. *Chem Rev* **2017**, *117* (3), 2059-2107.
32. Cowley, R. E.; Tian, L.; Solomon, E. I., Mechanism of O<sub>2</sub> activation and substrate hydroxylation in noncoupled binuclear copper monooxygenases. *Proc Natl Acad Sci U S A* **2016**, *113* (43), 12035-12040.
33. Stoll, S.; Schweiger, A., EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *Journal of Magnetic Resonance* **2006**, *178* (1), 42-55.

# On the Mechanism of Substrate-O<sub>2</sub> Coupling in Lytic Polysaccharide Monooxygenases: EPR, X-Ray Crystallography, and DFT Studies of Azide Binding to a Chitin-Active LPMO

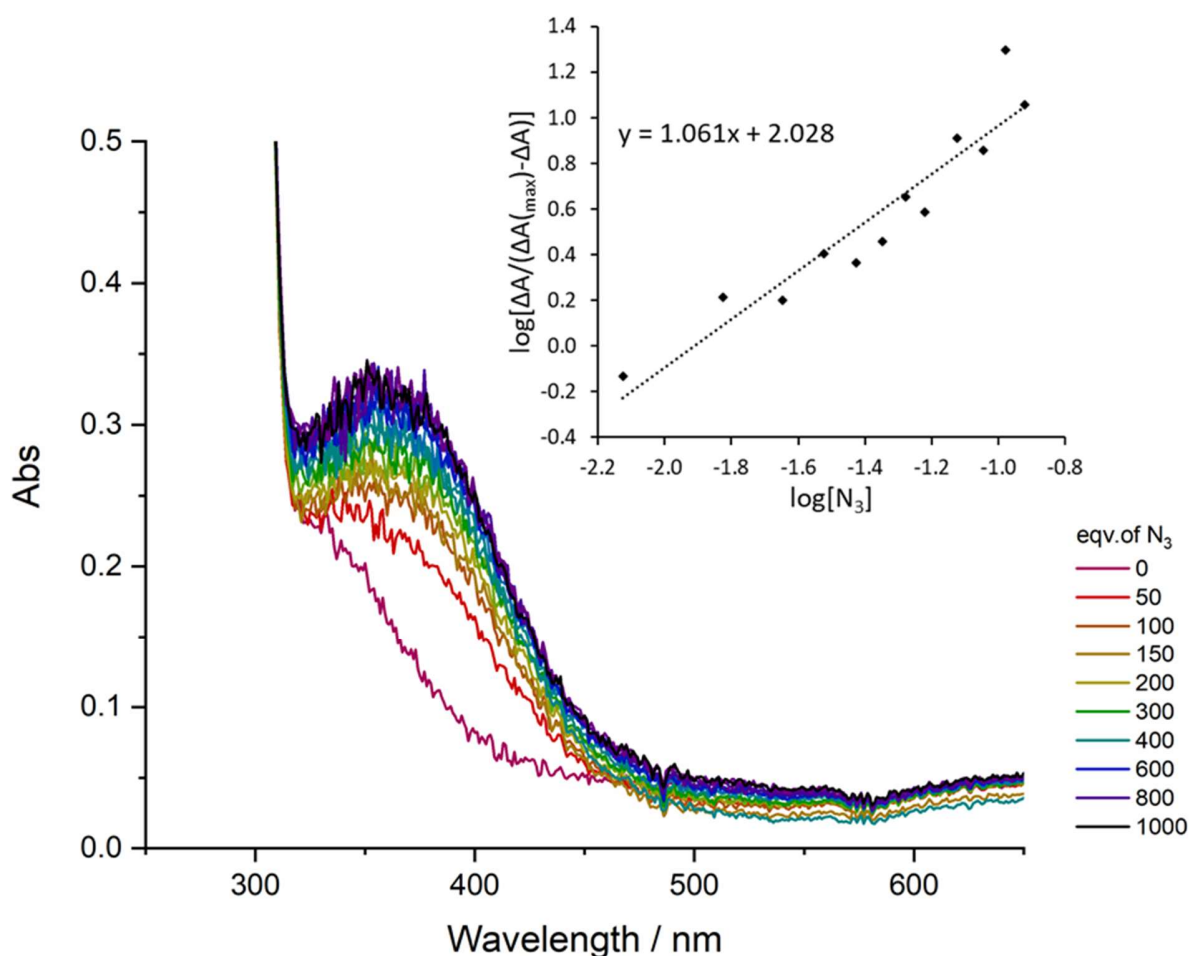
## Supplementary Information

### Contents

<b><u>UV-visible Spectra and Analysis of <i>BaAA10</i> + azide</u></b> .....	224
<b><u>Calculation of Outer-Sphere Electron Transfer Rate and Rate of Reoxidation of (Cu<sup>I</sup>)-<i>BaAA10</i></u></b> .....	226
<b><u>X-band EPR spectrum and Analysis of <i>BaAA10</i> + N<sub>3</sub> + Substrate</u></b> .....	228
<b><u>Purification of <i>BaAA10</i></u></b> .....	230
<b><u>Computational Details</u></b> .....	231
<i>Geometry Optimizations</i> .....	231
<i>EPR Property Calculations</i> .....	232
<i>Cartesian Reference System used in EPR and TD-DFT Property Calculations</i> .....	233
<i>Time-Dependent DFT (TD-DFT) Calculation</i> .....	234
<i>Single Point Calculations</i> .....	234



## UV-visible Spectra and Analysis of BaAA10 + azide



**Figure 9.** UV-Visible spectra of BaAA10 (0.15 mM in MES buffer; 20 mM pH 6) as a function of  $[N_3^-]$ . Insert shows Hill plot analysis of azide binding to BIAA10 active site

Hill plot was generated using the UV-Visible absorption data based off the equation:

$$\log \left[ \frac{\Delta A}{\Delta A_{max} - \Delta A} \right] = n \log[N_3^-] + \log K$$

Where  $\Delta A$  is the change in absorbance of the LMCT maximum (380 nm),  $\Delta A_{max}$  is the absorbance change of the LMCT band when azide is added to excess and absorbance maximum is seen to remain constant,  $n$  (the gradient) informs the number of azide anions that bind the complex and the  $\log K$  is the y-intercept and informs about the equilibrium constant. Here,  $n$ , is seen to approximately equal 1 showing that azide

coordinates the LPMO-Cu(II) with 1:1 stoichiometry. Figures S1 and subsequent equations were generated in line with the work of Serratice *et al.*<sup>1</sup>

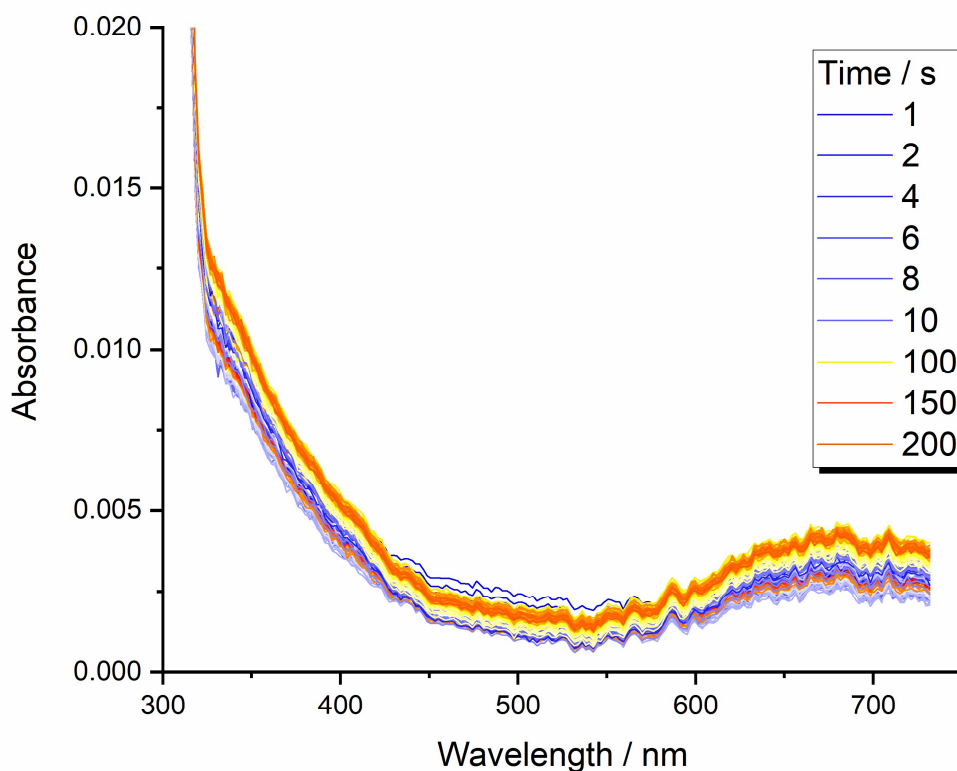
The regression line for the insert resulted as:

$$y = 1.061x + 2.028$$

Where the gradient is equal to the number of azide ions binding the Cu(II)-LPMO,  $n$ , and the intercept is equal to  $\text{Log}_{10}K$ .

This yields a value of  $K = 107$

and associated Gibbs free energy change at 293 K of  $\Delta G_{298} = -11.4 \text{ kJ mol}^{-1}$



**Figure 10. Reoxidation of BaAA10-Cu(I) in oxygenated buffer monitored by stopped-flow UV-Vis over 200 s**

Stopped flow UV-Vis spectra were obtained for the reaction of 300  $\mu\text{M}$  (Cu(I))-BaAA10 with aerated buffer (20 mM MES, pH 6.0). Maximum aeration of the buffer was achieved by cooling the solution down on ice and bubbling with compressed air for 30 minutes prior to data collection. The reoxidation of the BaAA10-Cu(I) solution was found to occur over approximately three minutes.

### Calculation of Outer-Sphere Electron Transfer Rate and Rate of Reoxidation of (Cu<sup>I</sup>)-BaAA10

The theoretical rate of outer-sphere electron transfer in LPMOs was estimated using Marcus theory via the following equations (also explained by Kjaergaard *et al.*).<sup>2</sup>

$$(1) k_{et} = Z e^{-\frac{\Delta G^\ddagger}{RT}}$$

$$(2) \Delta G^\ddagger = \left(\frac{\lambda}{4}\right) \left(\frac{\lambda + \Delta G^\circ}{\lambda}\right)^2$$

Where:

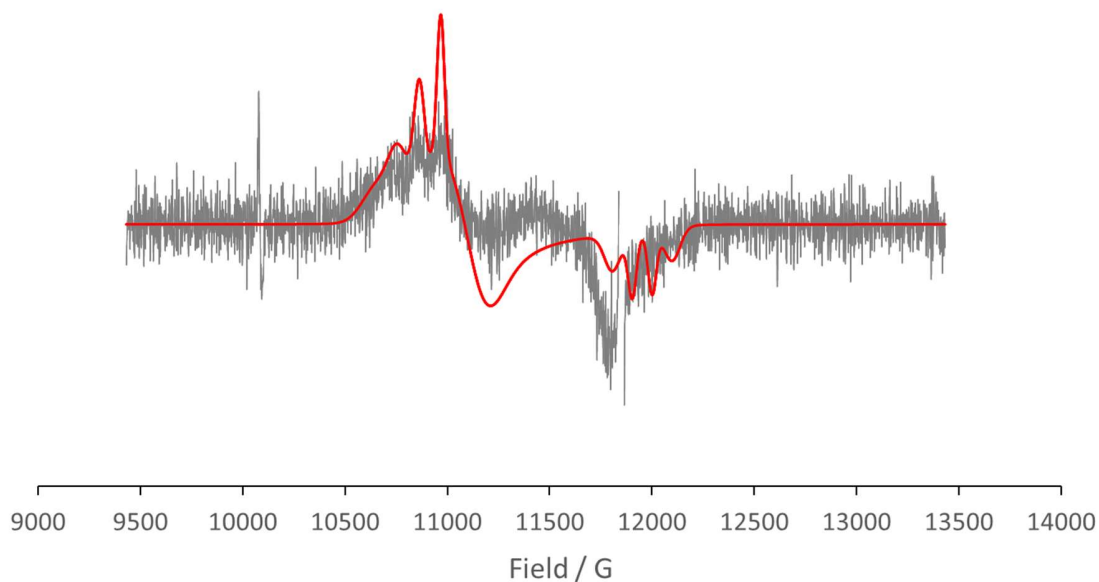
Z is the frequency factor,  $10^{11} \text{ M}^{-1} \text{ s}^{-1}$ .

$\lambda$  is the total reorganisation energy =  $(\lambda_{\text{Cu donor}} + \lambda_{\text{O}_2 \text{ acceptor}})/2 = 1.74 \text{ eV}$ , where  $\lambda_{\text{O}_2 \text{ acceptor}} = 1.89 \text{ eV}$  and  $\lambda_{\text{Cu donor}} = 1.58 \text{ eV}$  (due to lack of experimental data, the reorganisation energy for a Cu(I/II)-complex with a tris(2-pyridyl-methyl)amine ligand was used).

$\Delta G^\circ = -nRT\Delta E^\circ$ , where  $\Delta E^\circ = -0.16\text{V} (\text{O}_2/\text{O}_2^-) - 0.275\text{V} (\text{Cu(I/II)-AA10}) = 0.435 \text{ eV}$ .

Using these values, the electron transfer rate is predicted to be  $0.45 \text{ M}^{-1}\text{s}^{-1}$ , assuming an oxygen

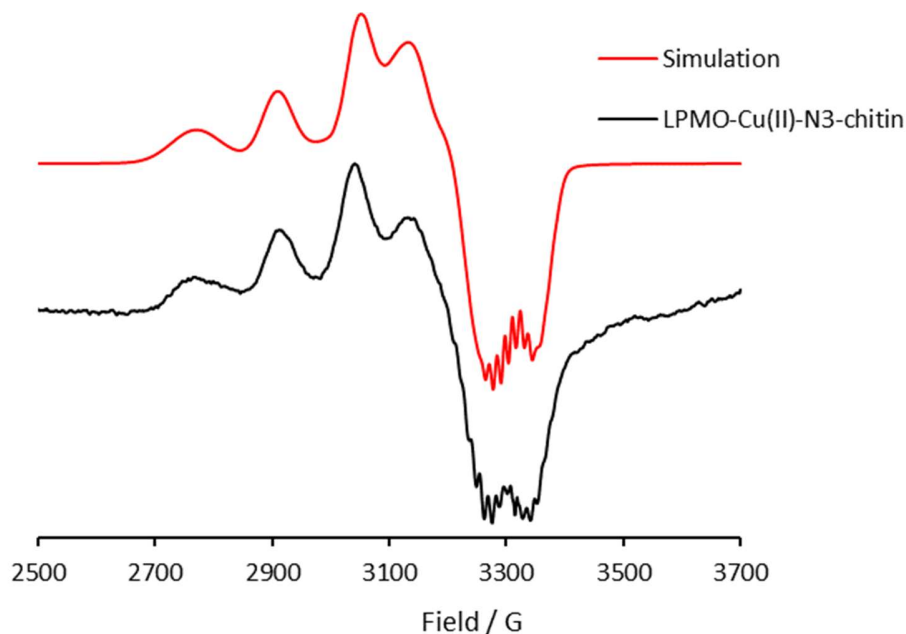
concentration in solution of 1 mM affords a first order rate constant of  $4.5 \times 10^{-4} \text{ s}^{-1}$ .



**Figure 11. Q-band EPR spectra of BaAA10 + N<sub>3</sub> (black) and simulation (red)**

A Q-band EPR spectrum was obtained for a sample of *BaAA10* (*ca.* 0.5 mM) in MES buffer (20 mM, pH 6) with 1000 molar equivalents of NaN<sub>3</sub>. The signal to noise ratio of this spectrum is comparatively poor due to practical challenges in concentrating the enzyme beyond 0.5 mM. Data was collected at the National EPR facility in Manchester by Murali Shanmugam.

### X-band EPR spectrum and Analysis of *BaAA10* + N<sub>3</sub> + Substrate



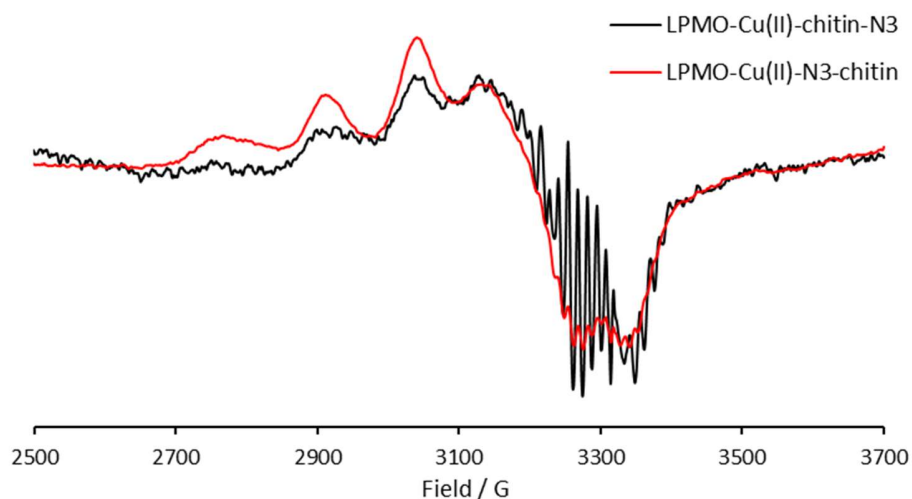
**Figure 12. X-band EPR spectra of *BaAA10* + N<sub>3</sub> + substrate (black) and simulation (red)**

Frozen solution, X-band EPR spectroscopy was performed on a solution of *BaAA10* + N<sub>3</sub> + chitin (black) using a Bruker MicroEMX spectrometer. Sample contained a protein concentration of *ca.* 0.5 mM and were in a buffer solution of 20 mM MES pH 6.0. A 1000-fold excess of sodium azide was added, and squid pen chitin was left to incubate with the sample for 24 hours prior to data collection. 20% of the intensity of the unbound species (*BaAA10* + N<sub>3</sub>) was removed to account for the fraction of protein unbound to the substrate. Simulation of this species (red dashed) was performed using the EasySpin 5.2.3 toolbox as implemented by Matlab 2020b.

**Table S1. Spin Hamiltonian parameters used to simulate the *BaAA10*-N<sub>3</sub>-substrate EPR spectrum**

g-matrix			Cu-hyperfine tensor (MHz)		
$g_z$	$g_y$	$g_x$	$A_z$	$A_y$	$A_x$
2.228	2.106	2.016	430	85	100

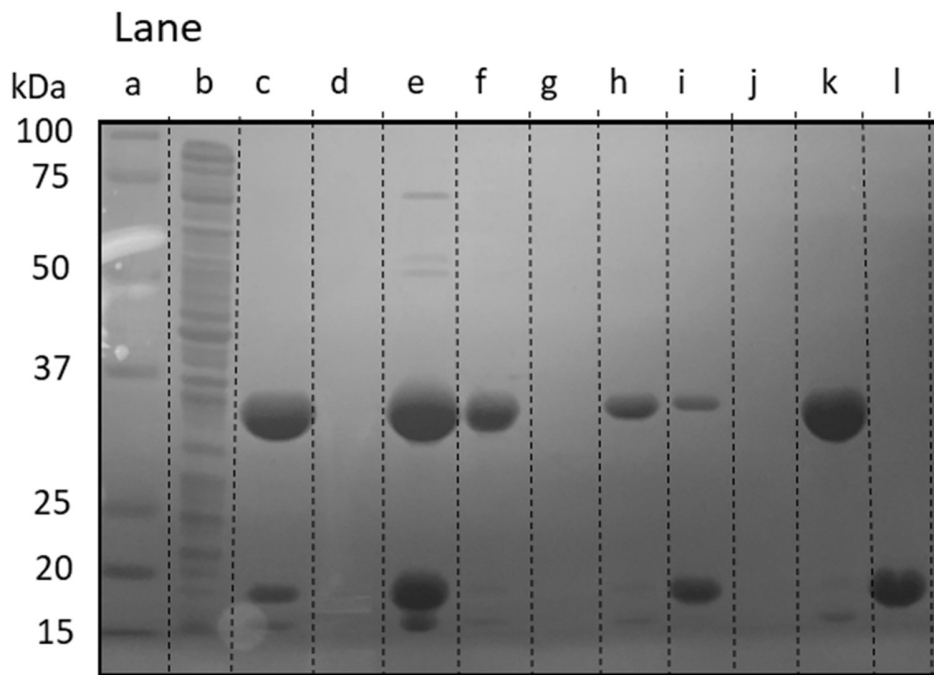
In addition to the SH parameters listed above, 3 nitrogen nuclei were included with a principal HFC value of 40, 40, and 30 MHz. However, without the employment of additional low frequency CW-EPR or pulsed EPR techniques, these values SHFC values carry a significant degree of error. Simulations performed with a line width of 0.5.



**Figure 13. Overlay of EPR spectra of BaAA10 + substrate + N<sub>3</sub> where the azide was added before and after substrate binding**

Frozen solution, X-band EPR spectroscopy was performed on a sample of *BaAA10* + chitin + N<sub>3</sub> (black) on a Bruker MicroEMX spectrometer. The sample contained a protein concentration of *ca.* 0.3 mM in a buffer solution of 20 mM MES, pH 6.0. Squid-pen chitin was added to the protein sample and left to incubate for 48 hours prior to data collection. After this time, a 1000-fold excess of sodium azide was added to each sample and left for an additional 24 hours to allow for diffusion. 50% of the intensity of the substrate-bound species without azide (*BaAA10* + chitin) was removed to account for the fraction of protein which the azide did not bind to. Considerable noise is evident in the spectrum due to the lowered concentration of protein following substrate addition. The SHFC pattern is also substantially different due to incoherence in the subtraction spectra. This spectrum was overlaid with the previous *BaAA10*+N<sub>3</sub>+chitin spectrum (where the substrate was added afterwards) highlighting the common single species between the two samples.

## Purification of *BaAA10*



**Figure 14. SDS-PAGE analysis of *BaAA10***

10% SDS-PAGE gel image showing various steps throughout the purification.

Lane contents:

- a) BioRad Precision Plus Protein Dual Colour Standard
- b) Supernatant of cells after resuspension
- c) Cell lysate contents
- d) Flow through
- e) Crude SUMO-*BaAA10* peaks from His-trap 1
- f) Crude SUMO-*BaAA10* peaks from His-trap 1 (different peak)
- g) Impurity peak from His-trap 1
- h) Pooled fractions from His-trap 1
- i) Crude cleaved *BaAA10* fractions from His-trap 2
- j) Impurity peak from His-Trap 2
- k) Uncleaved SUMO-*BaAA10* from His-trap 2
- l) Purified pooled fractions of *BaAA10* after SEC

## Computational Details

The initial atomic coordinates for the following computational models were taken from a *BaAA10*-Cu(II) crystal structure (PDB:5IJU, 1.7 Å) from the Carbohydrate-Active enZYmes (CAZy) database.<sup>3</sup> The models were truncated to include atoms from 9 important amino acid residues (His28, Glu68, Gln92, Ala123, Pro124, His125, Thr127, Trp187 and Phe196 (numbering starting at the first histidine is position 28) as well as the central copper ion and exogenous ligands (O487 and O412) and distal water molecule (O508). Valence hydrogen atoms were added to the atoms and the amino acids were truncated in the following way to minimize computational expense:

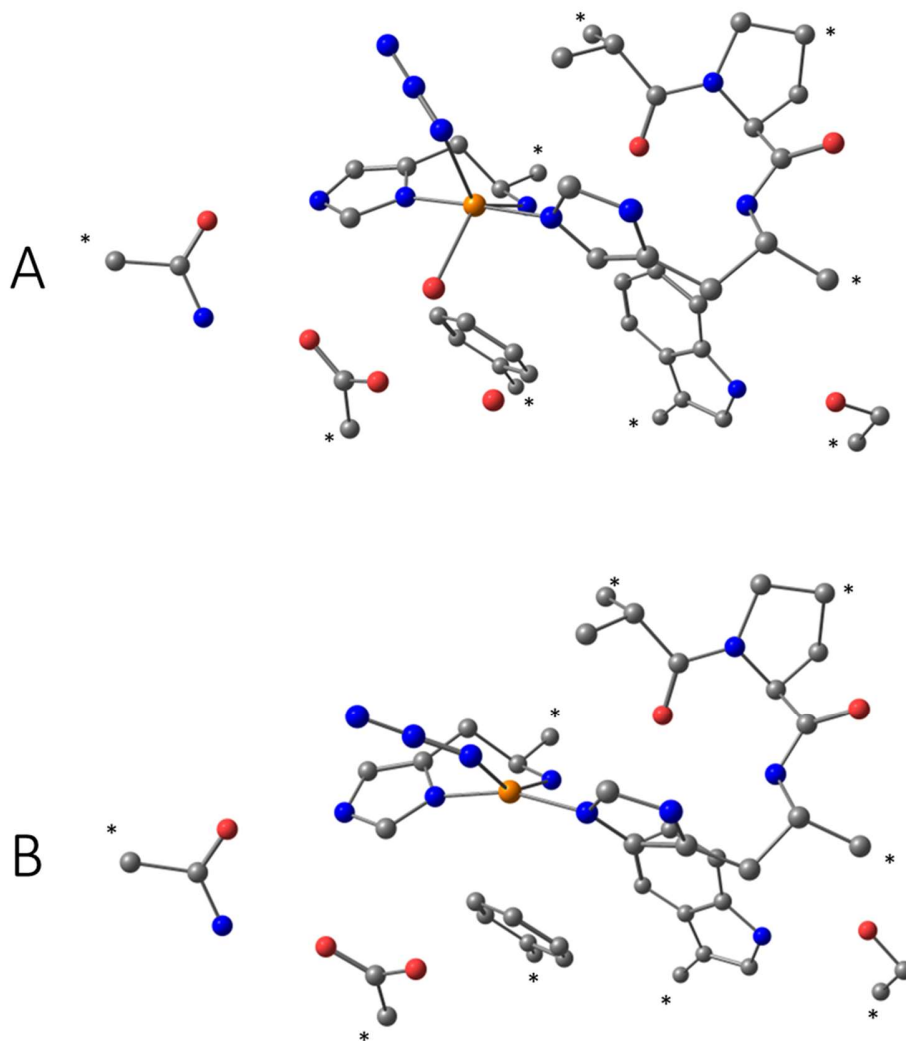
His28 and His125 were truncated at the carbonyl carbon, which was replaced by a methyl group, Glu68 and Gln92 were truncated with methyl substitution of the C $\gamma$ , the nitrogen of the amide bond between Ala123 and Thr122 was replaced by methyl groups, Trp187 and Phe196 were truncated with methyl substitution of the C $\beta$ , Thr127 was truncated with methyl substitution of the C $\alpha$  and the methyl group of C $\beta$  was removed. For Model A, the water molecule situated above the basal plane of the histidine brace (O487) was replaced with an azide anion (N $_3^-$ ), ligating the copper with an end-on coordination mode and the two remaining water molecules (O412 and O508) were retained. Model A was produced based on the 5-coordinate LPMO-Cu(II)-N $_3$  complex seen in the EPR and crystallographic study. For model B, all water molecules (O487, O412 and O508) were removed and replaced by a single azide ion with an end-on coordination mode placed in the now vacant equatorial position to model the azide LPMO-Cu(II)-N $_3$  complex after the addition of substrate.

### *Geometry Optimizations*

Geometry optimizations were performed using the Gaussian-09 software package at the DFT level of theory. Geometry optimizations of all models were performed using the generalized gradient approximation (GGA) functional, uBP86. Ahlrichs's Def-2-TZVP basis set was used to describe the copper, the first coordination sphere nitrogen atoms (including all three azide atoms) and oxygen atom of the copper-ligating water molecule. On all remaining atoms a Def2-SVP basis set was used. Solvation effects were accounted for using the polarizable continuum model with a dielectric constant of 80.0, as implemented by Gaussian 09. Empirical dispersion corrections were accounted for using Grimme's empirical dispersion (GD3). Certain atoms were kept frozen throughout the calculation to prevent movement that would be impossible if the constraints of the full size of the protein were accounted for (atoms frozen in calculations are highlighted with asterisks in **Error! Reference source not found.**). Analogous 5- and 4-coordinate LPMO-



Cu(II)-superoxide models were assembled and geometry optimized on the triplet potential energy surface under the same functional and basis set regime (models C & D, respectively).



**Figure 15. Optimized geometries of model A (top) and model B (bottom) highlighting constrained atoms with asterisks. Carbon atoms coloured with grey, nitrogen with blue, oxygen with red, and copper with orange. Hydrogen atoms omitted from figure for clarity.**

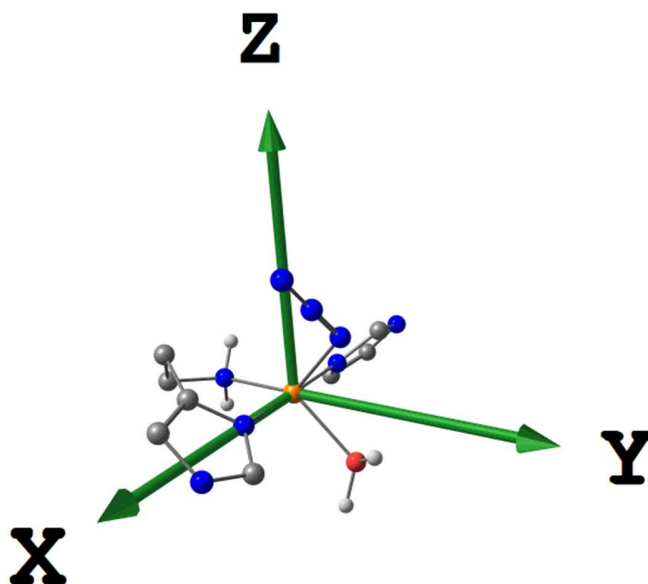
#### *EPR Property Calculations*

EPR property calculations were performed on the optimized geometries of Models A and B using the ORCA 4.2.0 program at the DFT level of theory. The cartesian reference system was oriented as such that

the His(NH<sub>2</sub>)-Cu(II) direction was aligned with the y-axis and the (His)N-Cu(II)-(His)N axis was oriented along x. The integration grid size was kept large (AngularGrid = 7 for all atoms and IntAcc = 7 for the Cu(II) ion) to ensure that the core density was correctly described. Solvation effects were accounted for in the property calculations by implementing the conductor-like polarized continuum model (CPCM) with a dielectric constant of 80.0 and a refractive index of 1.33 (water). The hyperfine coupling calculations included the Fermi-contact, spin dipolar and spin orbit contributions. The EPR properties of these models was performed using the hybrid functional, B3LYP, with an adjusted fraction of Hartree-Fock exchange (38%). A large core-properties basis set (CP(PPP)) was used to describe the central copper ion. The IGLO-III basis set was used for the ligating atoms (including all N<sub>3</sub> atoms) due to its added flexibility in the core compared to Def2-TZVP. And a reduced Def-SVP basis set was used to treat all remaining atoms. This functional and basis set regime was chosen as it has been previously shown to be successful in reproducing the experimental EPR parameters of AA10 LPMOs.<sup>4</sup>

#### *Cartesian Reference System used in EPR and TD-DFT Property Calculations*

The coordinates were translated so that the copper was the origin. The complex was rotated, so the x-axis aligned along the N(His)-Cu-N(His) bonds, the y-axis aligned with the N(NH<sub>2</sub>)-Cu bond and the z-axis perpendicular. With this reference system, the computed hyperfine tensors and g-matrices showed negligible deviation/rotation.



**Figure 16. Cartesian reference system (green) used for DFT property calculations demonstrated on the optimized coordinates for model A.**

### *Time-Dependent DFT (TD-DFT) Calculation*

Time dependent DFT calculations were performed using the optimized coordinates of model A using the ORCA 4.2.0 program. The hybrid B3LYP functional was used calculate the UV-Visible spectrum model A applying the RIJCOSX approximation (as it has been shown to minimize the computational expense of these types of calculations without incurring decreases in accuracy). The Def2-TZVP basis set and auxiliary def2/J basis set was used on all atoms. The grid size and final grid size were increased to their maxima of 5 and 6, respectively. Solvation effects were accounted for using the conductor-like polarizable continuum model, modelling water. A total of 30 roots were calculated to assess the total UV-visible spectrum. The origin of each root was assessed by looking at the parent and target orbitals. The LMCT transition was identified, and the difference density was plotted.

### *Single Point Calculations*

The optimized coordinates for models A-D provided the basis for DFT single-point calculations using the hybrid B3LYP functional as implemented by ORCA 4.2.0. The triple-zeta basis set, Def2-TZVP, was used to describe all atoms to provide greater accuracy over the previous split valence basis set. The RIJCOSX approximation was again used to improve the speed of the calculation as well as the auxiliary def2/J basis set. The conductor-like polarizable continuum model was used to account for solvation effects (water). The resulting orbital energies, compositions and overall spin populations of these calculations are discussed in the main text.

1. Serratrice, G.; Béguin, C.; Chautemps, P.; Cogne, C.; Pierre, J.-L., Biomimetic studies related to the azide-inhibited Cu,Zn superoxide dismutases. *New Journal of Chemistry* **2001**, 25 (5), 696-699.
2. Kjaergaard, C. H.; Qayyum, M. F.; Wong, S. D.; Xu, F.; Hemsworth, G. R.; Walton, D. J.; Young, N. A.; Davies, G. J.; Walton, P. H.; Johansen, K. S.; Hodgson, K. O.; Hedman, B.; Solomon, E. I., Spectroscopic and computational insight into the activation of O<sub>2</sub> by the mononuclear Cu center in polysaccharide monooxygenases. *Proceedings of the National Academy of Sciences* **2014**, 111 (24), 8797-8802.
3. Gregory, R. C.; Hemsworth, G. R.; Turkenburg, J. P.; Hart, S. J.; Walton, P. H.; Davies, G. J., Activity, stability and 3-D structure of the Cu(ii) form of a chitin-active lytic polysaccharide monooxygenase from *Bacillus amyloliquefaciens*. *Dalton Trans* **2016**, 45 (42), 16904-16912.
4. Courtade, G.; Ciano, L.; Paradisi, A.; Lindley, P. J.; Forsberg, Z.; Sørliie, M.; Wimmer, R.; Davies, G. J.; Eijssink, V. G. H.; Walton, P. H.; Aachmann, F. L., Mechanistic basis of substrate-O<sub>2</sub> coupling within a chitin-active lytic polysaccharide monooxygenase: An integrated NMR/EPR study. *Proceedings of the National Academy of Sciences* **2020**, 117 (32), 19178-19189.



## 6 Concluding Remarks

This thesis presents a collection of publications which build on the current understanding of the active site electronic structure in AA10 LPMOs. Herein, we have examined the sensitivity of the LPMO electronic structure to several factors. We have shown how adjustments to the pH leads to deprotonation of both exogenous and endogenous ligands that coordinate the copper ion, leading to marked changes in the overall electronic structure. Deprotonation of the coordinating water molecules in *B*/AA10 saw a change in ligand field from a rhombic, distorted square pyramidal geometry to an axial, bis-hydroxide 2N2O coordination geometry with a rotated SOMO. The pKa of the active site water molecule was found to be 8.7; an accessible alkalinity for biological conditions. Further increases in basicity were shown to lead to a second novel species with uncharacteristically high  $A_z$  (>610 MHz) and low  $g_z$  (2.18) spin Hamiltonian parameters compared to those already characterised in the LPMO literature. These derived SH parameters suggest that this species has the strongest ligand field oriented in the  $x/y$  plane that has been observed so far in the LPMO literature. Whilst this second species is formed outside of the common biological pH range, such findings are important in understanding the behaviour of the coordinating ligands in AA10 LPMOs.

The latter two research papers explored the fact the addition of substrate leads to an altered ligand field in chitin active AA10 LPMOs, changing from a rhombic 3N2O copper coordination geometry to an axial (near-square planar), 3N1O geometry. The changes to the electronic structure upon substrate binding have shown to be the most significant among the chitin-active AA10s when compared to the other families of LPMO, showing stark differences in their spin-Hamiltonian parameters. We used continuous wave and pulsed EPR techniques to show both the copper and nitrogen couplings are affected following substrate binding. We also revealed the formation of a hydrogen bond between the remote nitrogen of His2 and the polysaccharide N-acetyl group by evaluation of the electronic field gradient and DFT calculations.

We further explored the AA10 LPMO-substrate interaction by using the azide anion to mimic the electronic behaviour of the superoxide ion. On this basis, we prepared a range of LPMO-Cu(II)-N<sub>3</sub> complexes at the active site of a chitin-active AA10, the electronic features of which are representative of the LPMO-Cu(II)-O<sub>2</sub> species. This study highlighted the formation of two unique chemical species with high levels of metal-ligand covalency, providing valuable information on the elusive intermediate. The addition of substrate to these complexes again led to significant alterations to the ligand field. Importantly, DFT calculations showed that the 4-coordinate geometry imposed by substrate binding leads to the generation

of more reactive intermediates, lowering the barrier for HAA or formation of the LPMO-Cu(II)-O species. Together these findings are of great importance in the field of LPMOs as they are indicative of a coupling mechanism existing between substrate binding and oxygen activation occurring at the AA10 active site. Such a mechanism explains how LPMOs are able to generate such potent reactive intermediates whilst showing exquisite chemical selectivity. The proposed coupling mechanism would also suggest that AA10 LPMOs are indeed *oxygenases*, whereby O<sub>2</sub> is preferred as the oxidative co-substrate as it allows for the generation of more potent intermediates when in the presence of the correct substrate. This proposed mechanism provides the enzyme with an element of control, so that it can reduce the production of ROS when there is no substrate to act on. Such a mechanism protects the enzyme from deleterious self-oxidations that have been shown to occur during the rapid uncoupled turnover from reactions with peroxide.

## 7 Appendix of Research Papers

### 7.1 Secreted pectin monooxygenases drive plant infection by pathogenic oomycetes

The following research article was published in Science, 2021. This publication reveals the discovery of a new family of LPMO from the plant-pathogenic oomycete, *Phytophthora infestans*. This newly discovered AA17 family was found to catalyse the oxidative cleavage of polygalacturonic acid, the backbone of pectin. This new family showed many structural parallels to the other characterised families but with a notable change to the substrate binding surface. A large, charged groove was found on the AA17 substrate binding surface, showing how the new family of LPMO can recognise the negatively charged, polysaccharide substrate. This publication has great impact as it shows that *Phytophthora infestans* relies on this pectin-degrading LPMO to achieve successful pathogenesis of potato or other solanaceous plants; highlighting new potential avenues for crop protection.

This was a large body of work to which the following authors contributed to: Federico Sabbadin, Saioa Urresti, Bernard Henrissat, Anna O Avrova, Lydia R J Welsh, Peter J Lindley, Michael Csukai, Julie N Squires, Paul H Walton, Gideon J Davies, Neil C Bruce, Stephen C Whisson, Simon J McQueen-Mason. The specific contributions of each author were as follows: F.S. carried out analysis of RNA-seq data, gene cloning, heterologous protein expression and purification, enzyme activity assays, and mass spectrometry analysis of reaction products and prepared figures and tables. S.U. and G.J.D. conceived the x-ray crystallography studies. S.U. crystallized the proteins, collected and analyzed crystallographic data, solved the crystal structures, and made structural figures and tables. P.H.W. and P.J.L. conceived the EPR study. P.J.L. carried out EPR experiments and simulations. B.H. performed bioinformatics analyses and alignments. M.C., S.C.W., and J.N.S. conceived and performed the RNA-seq experiments and analyzed the data. S.C.W. and L.R.J.W. performed the stable gene silencing experiments. L.R.J.W. performed RT-qPCR. A.O.A. conceived and performed the transient gene silencing experiments. F.S., S.C.W., N.C.B., and S.J.M.-M. organized the data and wrote the manuscript. All authors contributed to production of the manuscript.

## PLANT SCIENCE

# Secreted pectin monoxygenases drive plant infection by pathogenic oomycetes

Federico Sabbadin<sup>1\*</sup>, Saioa Urresti<sup>2</sup>, Bernard Henrissat<sup>3,4,5</sup>, Anna O. Avrova<sup>6</sup>, Lydia R. J. Welsh<sup>6</sup>, Peter J. Lindley<sup>2</sup>, Michael Csukai<sup>7</sup>, Julie N. Squires<sup>8</sup>, Paul H. Walton<sup>2</sup>, Gideon J. Davies<sup>2</sup>, Neil C. Bruce<sup>1</sup>, Stephen C. Whisson<sup>6</sup>, Simon J. McQueen-Mason<sup>1\*</sup>

The oomycete *Phytophthora infestans* is a damaging crop pathogen and a model organism to study plant-pathogen interactions. We report the discovery of a family of copper-dependent lytic polysaccharide monoxygenases (LPMOs) in plant pathogenic oomycetes and its role in plant infection by *P. infestans*. We show that LPMO-encoding genes are up-regulated early during infection and that the secreted enzymes oxidatively cleave the backbone of pectin, a charged polysaccharide in the plant cell wall. The crystal structure of the most abundant of these LPMOs sheds light on its ability to recognize and degrade pectin, and silencing the encoding gene in *P. infestans* inhibits infection of potato, indicating a role in host penetration. The identification of LPMOs as virulence factors in pathogenic oomycetes opens up opportunities in crop protection and food security.

Oomycetes are fungal-like microorganisms that cause severe diseases in agriculture and aquaculture and pose a recurrent threat to global food security (1). The *Phytophthora* genus comprises more than 140 species (2), including the late blight pathogen *Phytophthora infestans*, which triggered the Irish potato famine in the mid-19th century. *P. infestans* infects both potato and tomato crops, causing economic losses in excess of \$6 billion annually (1).

The plant cell wall, which represents the first protective barrier against pathogens, is a composite material made of cellulose microfibrils embedded in a matrix of hemicellulose and pectin (3). Pectin forms an acidic polysaccharide network and makes up the scaffold of the middle lamella that binds plant cells to one another. Plant pathogens deploy cell wall-degrading enzymes to penetrate this protective barrier (4). Once the plant cell wall is breached, *P. infestans* secretes intracellular effector proteins (5) that manipulate the host and its defense mechanisms. Despite their importance, cell wall-degrading enzymes have received less attention than intracellular effectors, leaving a gap in our understanding of the mechanisms underlying infection. Here, we show that phytopathogenic oomycetes harbor a family of secreted copper-bound lytic polysaccharide monoxygenases (LPMOs) that cleave

pectin, enabling host penetration and infection by *P. infestans*.

LPMOs are copper-containing enzymes known for their ability to degrade recalcitrant and crystalline polysaccharides such as cellulose and chitin through oxidative attack (6–9). LPMOs require external electron donors in the form of small compounds (9) or redox protein partners (10) and have been characterized from fungi, bacteria, viruses, and invertebrates (9, 11), mostly in the context of plant biomass degradation.

## Identification of an expanded LPMO family in phytopathogenic oomycetes

We identified a group of proteins that have no recognized functional domains but share substantial homology and carry predicted N-terminal signal peptides for secretion. The mature proteins feature an N-terminal histidine and a further conserved histidine, potentially forming a “histidine brace” (7), the hallmark of LPMOs. Homology searches against the National Center for Biotechnology Information (NCBI) nr databases (for nonredundant protein sequences) revealed that this gene family is only found in oomycetes and is expanded in hemi-biotrophic and necrotrophic plant pathogenic species, averaging 46 family members in *Phytophthora* spp. (Fig. 1A). This LPMO family now appears as auxiliary activity 17 (AA17) in the online Carbohydrate-Active enZYmes (CAZy) database (12).

Whereas most AA17 genes code for the catalytic domain alone, sequences from *Saprolegniales* (fish pathogens) typically have putative C-terminal cellulose-binding domains (13) and putative protein-protein or protein-carbohydrate interaction domains (14), and some from *Phytophthora parasitica* feature putative chitin-binding domains (15) (fig. S1A). With few exceptions, AA17 genes in plant pathogenic oomycetes form genomic clusters surrounded by transposable

elements, similar to bacterial high-density pathogenicity islands involved in infection (16) (fig. S1B).

## AA17 genes are up-regulated during plant infection

Analysis of the published transcriptome of *Phytophthora palmicora*, a relative of *P. infestans* infecting *Nicotiana benthamiana* (17), revealed 42 AA17-coding genes, and more than half of them are induced >10-fold during infection (fig. S2 and data S1). We performed differential gene expression analysis of the transcriptome of *P. infestans* as it infected tomato plant leaves and found that AA17s were the most-induced CAZy family apparent during early infection (Fig. 1B and data S2), with PITG\_04949 being the most up-regulated (Fig. 1C).

We used reverse transcription quantitative polymerase chain reaction (RT-qPCR) to investigate the induction of LPMO-coding genes (both AA16 and AA17) during infection of potato leaves (data S3). Known transcriptional markers for developing infection (*P. infestans* putative haustorium-specific membrane protein *Pihmp1*, *P. infestans* avirulence protein 3a *PiAvr3a*, and PITG\_12808) showed up-regulation, as expected (18–20). Transcripts for six AA16 LPMOs were expressed at low levels or were undetectable. Of the 31 AA17 LPMO genes tested, the expression of 15 was detected in all replicates, and 11 exhibited greater than twofold up-regulation during infection, compared with expression in sporangia. The AA17 LPMO genes exhibiting greatest transcript abundance were PITG\_04949, 04947, 20631/20312, 01966, and 13520.

Our data show that PITG\_04949 (hereafter called *PiAA17C*) dominates in terms of gene induction in *P. infestans* infecting both tomato and potato leaves. *PiAA17C* is found in a gene cluster with three additional AA17 genes, PITG\_04947 and PITG\_04948 (henceforth called *PiAA17A* and *PiAA17B*, also induced during infection) and PITG\_04953 (not substantially expressed during infection).

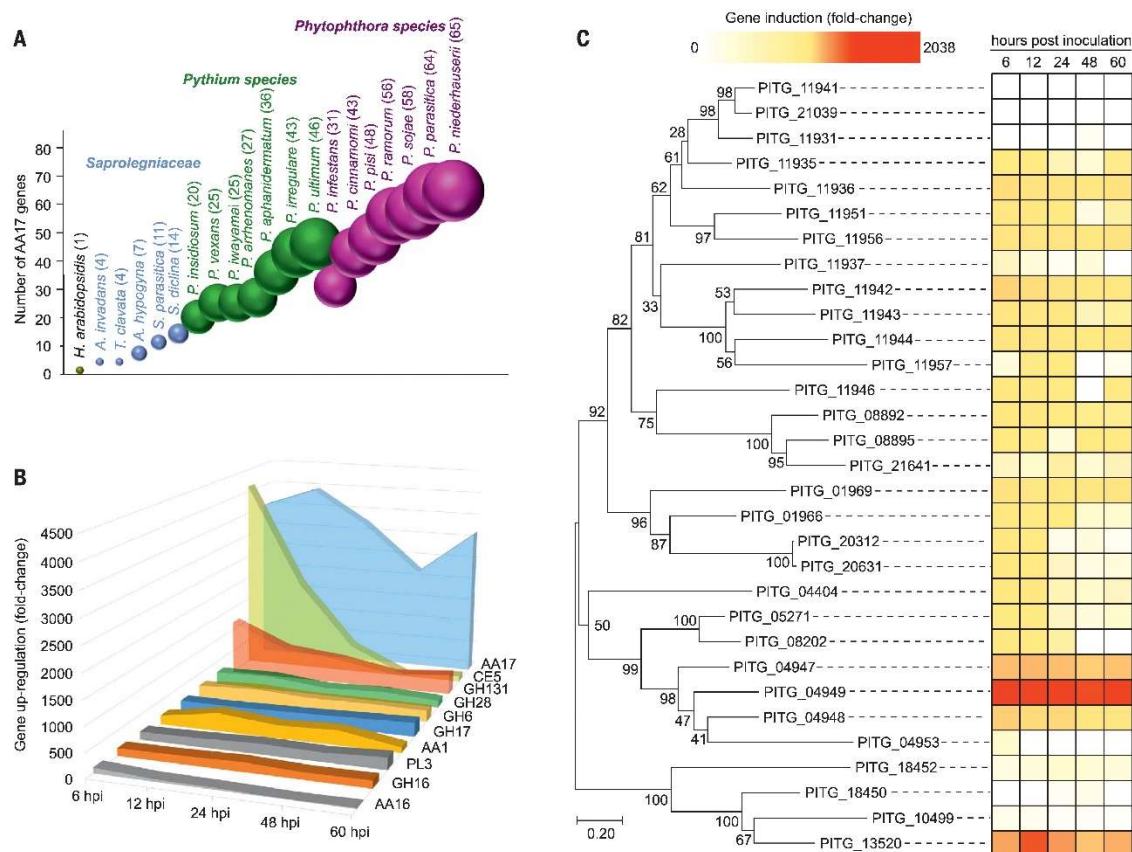
## *PiAA17C* has a canonical type-2 copper center

The predicted mature proteins that are coded by AA17 genes in *P. infestans* typically feature an N-terminal LPMO domain followed by a variable polypeptide rich in serine, proline, and threonine residues. Analysis of these C-terminal extensions using IUPred2A (intrinsically unstructured/disordered proteins and domains prediction tool) (21) predicts them to be intrinsically disordered. We cloned the LPMO domains of *PiAA17A*, -B, and -C; expressed them heterologously in *Escherichia coli*; and purified them using established methods (fig. S3A) (9). Correct protein folding and copper binding was confirmed by using thermal shift assays (fig. S3, B to D).

In vitro activity assays that used copper-loaded *PiAA17A*, -B, and -C with a range of

<sup>1</sup>Centre for Novel Agricultural Products, Department of Biology, University of York, York YO10 5DD, UK. <sup>2</sup>Department of Chemistry, University of York, York YO10 5DD, UK. <sup>3</sup>Architecture et Fonction des Macromolécules Biologiques (AFMB), UMR 7257 CNRS, Université Aix-Marseille, 163 Avenue de Luminy, 13288 Marseille, France. <sup>4</sup>INRA, USC 1408 AFMB, 13288 Marseille, France. <sup>5</sup>Department of Biological Sciences, King Abdulaziz University, Jeddah 21589, Saudi Arabia. <sup>6</sup>Cell and Molecular Sciences, James Hutton Institute, Invergowrie, Dundee DD2 5DA, UK. <sup>7</sup>Syngenta, Jealott's Hill International Research Centre, Bracknell, Berkshire RG42 6EY, UK. <sup>8</sup>Corresponding author. Email: federico.sabbadin@york.ac.uk (F.S.); simon.mcqueenmason@york.ac.uk (S.J.M.-M.)





**Fig. 1. Taxonomic distribution of AA17 LPMO genes and induction during infection of tomato leaves.** (A) Number of AA17 genes across representative oomycete species. (B) Cumulative up-regulation of CAZy families, obtained from transcriptome sequencing of *P. infestans* infecting tomato leaves at 6, 12, 24, 48, and 60 hours postinoculation (hpi), relative to sporangia. (C) Neighbor joining phylogenetic tree of *P. infestans* AA17

LPMOs (predicted catalytic domains only) and gene induction levels on the basis of RNA sequencing (RNA-seq) data of *P. infestans* infecting tomato leaves during the time course experiment (6 to 60 hpi). The maximum fold change (2038) corresponds to gene PITG\_04949 at 60 hpi. Bootstrap values (calculated using Mega7 with 1000 cycles) are shown at branching points.

polysaccharides and reducing agents (ascorbic acid, gallic acid, and pyrogallol), followed by analysis by matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF MS) and electrospray ionization mass spectrometry (ESI MS), failed to reveal any products released from the substrates tested, including those oxidized by characterized LPMOs (cellulose, cello-oligosaccharides, chitin, xylan, xyloglucan, and glucomannan) (6–9, 11, 22).

Despite the lack of activity by *PiAA17C* on these substrates, electron paramagnetic resonance spectroscopy (EPR) revealed a spectral envelope consistent with a type-2 copper center with near axial symmetry ( $g_z > g_y \approx g_x > g_w$ , where the  $g$  value is the proportionality factor) (Fig. 2A and table S1) (23), confirming that *PiAA17C* is an LPMO with a canonical histidine brace active site (7). The spin-Hamiltonian

values are similar to those obtained for fungal AA9 LPMOs (22), indicating active high-valent copper-oxygen intermediates (e.g., copper-oxyl) as part of their catalytic cycle (23). Visible superhyperfine coupling is resolved in the perpendicular region of the spectrum and is attributed to coupling to the nitrogen nuclei in the  $xy$  plane, which ranges from 34 to 38 MHz, typical for an LPMO (24).

#### The structure of *PiAA17C* suggests an interaction with charged polysaccharides

To shed light on the potential substrate specificity and activity of *PiAA17C*, we solved its x-ray crystal structure to 1.01-Å resolution. Structural similarity searches using Dali (25), which compares the C-alpha of the query structure with others available in the Protein Data Bank (PDB), gave the best match with chitin-active

*CjLPMO10A* from *Cellvibrion japonicus* [PDB ID: 5FJQ(A)], although the overall similarity is low (Dali  $z$ -score 9.5, 11% identity, 2.7-Å root mean square deviation over 118 aligned C-alphas). Akin to other LPMO families, the *PiAA17C* structure reveals a central  $\beta$ -sandwich fold decorated with several loops and stabilized by three disulfide bonds (Fig. 2B and table S2). The active site features a histidine brace (7) formed by His<sup>1</sup> and His<sup>96</sup>, which is required for copper coordination. The electrostatic surface potential and residue charge distribution of *PiAA17C* are notably distinct from those of typical LPMOs active on crystalline cellulose or chitin, which display a flat surface surrounding the active site. By contrast, the His-brace of *PiAA17C* lies within a cleft (Fig. 2C). Sequence conservation analysis (Fig. 2D) (26) highlighted conservation at the active site

as well as several polar and negatively charged residues across the AA17 family, which suggests the ability to bind charged polysaccharides (Fig. 2E). The surface surrounding the active site also lacks the aromatic residues required for the recognition of uncharged polysaccharides in canonical LPMOs (24). Another unusual feature of *PiAA17C*'s structure is a large  $\alpha$  helix (residues 36 to 51) (fig. S4) with a negatively charged groove engraved beneath it, leading to the active site. These negatively

charged residues are reminiscent of the aspartate side chains involved in calcium coordination in homogalacturonate-active enzymes (27–29), and this prompted us to test our purified enzymes against negatively charged substrates.

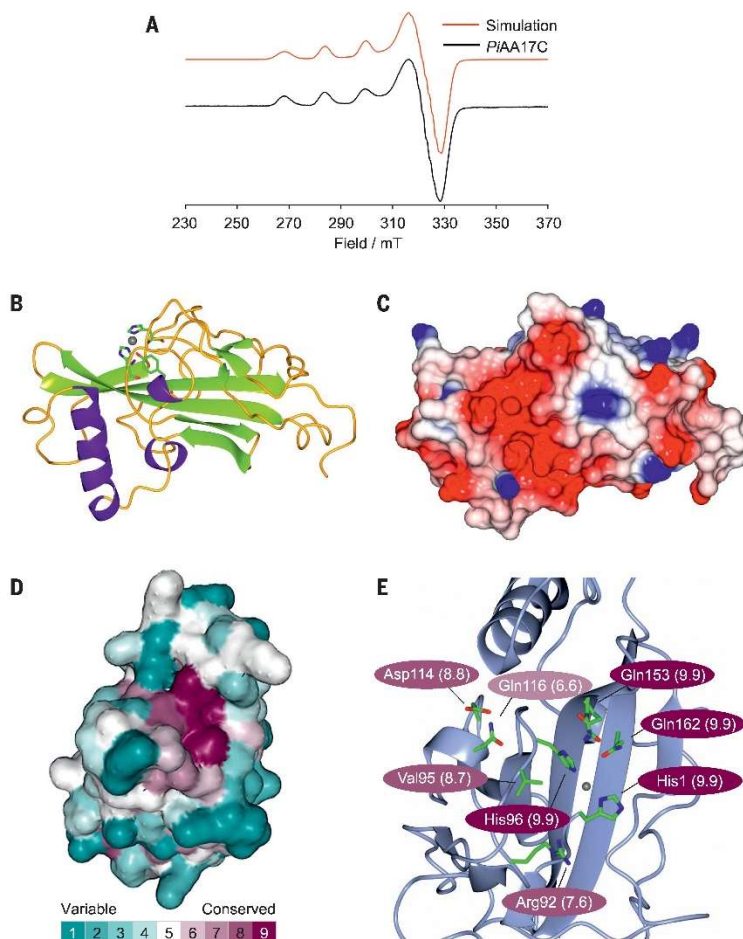
#### AA17 LPMOs oxidatively cleave homogalacturonan

*PiAA17A*, -B, and -C were incubated with a panel of charged polysaccharides in the presence of a range of reducing agents (ascorbic

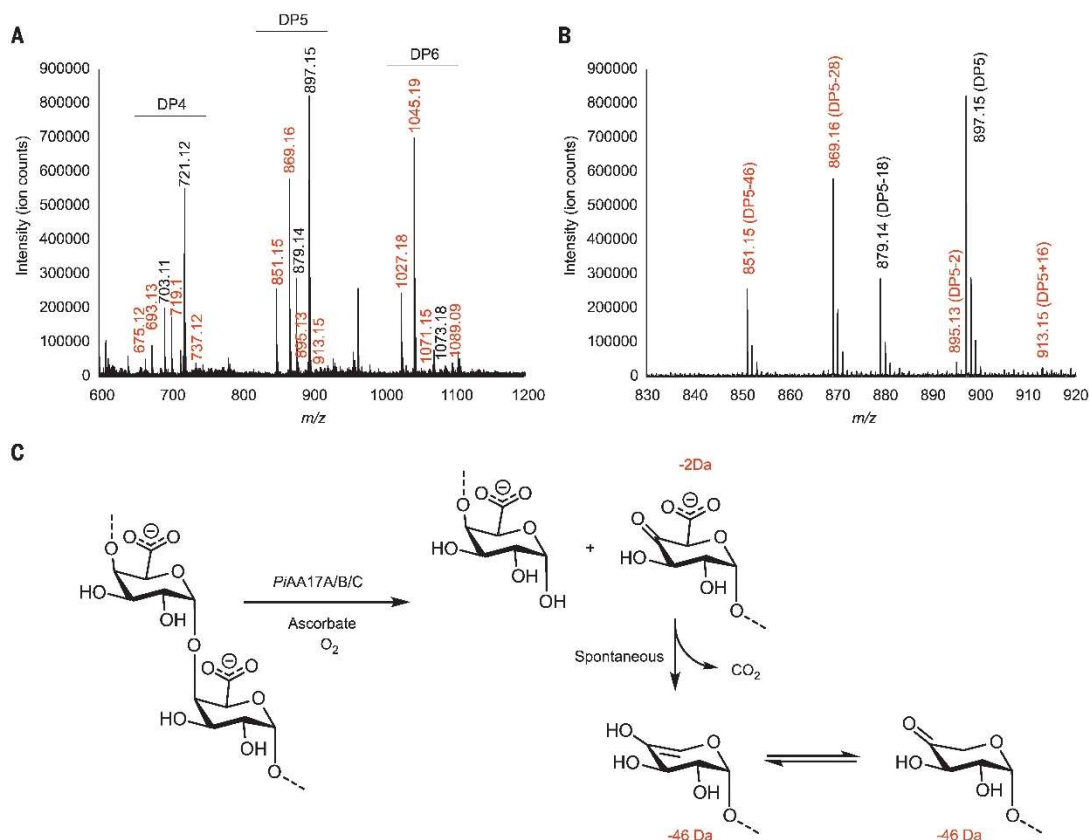
acid, gallic acid, and pyrogallol). MALDI-TOF MS and ESI MS analysis of the released products revealed distinct oxidized and native product peaks when using homogalacturonan (Fig. 3, A and B, and fig. S5, D, F, and H) and oligogalacturonides [degree of polymerization (DP) 10 to 15] (fig. S6, D, F, and H) as substrates. Homogalacturonan forms the backbone of pectin and consists of a linear chain of (1,4)-linked  $\alpha$ -D-galacturonic acid units with variable degrees of methyl esterification (3). Activity on charged polysaccharides has not previously been reported for LPMOs. We observed that AA17s can accept electrons from ascorbic acid, but not from small phenolic compounds, whereas both types of molecules were used successfully with published LPMOs (9). Treatment of the released products with exo-polygalacturonase (GH28) acting on the nonreducing end (C4) degraded native oligogalacturonides, whereas peaks corresponding to the putative oxidized species were not altered (fig. S7), suggesting that *PiAA17C* predominantly carries out C4-specific oxidation of homogalacturonan and that the resulting C4-oxidized oligogalacturonides are not accessible to the C4-acting exo-polygalacturonase. C4 oxidation also best explains the identity of two species (-46 and -28) not previously observed with characterized LPMOs. On the basis of peak masses from both MALDI-TOF MS and ESI MS analyses, we propose that *PiAA17A*, -B, and -C carry out a C4-oxidative cleavage of polygalacturonic acid, generating a C4-ketone in a  $\beta$  position relative to one carboxylic group and resulting in an unstable  $\beta$ -keto acid that undergoes spontaneous decarboxylation and tautomerization (Fig. 3C). This mechanism is analogous to that proposed for uridine 5'-diphosphate (UDP)-glucuronic acid decarboxylase, for which the oxidation of UDP glucuronic acid in the C4 position generates a ketone that undergoes decarboxylation with formation of a UDP-4-keto pentose (30).

#### AA17 LPMOs recognize the carboxylic groups of de-esterified pectin

We found that *PiAA17C* was not active on esterified citrus pectin (fig. S8, E and F). However, preincubation of esterified pectin with pectin methyltransferase, followed by addition of *PiAA17C*, resulted in substrate degradation (fig. S8H), suggesting that carboxylic groups exposed through de-esterification are important for substrate recognition and cleavage by the LPMO. In addition, thermal shift analysis of *PiAA17A*, -B, and -C revealed a specific interaction with homogalacturonan and oligogalacturonides, whereas interaction with highly esterified pectin was only detected after demethylation of the substrate (fig. S9 and table S3). The role of AA17 LPMOs in pectin degradation is supported by their co-induction with several putative pectin methyltransferases (from



**Fig. 2. Structural and spectroscopic characterization of *PiAA17C*.** (A) Continuous-wave X-band EPR spectrum of *PiAA17C* (~0.5 mM) in sodium phosphate buffer (pH 7.0, 20 mM) collected at 150 K (black) and spectral simulation (red). (B) Overall structure of *PiAA17C*, showing the antiparallel  $\beta$  sheet (in green) and histidine brace (sticks), featured in all LPMOs, and the long  $\alpha$  helix (in purple), which is not present in other LPMO families (see fig. S4). (C) Electrostatic surface potential of *PiAA17C*, showing the cleft surrounding the active site and the negatively charged groove (red) leading toward it. (D) Sequence conservation analysis (ConSurf, based on 401 AA17 sequences) of *PiAA17C* looking down on the active site. The surface is colored by ConSurf score according to the indicated scoring scheme. (E) Highly conserved residues around the active site of *PiAA17C*, based on ConSurf analysis of 401 AA17 sequences (LPMO domain only). Color shades indicate the level of conservation, calculated using ConSurf.



**Fig. 3. Biochemical characterization of *PiAA17C* and proposed mechanism of action.** (A) Negative-mode ESI MS spectrum of products released by 2  $\mu\text{M}$  *PiAA17C* from 4  $\text{mg ml}^{-1}$  homogalacturonan in the presence of 4 mM ascorbic acid after 24-hour incubation (see materials and methods). Masses are indicated by numbers for the main peaks corresponding to native (black) and oxidized (red) oligogalacturonides. Control reactions with substrate only, substrate plus ascorbic acid, and substrate plus *PiAA17C* did not generate detectable amounts of oligogalacturonides (see fig. S5).  $m/z$ , mass/charge ratio. (B) Expanded ESI MS spectra for DP5, showing the main peaks and their identity.  $m/z$  897.15: native oligogalacturonide.  $m/z$  895.13: -2 species, oxidized (ketone).  $m/z$  913.15: +16 species, oxidized (gemdiol).  $m/z$  879.14: -18 species,

dehydrated native oligogalacturonide;  $m/z$  851.15: -46 species, oxidized + decarboxylated oligogalacturonide (C4 ketone);  $m/z$  869.16: -28 species, oxidized + decarboxylated oligogalacturonide (hydrated version of the C4 ketone, corresponding to 851.16 + 18 mass units, a gemdiol). The keto sugar is in equilibrium with the C4 gemdiol in aqueous solution, a feature common to keto saccharides, including C4 ketones generate by characterized LPMOs (40). (C) Proposed mechanism of action for *PiAA17A*, -B, and -C acting on polygalacturonic acid. In the presence of ascorbic acid, the LPMO carries out a C4 oxidation, leading to the formation of a ketone in C4, followed by spontaneous decarboxylation of the resulting  $\beta$ -keto acid. The C4-enol undergoes tautomerization and formation of the more stable ketone form.

families CE8 and CE13), as well as other pectin-active enzymes (families GH28, PL3, and PL4) in our transcriptomic studies (data S2) and published data from other *Phytophthora* species (data S1) (17, 31).

#### *PiAA17C* is necessary for successful plant infection by *P. infestans*

We assessed the role of *PiAA17C* in plant infection through transient gene silencing by delivering in vitro synthesized double-stranded RNA (dsRNA) into protoplasts of *P. infestans* (32), followed by colony regeneration and infection of potato leaves. Infection phenotypes recorded 72 to 96 hours postinoculation showed that the introduction of dsRNA

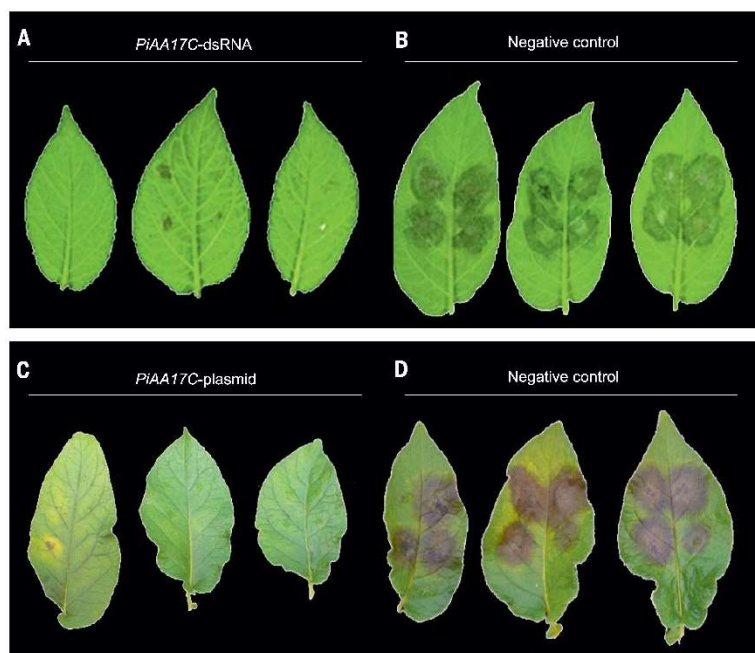
targeting *PiAA17C* reduced virulence compared with control lines (Fig. 4, A and B). We verified this result using stable gene silencing in which *PiAA17C* was transformed into *P. infestans* as an inverted repeat (33), resulting in the removal of expression for the gene through the formation of heterochromatin. Six genocin-resistant *P. infestans* lines that exhibited silencing of *PiAA17C* were selected (data S4) and showed a marked reduction in lesion size upon infection of potato leaves (Fig. 4, C and D, and fig. S10).

As seen previously for other genes silenced in *P. infestans* (34), the expression of nearby genes was also affected by the silencing of *PiAA17C*, with the greatest effect on the closest gene *PiAA17B* and a lesser impact on the more

distant *PiAA17A* (data S4). Transcripts of *PiAA17C* are the most abundant in *P. infestans* during infection, and lesion sizes caused by the transgenic lines closely mirror the level of silencing of *PiAA17C* (fig. S10), whereas the differing levels of expression of *PiAA17A* and *PiAA17B* between replicates effectively rule them out as major contributors to the observed effect, validating the results of the transient gene silencing experiment.

#### Discussion

Our results suggest that the evolutionary arms race with plants has spurred oomycetes to evolve LPMOs as virulence factors to overcome host defenses. Although AA17s likely have



**Fig. 4.** In vivo *PiAA17C* gene silencing during infection of *P. infestans* on plant leaves. (A to D) Transient (A and B) and stable (C and D) silencing of *PiAA17C*. (A) Potato leaves infected with *P. infestans* isolate 88069 treated with dsRNA targeting *PiAA17C*. (B) Leaves infected with *P. infestans* isolate 88069 treated with nonhomologous dsRNA (negative control). (C) Leaves treated with *P. infestans* isolate 3928A stably transformed with silencing plasmid targeting *PiAA17C* (line IR6). (D) Leaves treated with wild-type *P. infestans* (negative control, isolate 3928A).

a role in facilitating host penetration by cleaving pectin and disrupting the plant cell wall network, we speculate that they further enhance *P. infestans* virulence by interfering with host immunity. Oligogalacturonides released by pathogen polygalacturonase enzymes are well-characterized inducers of plant responses to pathogens, and the oxidation of oligogalacturonides by berberine bridge enzyme-like proteins has been shown to hinder their recognition by the plant, which prevents the activation of immune responses (35). AA17s generate oxidized and decarboxylated oligogalacturonides while degrading the pectin backbone and might play a role in avoiding the plant immune response while simultaneously breaching the host cell wall.

Transcriptomic analyses have shown up-regulation of some AA9-coding genes during fungal invasion of plant tissue (36), which suggests a role in pathogenesis. Recently, a marked induction of immunity-related genes was observed in *Arabidopsis thaliana* treated with native and oxidized cellulose oligosaccharides produced by a fungal AA9 LPMO (37). The different physicochemical properties of charged oligogalacturonides released by berberine bridge enzyme-like proteins and AA17s, compared

with uncharged cello-oligosaccharides released by AA9s, may hold the key to the two seemingly contrasting effects on the host immune system.

The amount of ascorbic acid that we found to be effective in our assays (1 to 4 mM) is comparable to physiological levels measured in potato and tomato plants (38, 39). The observation that *PiAA17A*, -B, and -C activity is driven by ascorbic acid but not phenolic compounds may reflect an adaptation of the *P. infestans* life cycle, in which it invades fresh host tissue (rich in cellular reductants, including ascorbic acid) rather than lignified tissue (a plentiful source of phenolics, compatible with characterized LPMO families from wood-decay fungi).

Our results shed light on the complex interactions between hosts and pathogens in the plant cell wall and illustrate how targeting LPMO genes through RNA-based approaches could provide a strategy to fight crop diseases and increase agricultural productivity.

#### REFERENCES AND NOTES

1. L. Derevnina et al., *Philos. Trans. R. Soc. B* **371**, 20150459 (2016).
2. X. Yang, B. M. Tjyer, C. Hong, *IMA Fungus* **8**, 355–384 (2017).
3. R. A. Burton, M. J. Gidley, G. B. Fincher, *Nat. Chem. Biol.* **6**, 724–732 (2010).
4. C. P. Kubicek, T. L. Starr, N. L. Glass, *Annu. Rev. Phytopathol.* **52**, 427–451 (2014).
5. S. C. Whisson et al., *Nature* **450**, 115–118 (2007).

6. G. Vaaje-Kolstad et al., *Science* **330**, 219–222 (2010).
7. R. J. Quinlan et al., *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15079–15084 (2011).
8. G. R. Hemsworth et al., *J. Am. Chem. Soc.* **135**, 6069–6077 (2013).
9. F. Sabbadin et al., *Nat. Commun.* **9**, 756 (2018).
10. T. C. Tan et al., *Nat. Commun.* **6**, 7542 (2015).
11. Z. Forsberg et al., *Curr. Opin. Struct. Biol.* **59**, 54–64 (2019).
12. V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho, B. Henrissat, *Nucleic Acids Res.* **42**, D490–D495 (2014).
13. A. Vámai et al., *Adv. Appl. Microbiol.* **88**, 103–165 (2014).
14. H. Tordai, L. Bányai, L. Patthy, *FEBS Lett.* **461**, 63–67 (1999).
15. T. Suetake et al., *J. Biol. Chem.* **275**, 17929–17932 (2000).
16. O. Gal-Mor, B. B. Finlay, *Cell. Microbiol.* **8**, 1707–1719 (2006).
17. E. Evangelisti et al., *BMC Biol.* **15**, 39 (2017).
18. A. O. Avrova et al., *Cell. Microbiol.* **10**, 2271–2284 (2008).
19. M. R. Armstrong et al., *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7766–7771 (2005).
20. M. Abrahamian, A. M. V. Ah-Fong, C. Davis, K. Ancreeva, H. S. Judelson, *PLOS Pathog.* **12**, e1006097 (2016).
21. B. Mészáros, G. Erdős, Z. Dosztányi, *Nucleic Acids Res.* **46**, W329–W337 (2018).
22. T. J. Simmons et al., *Nat. Commun.* **8**, 1064 (2017).
23. G. R. Hemsworth, L. Clano, G. J. Davies, P. H. Walton, *Methods Enzymol.* **613**, 63–90 (2018).
24. K. E. H. Frandsen et al., *Nat. Chem. Biol.* **12**, 298–303 (2016).
25. L. Holm, P. Rosenström, *Nucleic Acids Res.* **38**, W545–W549 (2010).
26. H. Ashkenazy, E. Erez, E. Martz, T. Pupko, N. Ben-Tal, *Nucleic Acids Res.* **38**, W529–W533 (2010).
27. S. J. Charnock, I. E. Brown, J. P. Turkenburg, G. W. Black, G. J. Davies, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12067–12072 (2002).
28. M. D. Yoder, N. T. Keen, F. Jurnak, *Science* **260**, 1503–1507 (1993).
29. R. Pickersgill, J. Jenkins, G. Harris, W. Nasser, J. Robert-Baudouy, *Nat. Struct. Biol.* **1**, 717–723 (1994).
30. M. Bar-Peled, C. L. Griffith, T. L. Doering, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12003–12008 (2001).
31. L. M. Blackman, D. P. Cullerne, P. Torreña, J. Taylor, A. R. Hardham, *PLOS ONE* **10**, e0136899 (2015).
32. S. C. Whisson, A. O. Avrova, P. van West, J. T. Jones, *Mol. Plant Pathol.* **6**, 153–163 (2005).
33. H. S. Judelson, S. Tani, *Eukaryot. Cell* **6**, 1200–1209 (2007).
34. A. L. Vu, W. Leesutthiphonchai, A. M. V. Ah-Fong, H. S. Judelson, *Mol. Plant Microbe Interact.* **32**, 915–927 (2019).
35. M. Benedetti et al., *Plant J.* **94**, 260–273 (2018).
36. G. Jagadeeswaran, L. Veale, A. J. Mort, *Trends Plant Sci.* **26**, 142–155 (2021).
37. M. Zarattini et al., *Commun. Biol.* **4**, 727 (2021).
38. J. S. Han, N. Kozukue, K. S. Young, K. R. Lee, M. Friedman, *J. Agric. Food Chem.* **52**, 6516–6521 (2004).
39. H. Gaultier, C. Massot, R. Stevens, S. Sério, M. Génard, *Ann. Bot.* **103**, 495–504 (2009).
40. T. Isaksen et al., *J. Biol. Chem.* **289**, 2632–2642 (2014).
41. P. H. Walton, P. Lindley, F. Sabbadin, G. J. Davies, EPR data for “Secreted pectin monoxygenases drive plant infection by pathogenic oomycetes.” York Research Database (2021); doi: 10.15124/0ee2a9c1-6d3b-4ee0-8ac2-c19061a40d94.

#### ACKNOWLEDGMENTS

We thank J. Agirre for his help with comparisons of *PiAA17C* with all available LPMO structures. We thank Diamond Light Source for access to beamline I04 (proposal number mx-9948), which contributed to the results presented here. **Funding:** This work was funded by the UK Biotechnology and Biological Sciences Research Council (grants BB/L001926/1, BB/J016500/1, BB/L021633/1, BB/V000365/1, and BB/V000675/1). The York Centre of Excellence in Mass Spectrometry was created with funds from Science City York, Yorkshire Forward, and the Northern Way Initiative and by EPSRC (EP/K039660/1; EP/M028127/1), A.O.A., S.C.W., L.R.J.W., and J.N.S. acknowledge Syngenta and the Scottish Government Rural and Environment Science and Analytical Services Division. **Author contributions:** F.S. carried out analysis of RNA-seq data, gene cloning, heterologous protein expression and purification, enzyme activity assays, and mass spectrometry analysis of reaction products and prepared figures and tables. S.U. and G.J.D. conceived the x-ray crystallography studies. S.U. crystallized the proteins, collected and analyzed crystallographic data, solved the crystal structures, and made structural figures and tables. P.H.W. and P.J.L. conceived the EPR study. P.J.L. carried out EPR experiments and simulations. B.H. performed bioinformatics analyses and alignments. M.C., S.C.W., and J.N.S. conceived and performed the RNA-seq experiments and analyzed the data. S.C.W. and L.R.J.W.

performed the stable gene silencing experiments. L.R.J.W. performed RT-qPCR. A.O.A. conceived and performed the transient gene silencing experiments. F.S., S.C.W., N.C.B., and S.J.M.-M. organized the data and wrote the manuscript. All authors contributed to production of the manuscript. **Competing interests:** The authors declare no competing interests. B.H. is now affiliated with the Technical University of Denmark. **Data and materials availability:** Coordinates and structure factors for the x-ray structure of P1AA17C were deposited in the PDB under accession

code 6Z5Y. Raw EPR data are available through the York Research Database (41). *P. infestans* raw RNA-seq data are available in the NCBI Sequence Read Archive under BioProject PRJNA739688, accession numbers SRR14871460 to SRR14871482. All other data are in the main paper or supplementary materials.

**SUPPLEMENTARY MATERIALS**

[science.sciencemag.org/content/373/6556/774/suppl/DC1](https://science.sciencemag.org/content/373/6556/774/suppl/DC1)  
Materials and Methods

Figs. S1 to S10  
Tables S1 to S3  
References (42–58)  
MDAR Reproducibility Checklist  
Data S1 to S4

[View/request a protocol for this paper from \*Bio-protocol\*.](#)

22 April 2021; accepted 6 July 2021  
10.1126/science.abj1342

## 7.2 C-type cytochrome-initiated reduction of bacterial lytic polysaccharide monooxygenases

The following research article was published in *Biochemical Journal*, 2021. This publication explores the ability of C-type cytochromes to act as the reducing co-partner for bacterial LPMO catalysis. The nature of the reductive partner which facilitates the Cu<sup>II/I</sup> redox couple in LPMO catalysis is one of the important questions still under debate. Herein, we explore the possibility of a small c-type cytochrome (CjX183) as the potential redox partner for AA10 LPMOs in the deconstruction of lignocellulosic materials. This c-type cytochrome showed to be effective in activating two LPMOs towards cellulose oxidation. Moreover, the use of the cytochrome as the reductive partner resulted in the generation of less H<sub>2</sub>O<sub>2</sub>, and therefore less oxidative damage to the enzyme, than the corresponding reactions with ascorbate as the reducing agent. This article provides new insight into the redox processes that occur in AA10 LPMO biochemistry.

This article was prepared by several co-authors to which I made a minor contribution. My specific contribution to this study was performing the MALDI-TOF/TOF mass spectrometry studies to identify the presence or absence of oxidised products from the polysaccharide activity assays.

The contributions of each author are as follows:

Glyn R. Hemsworth: Conceptualization, Resources, Data curation, Formal analysis, Supervision, Funding acquisition, Validation, Investigation, Visualization, Methodology, Project administration, Writing — review and editing. Jessie Branch: Conceptualization, Resources, Data curation, Formal analysis, Supervision, Funding acquisition, Validation, Investigation, Visualization, Methodology, Writing — original draft, Project administration, Writing — review and editing. Badri S. Rajagopal: Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing — original draft, Writing — review and editing. Alessandro Paradisi: Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing — review and editing. Nick Yates: Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing — review and editing. Peter J. Lindley: Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing — review and editing. Jake Smith: Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing — review and editing. Kristian Hollingsworth: Conceptualization, Resources, Data curation, Formal analysis, Supervision, Funding acquisition, Validation, Investigation, Visualization, Methodology, Project administration, Writing — review and editing. Bruce Turnbull: Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing — review and editing. Bernard Henrissat: Conceptualization, Data curation, Software, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing — review and editing. Alison Parkin: Conceptualization, Data curation, Software, Formal analysis, Supervision, Validation, Investigation, Visualization, Methodology, Writing — review and editing. Alan berry: Conceptualization, Data curation, Formal analysis, Supervision, Validation, Investigation, Visualization, Methodology, Writing — review and editing.

Research Article

# C-type cytochrome-initiated reduction of bacterial lytic polysaccharide monoxygenases

Jessie Branch<sup>1</sup>, Badri S. Rajagopal<sup>1</sup>, Alessandro Paradisi<sup>2</sup>, Nick Yates<sup>2</sup>, Peter J. Lindley<sup>2</sup>, Jake Smith<sup>2</sup>, Kristian Hollingsworth<sup>3</sup>, W. Bruce Turnbull<sup>3</sup>, Bernard Henrissat<sup>4,5,6</sup>, Alison Parkin<sup>2</sup>, Alan Berry<sup>1</sup> and Glyn R. Hemsworth<sup>1</sup>

<sup>1</sup>Astbury Centre for Structural Molecular Biology and School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, U.K.; <sup>2</sup>Department of Chemistry, University of York, York YO10 5DD, U.K.; <sup>3</sup>School of Chemistry and Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, U.K.; <sup>4</sup>Architecture et Fonction des Macromolécules Biologiques (AFMB), CNRS, Aix-Marseille Université, Marseille, France; <sup>5</sup>AFMB, INRA, USC 1408, Marseille, France; <sup>6</sup>Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

Correspondence: Glyn R. Hemsworth (g.r.hemsworth@leeds.ac.uk)



The release of glucose from lignocellulosic waste for subsequent fermentation into biofuels holds promise for securing humankind's future energy needs. The discovery of a set of copper-dependent enzymes known as lytic polysaccharide monoxygenases (LPMOs) has galvanised new research in this area. LPMOs act by oxidatively introducing chain breaks into cellulose and other polysaccharides, boosting the ability of cellulases to act on the substrate. Although several proteins have been implicated as electron sources in fungal LPMO biochemistry, no equivalent bacterial LPMO electron donors have been previously identified, although the proteins Cbp2D and E from *Cellvibrio japonicus* have been implicated as potential candidates. Here we analyse a small c-type cytochrome (CjX183) present in *Cellvibrio japonicus* Cbp2D, and show that it can initiate bacterial Cu<sup>II/I</sup> LPMO reduction and also activate LPMO-catalyzed cellulose-degradation. In the absence of cellulose, CjX183-driven reduction of the LPMO results in less H<sub>2</sub>O<sub>2</sub> production from O<sub>2</sub>, and correspondingly less oxidative damage to the enzyme than when ascorbate is used as the reducing agent. Significantly, using CjX183 as the activator maintained similar cellulase boosting levels relative to the use of an equivalent amount of ascorbate. Our results therefore add further evidence to the impact that the choice of electron source can have on LPMO action. Furthermore, the study of Cbp2D and other similar proteins may yet reveal new insight into the redox processes governing polysaccharide degradation in bacteria.

## Introduction

The production of second-generation biofuels from lignocellulosic waste in a cost-effective manner is a much-sought-after goal of the industrial biotechnology sector [1,2]. One of the major hurdles to overcome has been the inherent difficulty in liberating individual glucose units from cellulose, a polymer of β-1,4-glucose, for subsequent fermentation into bioethanol. This is a difficult obstacle to overcome because the polysaccharide substrate is highly recalcitrant. As such, efforts to use cellulose as a feedstock for industrial bioethanol production often faltered [3]. However, the discovery of lytic polysaccharide monoxygenases (LPMOs) has caused considerable renewed interest in this area because LPMOs boost the activity of cellulase cocktails, making cellulose deconstruction significantly more cost effective [4–9].

LPMOs are widespread and have been discovered in all kingdoms of life. Based on their protein sequences, these enzymes are classified into seven distinct families in the CAZy database of carbohydrate-active enzymes (www.cazy.org) — auxiliary activity (AA) families AA9 to AA11 and AA13 to AA16 [6–17]. Much of the research on LPMOs has focussed on understanding how these

Received: 25 May 2021  
Revised: 6 July 2021  
Accepted: 9 July 2021

Accepted Manuscript online:  
9 July 2021  
Version of Record published:  
28 July 2021

enzymes bring about the observed boosting of cellulase activity. The molecular basis for this is thought to result from their ability to introduce chain breaks into the crystalline regions of cellulose, thereby introducing new points at which cellulases can act ([7,9,18–21] and see [22] for a recent review). LPMOs do this by reductively activating oxygen or hydrogen peroxide to specifically attack at either the C1 or C4 carbon of the glucose ring resulting in hydroxylation at that position and subsequent breakage of the glycosidic bond [7,9,18–20,23]. The precise mechanism of reaction that these enzymes use is a hotly debated topic, which continues to fuel much of the research on these enzymes at present (see [22,24,25] for recent reviews).

Though LPMO families share little sequence identity, they are united by the presence of the ‘histidine brace’ motif [7,26–29]. This motif coordinates a single copper ion in the active site which is essential for activity. Probing how this mononuclear metal site can activate dioxygen has been the key driver behind a number of recent mechanistic studies [18,27,28,30–34]. It was generally agreed that LPMOs utilise either an O<sub>2</sub>-derived Cu(II)-superoxide or a Cu(II)-oxyl intermediate to oxidise their substrates, with the generation of the latter requiring multiple electrons in order to cleave the O–O bond (see reviews [22,35,36]). Recent research, however, has called into question the physiological relevance of these mechanisms by proposing that H<sub>2</sub>O<sub>2</sub> (the redox-state equivalent of O<sub>2</sub> + 2H<sup>+</sup> + 2e<sup>−</sup>) may be the true co-substrate for LPMOs [23,37–39]. Such proposals were originally based on an observed increased rate of LPMO reaction when using H<sub>2</sub>O<sub>2</sub> for activity as opposed to O<sub>2</sub> [23,37–39]. Further support for this proposal has also come from computational studies which suggest that using H<sub>2</sub>O<sub>2</sub> as a co-substrate is more energetically favourable than using O<sub>2</sub> [40–43]. Whether H<sub>2</sub>O<sub>2</sub> is the true co-substrate, or a ‘reaction-shunt’ (i.e. H<sub>2</sub>O<sub>2</sub> provides a shortcut to forming the same key reactive intermediate required in an O<sub>2</sub> and electron dependent reaction) is difficult to dissect, not least because hydrogen peroxide is known to cause significant oxidative damage to LPMOs [23,37,44,45], and also because lab-based studies can be misleading in terms of the multiple components that might be involved in activating LPMOs [46].

Whether O<sub>2</sub> or H<sub>2</sub>O<sub>2</sub> is the co-substrate, LPMOs require a source of electrons for activity in order to induce the reduction of the active site Cu<sup>II</sup> to Cu<sup>I</sup>. In laboratory experiments, small molecule reducing agents are typically used to do this and trigger LPMO action. In fungal LPMO systems, several enzymes including cellobiose dehydrogenase (CDH) [47–50], glucose-methanol-choline (GMC) oxidoreductases [50,51] and AA12 pyrrolo-quinoline quinone (PQQ)-dependent pyranose dehydrogenases [52], have also been demonstrated to donate electrons to LPMOs thereby activating them for action. Since CDH and AA12 enzymes have an appended *b*-type cytochrome domain (AA8) which mediates the electron transfer to the LPMO, a direct interaction between enzymes must occur [52,53]. *In silico* docking [52,53] and protein–protein interaction studies [54] so far suggest that these enzyme partners directly contact the copper ion to allow its reduction, but the transient nature of the interaction and the difficulty in assessing these interactions in the presence of the substrate, has made it challenging to characterise this further. Indeed, several studies have proposed the presence of electron hopping pathways through the core of the LPMO [6,30,32,55,56], which may be of particular relevance in cases where an LPMO activating enzyme partner has been invoked as the electron source. There are many open questions that remain to be addressed in this area, but what is clear is that the activity of LPMOs is only fully revealed when they act as part of a consortium of enzymes, not in isolation. Determining how the individual components of the consortium interact and what their roles are is perhaps the most important question facing the LPMO field, and whether such systems also operate outside of the fungal kingdom is a key outstanding question to be addressed.

While a native bacterial LPMO protein partner has yet to be identified, two bacterial proteins have been suggested as candidates, these are Cbp2D and Cbp2E from *Cellvibrio japonicus* [57]. When Gardner and co-workers knocked out the genes coding for these proteins, the growth of the resulting *C. japonicus* strain was significantly retarded on cellulose filter paper, suggesting they played a prominent role in cellulose metabolism. Additionally, both proteins have been detected in secretomic studies during growth on chitin, suggesting a potential role in the general polysaccharide metabolism of this organism [58]. Both Cbp2D and E contain a carbohydrate-binding module (CBM2) and a YceI-like ubiquinone-8 binding domain, designated an X158 domain, as described by Vincent *et al.* [57,59]. Ubiquinone-8 is well known for its role in the electron transport chain hinting that proteins harbouring this domain may have a redox function [59]. Additionally, Cbp2D also harbours two predicted *c*-type cytochrome domains at its C-terminus suggesting an electron transfer function for this protein, but direct links between these proteins to LPMO biochemistry have not yet been demonstrated experimentally [57].



We therefore set out to characterise Cbp2D structurally and biochemically to better understand its likely function. Here, we present the structure of a domain from Cbp2D, which we dub *CjX183*, confirming that it is indeed a small *c*-type cytochrome domain, with an exposed heme cofactor. We further show that the isolated domain is redox active and can activate two bacterial LPMOs for oxidative action on cellulose. By comparing the results of LPMO activity using *CjX183* as a reductant to those using ascorbate, we highlight differences that may relate to how a small molecule may act compared with a protein-based electron source. Our results therefore highlight the need for a better understanding of the roles of redox partners to LPMOs which may help unravel further details of how these enzymes are harnessed in nature.

## Results

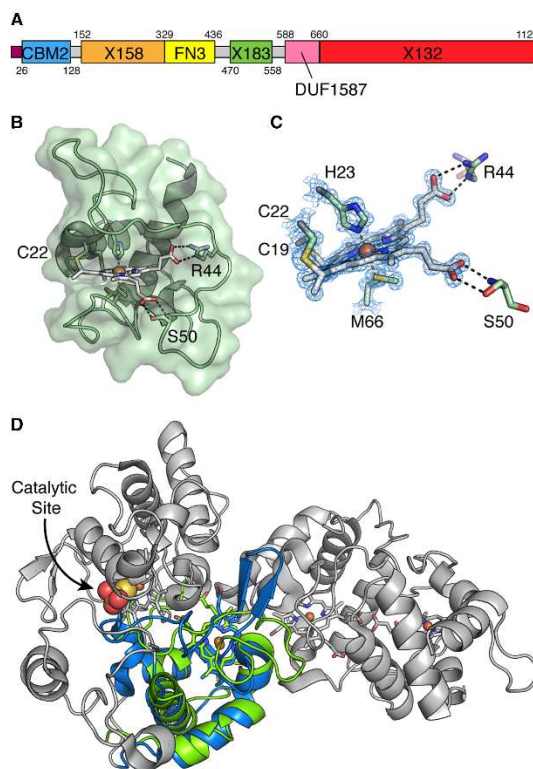
### Bioinformatic analysis of Cbp2D reveals several domains of unknown function

Gardner *et al.* have shown previously that *cbp2D* codes for a protein that is predicted to contain a carbohydrate-binding module, a YceI-like quinone binding domain, a fibronectin type III (FN3) domain and two *c*-type cytochrome domains [57]. To better define the domain boundaries for these, we performed a sequence analysis of Cbp2D to generate the annotation shown in Figure 1A. Vincent *et al.* defined the YceI-like domain in a protein similar to Cbp2E from *Saccharophagus degradans* as an X158 domain. This domain has been demonstrated to bind to polyisoprenoids, with one of the two structures determined containing ubiquinone-8 (UQ-8), suggesting a potential redox role for proteins containing these domains [59]. Likewise, Gardner *et al.* [57] determined that Cbp2D contains two domains at the C-terminus which encompass distinct sequences of different lengths, but both harbour a single CXXCH motif indicative that they are *c*-type cytochrome-like domains. We have defined these as X183 and X132 domains, in line with the naming convention for X158, as these represent domains of unknown function but are regularly associated with CBMs and other domains that are known to CAZy (for further examples of proteins harbouring these domains see the Supplementary Spreadsheet S1) [17]. Given the apparent redox nature of these proteins, we set out to better structurally and biochemically characterise Cbp2D from *Cellvibrio japonicus* with the aim of determining this protein's potential to act as an LPMO electron donor. We were unable to produce the full-length protein or express a range of constructs coding for the individual X-domains encoded within the linear sequence. The X183 construct, however, produced protein at high-levels and herein we describe our characterisation of this domain.

### *CjX183* is a small *c*-type cytochrome-like domain

*CjX183* was expressed in the periplasm of *E. coli* with the aid of the pEC86 vector which codes for the cytochrome maturation machinery to allow the insertion of heme into the protein [61]. The red coloured protein was subsequently purified to homogeneity, crystallized and the structure determined to 1.2 Å resolution (see Supplementary Table S1 for structure solution and refinement statistics). The structure reveals *CjX183* to be an  $\alpha$ -helical *c*-type cytochrome domain (Figure 1B) harbouring a single heme molecule, which is covalently linked to the conserved CXXCH motif via thioether linkages to Cys19 and Cys22. Two additional cysteines (residues 74 and 78) are present towards the C-terminus of the domain and these form a disulfide bond, presumably to stabilise the protein in the extracellular environment. The central heme iron atom is coordinated by a proximal histidine residue, His23, and a distal methionine residue, Met66 (Figure 1C) in the typical configuration observed for class one *c*-type cytochromes [62]. Interestingly, the edges of the heme molecule appear to be largely solvent exposed (Figure 1), however, the propionate groups form hydrogen bonds with Arg44 and Ser50 located in a surface loop which extends around this part of the molecule. These interactions may account for the relative stability of the reduced state discussed later.

Structural comparisons against the PDB were performed using the DALI server [63]. *CjX183* aligns with a number of other small *c*-type cytochrome domain-containing proteins with the closest match being a single domain from the thiosulfate dehydrogenase/tetrathionate reductase, TsdBA, from *Marichromatium purpuratum* (PDB ref 4v2k, [60]). The superposition reveals an rmsd of 2.3 Å between the proteins over 72 C $\alpha$  positions and a sequence identity of only 31% (Figure 1D). TsdBA in this organism is an unusual gene fusion in which the genes that typically code for the proteins TsdA and TsdB have been fused, coding for their expression in a single polypeptide chain. A series of *c*-type cytochrome domains in these proteins form an electron relay, which feeds electrons generated by the oxidation of thiosulfate into either photosynthetic or respiratory electron transfer chains [60]. Within TsdA, there are two *c*-type cytochrome domains, one of which has unusual



**Figure 1. Structural analysis of CjX183.**

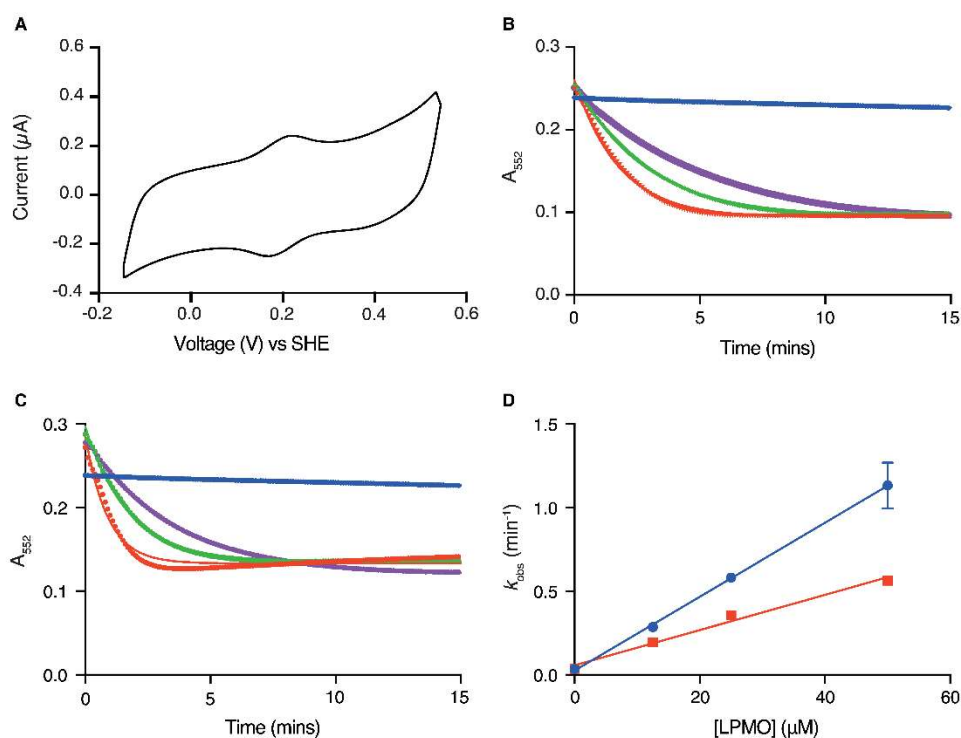
(A) Domain annotation for Cbp2D used throughout this work. Numbers indicate the point in the linear sequence at which the domain boundaries reside. X domains represent domains of unknown function whilst FN3 and DUF1587 domains are annotated in pfam (see Supplementary Spreadsheet S1 for further examples of proteins that harbour these domains). (B) Overall structure of CjX183 displayed as a cartoon with the protein surface shown in green. The heme molecule bound to the protein is shown as sticks with white carbon atoms and the central iron atom shown as an orange sphere. Important residues that interact with the heme molecule are also shown as sticks coloured by atom type with green carbon atoms. (C) Close up view of the heme molecule in CjX183. The  $2F_o - F_c$  map is shown as a blue mesh around the heme molecule and coordinating residues contoured at  $1\sigma$ . (D) Superposition of CjX183 (green) with TsdBA from *Marichromatium purpuratum* (PDB ref 4v2k, [60]). The domain in TsdBA with which CjX183 superposes is coloured blue with the rest of the protein coloured grey. The catalytic site in TsdBA is indicated by the molecule of thiosulfate shown as spheres. The chain of heme molecules in TsdBA, shown with grey carbons, are proposed to shuttle electrons away from the active site following catalysis [60].

coordination of the heme iron via a cysteine and histidine and is the catalytic site at which thiosulfate is converted to tetrathionate. The second domain is responsible for shuttling electrons away following catalysis to be utilised elsewhere in respiration/photosynthesis. CjX183 aligns to the electron transfer domain within TsdA, suggesting a similar electron transfer role for CjX183 as is typical of this cytochrome class (Figure 1D). It should be noted that in TsdBA the heme propionate groups of the domain that CjX183 superposes with are ideally positioned for receiving electrons from the TsdA active site heme. It is therefore possible that CjX183 may be similarly positioned with reference to the other domains within the full-length Cbp2D, playing a role in shuttling electrons through the protein. Without a structure of the full-length protein, we cannot determine whether this is the case, and so we decided to investigate the redox properties of the domain in isolation as a first step to determining whether Cbp2D may have a role in LPMO activation.

## CjX183 can donate electrons to LPMOs and activate them for cellulose degradation

Protein film voltammetry was used to assess the likelihood that CjX183 could activate LPMOs. As shown in Figure 2A, when CjX183 is adsorbed onto a graphite electrode and the voltage is swept between  $-0.15$  and  $+0.54$  V vs SHE at pH 7.0 and  $5^{\circ}\text{C}$ , clear peaks centred around  $\sim 0.2$  V vs SHE are observed for the oxidation (positive current) and reduction (negative current) of the protein domain. Analysis of these signals (Supplementary Figure S1 and Supplementary Notes [64,65]) indicates that they arise from a one-electron redox reaction that is attributed to the reversible  $\text{Fe}^{\text{III/II}}$  chemistry of the heme within CjX183. This experiment suggested that in principle CjX183 is capable of activating LPMOs because the extracted midpoint potential ( $+193$  mV vs SHE, Supplementary Table S2) sits below the reduction potentials that have been reported for LPMOs (typically greater than  $+220$  mV vs SHE)[28,31,66–68]. We sought to investigate this further biochemically.

CjX183 demonstrates a typical UV-vis absorption spectrum for a c-type cytochrome with an intense Soret band at 413 nm in the oxidised state, which shifts to 419 nm upon reduction, concomitant with the appearance of the  $\alpha$  and  $\beta$  bands which have absorption maxima at 523 nm and 552 nm, respectively (Supplementary Figure S2). These spectroscopic properties make it easy to monitor the redox state of CjX183 by measuring the absorbance at 552 nm ( $A_{552}$ )[62]. CjX183 was reduced using ascorbate and then passed down a PD-10



**Figure 2. Analysis of CjX183 redox reactivity.**

(A) Cyclic voltammetry measured for CjX183 on a graphite electrode reveals a reduction potential of  $+193$  mV vs SHE. (B) Monitoring of the reduced state in CjX183 as measured by the  $A_{552}$  over time in the absence of CjAA10BΔCBM (blue), and then with 1.25 fold (purple), 2.5 fold (green) and 5 fold (red) excess of CjAA10BΔCBM. (C) Equivalent plot as panel B for CfAA10. (D) Plot of oxidation rates for CjX183 against CjAA10BΔCBM (red) or CfAA10 (blue) concentration show a linear relationship.

desalting column to remove any excess reducing agent. The UV-Vis spectrum was then monitored over time whilst the sample was exposed to air. The complete oxidation of the heme in CjX183 took ~2 h demonstrating that it remains reduced in the presence of O<sub>2</sub> for an extended period, even though the structure reveals that the heme is quite solvent exposed. This is likely the result of the heme propionate groups showing close interactions with a loop that may help stabilise the reduced state.

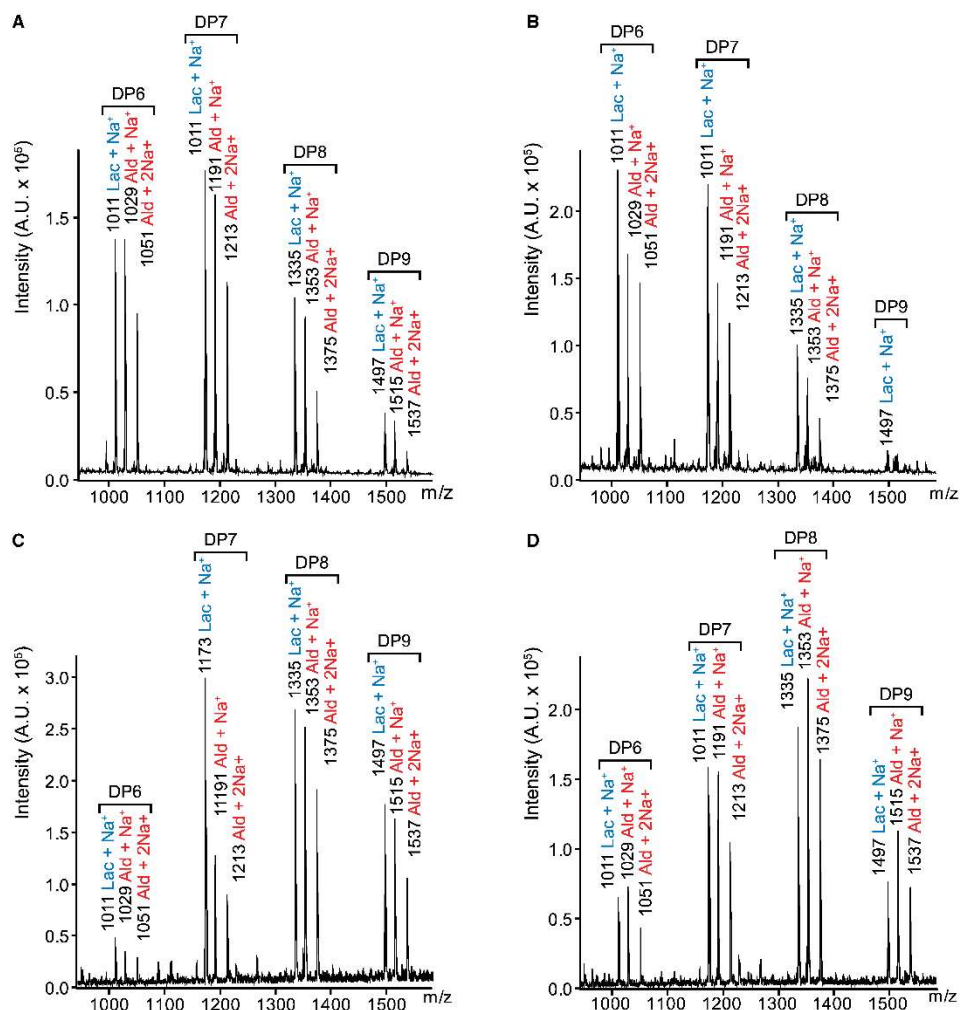
Since CjX183 appeared to maintain its reduced state in an aerobic environment, we used the UV-Vis spectrum to investigate the effect of the addition of LPMOs on the rate of heme oxidation. For these and subsequent experiments, two LPMOs were used. The cellulose active LPMO from *Celivibrio japonicus*, CjAA10B, which has been previously investigated by Gardner *et al.* [57], and an LPMO from *Cellulomonas fimi*, CfAA10, characterised by Crouch *et al.* [69]. Both enzymes typically possess C-terminal carbohydrate binding modules but in order to produce enough enzyme for our work we expressed a truncated version of the *C. japonicus* enzyme lacking its CBM which will be referred to as CjAA10BΔCBM henceforth. Upon addition of each LPMO, CjX183 demonstrated an increased rate of oxidation as demonstrated by a more rapid decrease in A<sub>552</sub> (Figure 2B,C). Fitting these data to a single exponential provided the rate of decay of the reduced state. The rate of CjX183 oxidation by oxygen is close to zero over this timescale so was assumed not to affect the assay. As the concentration of CjAA10BΔCBM or CfAA10 was increased from 0 to 50 μM, the oxidation rate of CjX183 also increased linearly up to a highest rate of 0.56 min<sup>-1</sup> for CjAA10BΔCBM and 1.13 min<sup>-1</sup> for CfAA10 (Figure 2D). These rates are considerably slower than those reported by Loose *et al.* [70] for electron transfer between MtCDH and a bacterial AA10 (observed electron transfer rate of 32 s<sup>-1</sup> when a 1 : 10 molar ratio of CDH:AA10 was used), which may suggest a different mode of interaction between the proteins here. The results of this assay, nonetheless, suggest an enzyme-dependent electron transfer event from reduced CjX183 to both of the LPMOs tested. The rate of oxidation appeared to be faster when using CfAA10 as opposed to the enzyme from the same species as CjX183. The reason for this is currently unclear, but these results appeared to further support the notion that CjX183 could activate these enzymes (and others) for cellulose degradation.

Activity assays with both LPMOs were carried out on phosphoric acid swollen cellulose (PASC) in which either ascorbate or reduced CjX183 were used as the electron source. Following overnight incubation of the reaction components, samples were analysed by MALDI-TOF mass spectrometry to detect the presence of oxidised oligosaccharide products arising from LPMO activity (Figure 3). Controls using ascorbate as the reducing agent clearly liberated typical lactone and aldonic acid terminated oligosaccharides (Figure 3A,B), as had been observed previously with these LPMOs [57,69]. Importantly, when chemically reduced CjX183 was used as the electron donor a similar product profile was observed with both LPMOs (Figure 3C,D). To check that the observed activity was due to electron donation from CjX183, and not any residual ascorbate that may remain following desalting, we also used electrochemically reduced CjX183, which had not been exposed to chemical reducing agents, to test the activity of CfAA10 (Supplementary Figure S3). The results of this experiment confirmed the same result, giving a clear demonstration that reduced CjX183 alone can induce LPMO activity on cellulose.

### Use of CjX183 as an electron source results in less hydrogen peroxide production by the LPMO

It is now well established that LPMOs will turnover oxygen to hydrogen peroxide in the presence of chemical reducing agents and the absence of a polysaccharide substrate [71,72]. This ‘side’ activity has been used as a proxy for measuring LPMO activity where direct kinetic measurements of cellulose degradation has not been possible, so we were interested to see how our X183 domain compared with small molecule reducing agents in such conditions. We therefore used the well-established Amplex Red H<sub>2</sub>O<sub>2</sub> assay [71–73] to detect H<sub>2</sub>O<sub>2</sub> production by both CjAA10BΔCBM and CfAA10 in the presence of ascorbate or reduced CjX183. We took great care to perform controls using the reducing agents in the absence of the LPMOs and to take into account any absorbance effects resulting from the presence of the CjX183 in the preparation of H<sub>2</sub>O<sub>2</sub> standard curves to quantitate the assay. Since the reductants were not being regenerated during these experiments, CjX183 was used as a one electron donor and ascorbate as a two-electron donor, consequently, CjX183 was used at twice the concentration of ascorbate to ensure the same number of electrons were available to the LPMO in each reaction.

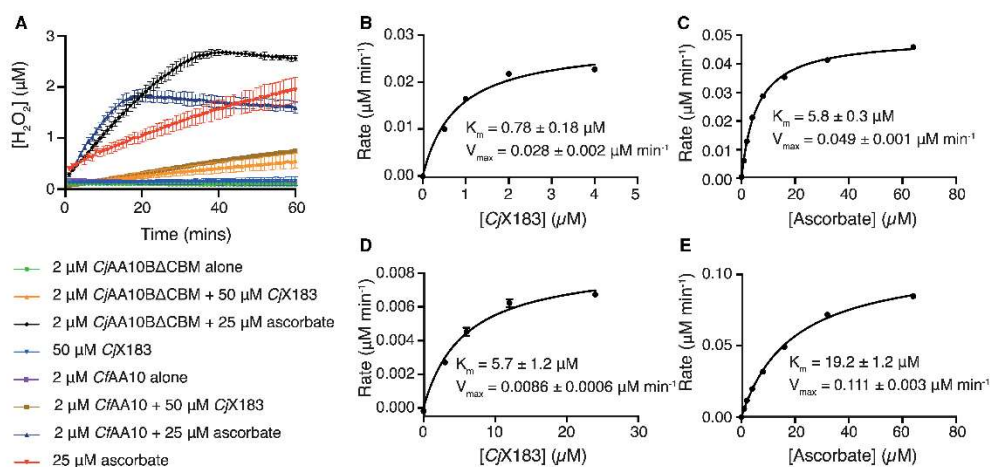
In the absence of either LPMO, the natural autooxidation of ascorbate produced significantly more H<sub>2</sub>O<sub>2</sub> compared with CjX183 alone (Figure 4A). This is in line with other’s findings which highlight that ascorbate



**Figure 3.** MALDI-ToF MS analysis of *CjAA10* and *CjAA10ΔCBM* enzymatic products.

C1 oxidised oligosaccharide products released from PASC when (A) using ascorbate to activate *CjAA10BΔCBM*; (B) using ascorbate to activate *CjAA10*; (C) using *CjX183* to activate *CjAA10BΔCBM*; and (D) using *CjX183* to activate *CjAA10*. In all panels the *m/z* of each peak is annotated together with the oxidised product formed – Lac representing oligosaccharides with a reducing end lactone, and Ald representing oligosaccharides with an aldonic acid at their reducing ends. DP $n$  (where  $n$  is an integer) indicates the length of the oligosaccharide products.

produces significant levels of  $H_2O_2$  which may be utilised by the LPMO in its reaction [23,46,72]. Our controls lacking LPMO, also further highlight *CjX183*'s ability to remain in its reduced state following removal of exogenous reducing agents. A marked increase in  $H_2O_2$  production stemming from LPMO activity was observed in the presence of *CjAA10BΔCBM* and *CjAA10* using both reducing agents (Figure 4A). For both *CjAA10BΔCBM* and *CjAA10*, activation using reduced *CjX183* produced significantly less  $H_2O_2$  after an hour of incubation compared with the use of ascorbate – 0.55  $\mu M$  and 0.75  $\mu M$  for *CjAA10BΔCBM* and *CjAA10*, respectively, compared with 2.56  $\mu M$  and 1.6  $\mu M$  when using ascorbate (Figure 4A).

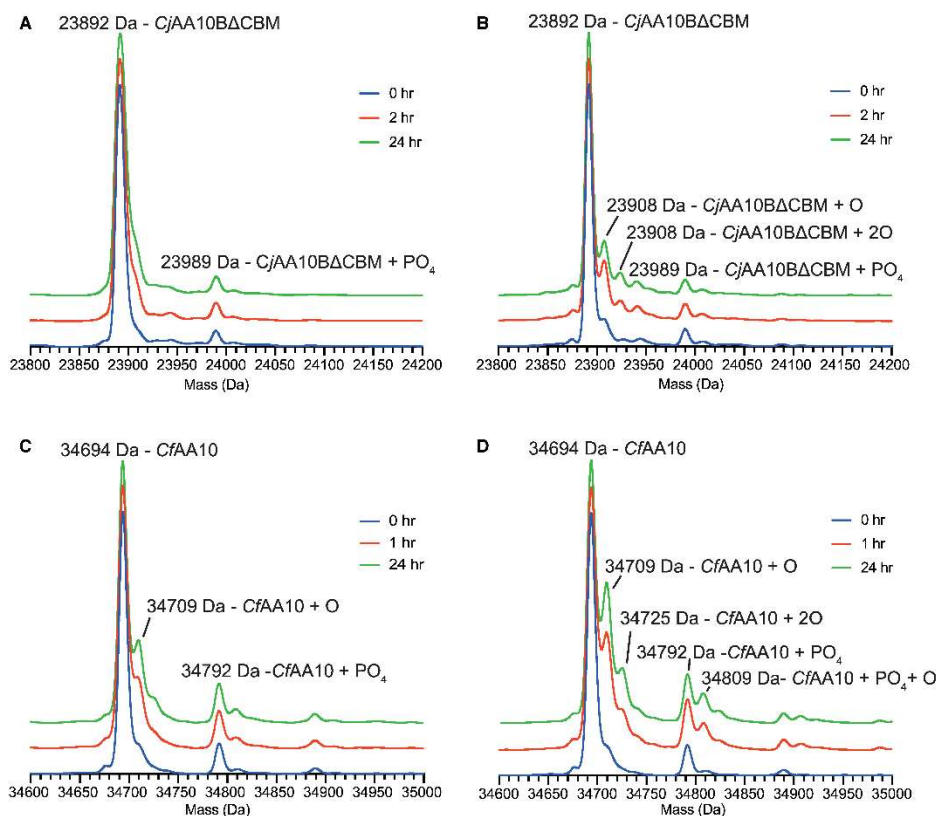


**Figure 4. Analysis of H<sub>2</sub>O<sub>2</sub> production using the Amplex Red H<sub>2</sub>O<sub>2</sub> assay.** (A) plot of the H<sub>2</sub>O<sub>2</sub> concentration against time for samples as indicated by the inset. (B) Michaelis–Menten plot for H<sub>2</sub>O<sub>2</sub> production by *CjAA10BΔCBM* when activated by *CjX183*, (C) *CjAA10BΔCBM* when activated with ascorbate, (D) *CfAA10* activated with *CjX183* (E) *CfAA10* activated by ascorbate.

To further dissect the differences in H<sub>2</sub>O<sub>2</sub> production by these enzymes we attempted to perform a more detailed kinetic analysis using this assay. *CjAA10BΔCBM* and *CfAA10* concentrations were held constant while the concentration of *CjX183*, or ascorbate, was varied around calculated  $K_M$  values, with H<sub>2</sub>O<sub>2</sub> production measured over time. Initial rates from the first 10 min were extracted and plotted against reductant concentration, giving the data shown in Figure 4B–E. Under these conditions, the rate of H<sub>2</sub>O<sub>2</sub> production by both enzymes appeared to obey Michaelis–Menten kinetics and the data fit well to such an analysis (Figure 4B–E and Supplementary Table S3). As was expected from previous experiments, the maximal rate of H<sub>2</sub>O<sub>2</sub> production by both LPMOs was achieved when ascorbate was used as the electron source. This was 2- and 10- fold lower when reduced *CjX183* was used for *CjAA10BΔCBM* and *CfAA10*, respectively. Comparing the  $K_M$  values that were extracted for each LPMO reductant pairing also shows that the  $K_M$  is consistently lower when *CjX183* is used as the reductant compared with ascorbate. Whilst this is a complicated reaction and there may be many explanations for the observed kinetic constants, if one assumes that the  $K_M$  reports on the affinity between the reductant and LPMO, these data may hint that there is a higher affinity, and potentially longer-lived interaction between *CjX183* and *CjAA10BΔCBM*, which may account for the differences in maximal H<sub>2</sub>O<sub>2</sub> production rates observed for this pairing as compared with ascorbate. Furthermore, the maximal rate of H<sub>2</sub>O<sub>2</sub> production is ~3-fold higher in the *CjX183*–*CjAA10BΔCBM* pairing compared with when *CjX183* is used to activate *CfAA10*. It is, therefore, tempting to speculate that these proteins may have evolved to interact with one another, but this will require further verification using other techniques. Whether this turns out to be true or not, there were clearly differences in the performance of the LPMO in producing H<sub>2</sub>O<sub>2</sub> when comparing the use of the *CjX183* c-type cytochrome to ascorbate, which both represent similarly finite electron sources which were not being replenished.

### Using *CjX183* as reductant results in less protein damage

Recent studies have suggested that there is a delicate balance to be struck between the concentrations of H<sub>2</sub>O<sub>2</sub> and reductant to ensure that the LPMO is maximally active for as long as possible [23,37–39,74]. Given the apparent lower level of H<sub>2</sub>O<sub>2</sub> production that we observed when using *CjX183* as the electron source, we reasoned that this would likely result in less damage to the LPMOs themselves. We therefore used electrospray ionisation mass-spectrometry (ESI-MS) to assess this by incubating the enzymes with reduced *CjX183* or ascorbate as would be done during an activity assay, but in the absence of cellulose. Accurate ESI-MS spectra were then taken for the proteins at set time points to assess whether the protein was accumulating oxidative damage



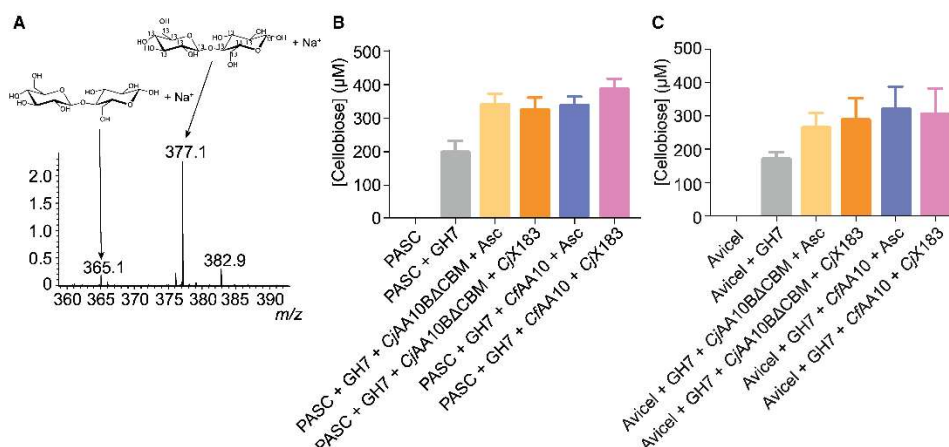
**Figure 5. Protein damage analysis by electrospray (ESI) mass spectrometry.**

(A) ESI mass spectra for *CjAA10BΔCBM* at the indicated time points following incubation with *CjX183*. (B) ESI mass spectra as taken for panel A in which the electron source was replaced with an equivalent quantity of ascorbate. (C) ESI mass spectra for *CfAA10* at the indicated time points after incubation with *CjX183*. (D) ESI mass spectra as taken for panel C in which the electron source was replaced with an equivalent quantity of ascorbate.

over time. The resulting spectra revealed the appearance of +16 Da species in both the *CjAA10BΔCBM* and *CfAA10* spectra, which build in over time, indicative of direct oxidation of the enzyme (Figure 5). Using the height of the oxidised protein peaks relative to the native protein peak as an indication of the level of oxidative damage to the sample, it appears that the *CfAA10* enzyme is much more prone to damage than the *CjAA10BΔCBM* protein (cf Figure 5A,B with 5C,D). Both enzymes also show fewer signs of damage when *CjX183* was used as the reductant compared with ascorbate (cf Figure 5A,C with 5B,D), with *CjAA10BΔCBM* appearing almost unaltered after 24 h when activated by the reduced *CjX183* domain (Figure 5A). These results correlate well with those from the Amplex Red  $H_2O_2$  assays in which *CfAA10* was found to produce the most  $H_2O_2$  under our assay conditions and had the highest  $V_{max}$  when ascorbate was used as the reductant, establishing a direct link between the production of  $H_2O_2$  and oxidative damage to the protein in line with other studies [23,37–39,74].

### Effective glycoside hydrolase boosting is maintained using *CjX183* as the electron source

Our analysis of the effect of using reduced *CjX183* compared with ascorbate on the production of  $H_2O_2$  by LPMOs suggested that the protein partner might be a less efficient LPMO activating agent, but that this may



**Figure 6. Activity boosting of LPMOs on cellulose.**

(A) Example spectrum for quantification of cellobiose (365 m/z species) by reference to a known quantity of  $^{13}\text{C}$ -labelled cellobiose (377.1 m/z species). (B) Histogram showing the quantity of cellobiose released by *T. longibrachiatum* GH7 in the absence and presence of CjAA10BΔCBM and CfAA10 with differing electron source on PASC and (C) on Avicel.

result in less oxidative damage to the enzyme. We therefore sought to perform boosting experiments with a cellulase to assess whether using CjX183 as the source of electrons to the LPMO had an overall influence on the efficiency with which cellulose can be degraded in an enzyme cocktail. Cellulose degradation assays were therefore performed on PASC and Avicel using a GH7 cellobiohydrolase from *Trichoderma longibrachiatum* in the presence of CjAA10BΔCBM or CfAA10 and either ascorbate or reduced CjX183 as the electron source. The cellobiose that was released from the substrate was then quantified by adding a known amount of  $^{13}\text{C}$ -labelled cellobiose to each sample and analysing the reaction products using ESI-MS (Figure 6A). Under the conditions we tested, we observed an ~2-fold increase in cellobiose release in the presence of both CjAA10BΔCBM and CfAA10 on both substrates used (Figure 6B,C), as has been demonstrated many times for LPMOs [4–9]. Furthermore, a similar amount of cellobiose was produced by the GH7 in the presence of each LPMO irrespective of the electron source (Figure 6B,C), and hence the level of activity boosting remained the same. These results suggest that even though the CjX183-driven activation of LPMO appears slower and results in less  $\text{H}_2\text{O}_2$  production, the overall effectiveness of the LPMO as a cellulase booster is maintained under these conditions when CjX183 is used as the reductant. This may therefore indicate that the LPMO remains active for longer when CjX183 is used as the activator resulting in a similar level of boosting over the period of the experiment.

## Discussion

Much of the focus on LPMO biochemistry has recently shifted towards a model in which  $\text{H}_2\text{O}_2$  is the active co-substrate used by these enzymes during the oxidative deconstruction of cellulose [23,38,39]. The rate at which LPMOs act is significantly improved through  $\text{H}_2\text{O}_2$  supply, but importantly, the  $\text{H}_2\text{O}_2$  concentration needs to be carefully controlled to ensure that the enzyme is not inactivated by oxidative damage [37]. The proclivity of the reductant to generate  $\text{H}_2\text{O}_2$  is another topic that has come into consideration, with studies demonstrating that ascorbate particularly generates significant levels of  $\text{H}_2\text{O}_2$  which can be harnessed by the LPMO for activity [23,46,72]. Protein partners have also been widely considered as electron sources for LPMOs, with CDH representing the key player considered in this context. CDH can generate  $\text{H}_2\text{O}_2$ , but also harbours a b-type cytochrome domain which allows it to directly transfer electrons to LPMOs [47–50,75]. Bacteria are not known to possess an enzyme equivalent to CDH and candidate enzymes that may play a similar role have not been widely investigated representing a key gap in our knowledge of how LPMOs are optimally harnessed in these microbes. Cbp2D and Cbp2E are two proteins that have been suggested as candidates to play this role in bacteria, but biochemical confirmation of this has not yet been possible [57]. We therefore set out to better



characterise the potential function for a portion of one of these multi-modular proteins through structural and biochemical analysis.

Cbp2D contains at least three domains (X158, X183 and X132) that appear to have a likely redox function. Here, we have determined the structure of the isolated X183 domain (CjX183) from this protein, confirming it as a typical *c*-type cytochrome that most likely has an electron transferring (rather than a catalytic) function. Whilst we were unable to study CjX183 as part of the larger Cbp2D protein, we have demonstrated that CjX183 alone can act as an electron donor to two bacterial LPMOs, supporting their cellulose oxidising activity. The apparent rates of electron transfer between CjX183 and the two LPMOs studied here were rather slow (Figure 2) and are considerably slower than those observed when MtCDH was used to activate an AA10 from *Streptomyces coelicolor* [70]. The observed differences in rates could be driven by the different reduction potentials of these proteins/domains. We measured the reduction potential for CjX183 to be +193 mV vs SHE, which is significantly higher than those reported for the haem in CDH which are typically between +90 and +150 mV vs SHE [47]. Given that LPMOs typically have reduction potentials above +220 mV vs SHE, the driving force for electron transfer between CDH and an LPMO will be greater compared with that for CjX183 which may account for this. In addition, there may be differences in the mode of interaction between the proteins which will influence the electron transfer rate. Where studied, the interaction between CDH and LPMOs has been proposed to occur by direct contact between the haem and the copper ion in the LPMO active site [53,54]. It is unknown whether CjX183 is likely to approach the LPMO in the same manner, and the surface properties of the protein may also influence the mode of interaction which could also affect the electron transfer rate. These data do not discount a potential role for this domain, or Cbp2D more broadly, as an electron donor to LPMOs and so we were interested to investigate how this apparent slow electron donation from the X183 may influence LPMO activity.

Since CjX183 was not linked to an enzymatic activity, we could directly compare this domain's ability to activate LPMOs with that of ascorbate, which would represent a similarly finite electron source, not replenished during LPMO activity assays. We reasoned that this may report on differences relating to the use of a protein-based electron donor compared with a small molecule reductant. We used the ability of the LPMO to produce H<sub>2</sub>O<sub>2</sub> as a means of assessing the efficiency with which electrons were delivered to the enzyme from the different reductants. The kinetic data from this analysis are complex and need careful interpretation given the range of potential rate-limiting steps under consideration. These include the rate of O<sub>2</sub> binding to the LPMO, the rate of electron transfer between electron source and LPMO, as well as the affinity and lifetime of a productive interaction between the electron source and LPMO. However, there are clear differences in the rate of H<sub>2</sub>O<sub>2</sub> production by each LPMO in response to the different electron sources studied. Of note, when CjX183 was used as the electron source the *K<sub>M</sub>* values extracted from the kinetic analysis were lower compared with those revealed for ascorbate. If one considers that these *K<sub>M</sub>*s may report on the affinity of the interaction between the electron source and LPMO, it may be that there is a higher affinity interaction between the proteins compared with the interaction between ascorbate and the LPMO. If this is the case, then the highest affinity interaction seems to occur between CjX183 and the LPMO from the same species: CjAA10BΔCBM. Taken together with the findings discussed above, these data suggest that there may be a specific interaction between these proteins, which we are currently investigating in greater detail. If we can confirm that a specific interaction exists between these proteins, then this may lend further support to the notion that Cbp2D's role is to activate LPMOs for action.

Having established that CjX183 can be used to activate two LPMOs to catalyse cellulose oxidation, we were interested in the potential consequences of CjX183's apparent diminished ability to induce H<sub>2</sub>O<sub>2</sub> production by the LPMO relative to ascorbate. Examining protein damage, we found that the reductant-LPMO pairing that resulted in least H<sub>2</sub>O<sub>2</sub> production also resulted in the least oxidative damage to the LPMO. This is also in line with H<sub>2</sub>O<sub>2</sub> dosing studies which have clearly established that this species directly damages the LPMO when used in reactions [23,37–39]. During our analyses we took care to ensure that the same number of electrons were available to the LPMO, irrespective of whether ascorbate or reduced CjX183 was used as the electron source, so it is not simply the number of electrons available that determine the H<sub>2</sub>O<sub>2</sub> production level by the LPMO. As might be expected, H<sub>2</sub>O<sub>2</sub> production must be influenced by other factors including the reduction potentials of the species involved and the mode of interaction between the LPMO and reductant. The use of a protein partner as an electron source may, therefore, have significant effects upon the LPMO reaction and so attention should continue to be paid to the role of protein partners in LPMO biochemistry.

Irrespective of whether ascorbate or CjX183 was used as the activator, there was no significant difference in the amount of cellobiose liberated from either PASC or Avicel by a GH7 cellobiohydrolase in the presence of the LPMOs used during this study. This is somewhat reflective of studies where boosting has been examined in the context of the effect of H<sub>2</sub>O<sub>2</sub> feeding on overall saccharification. Whilst LPMO activity has been shown to be significantly improved in terms of kinetics when H<sub>2</sub>O<sub>2</sub> is provided as the co-substrate [23,38,39], only modest improvements in boosting were observed by Müller *et al.* [37] in carefully controlled H<sub>2</sub>O<sub>2</sub> feeding experiments. This likely reflects the fact that it is the cellulase kinetics that ultimately define the rate at which cellulose is degraded overall. The LPMO reaction clearly benefits these enzymes but there may be a limit to those benefits, and there is a likely balance to be struck between the level of substrate oxidation and protein damage occurring during LPMO use.

In conclusion, we have demonstrated that a domain from Cbp2D in *C. japonicus* is capable of activating LPMOs for cellulose deconstruction and have used this to examine differences in LPMO activity resulting from the use of a finite electron source from either a protein or small molecule. We have been able to demonstrate that the choice of electron source can influence H<sub>2</sub>O<sub>2</sub> production by the LPMO, which was directly correlated to protein damage under the conditions that we tested. Whether Cbp2D's physiological function is to activate LPMOs remains an open question that we are currently examining, and how such an activity could be coupled to an electron generating reaction remains a key question. What is clear, is that Cbp2D must have some electron transfer capacity and based on other studies has a key role during cellulose deconstruction [57], so further study of this and related proteins should continue to be a high priority for the field.

## Experimental procedures

### Expression and purification of CjX183

CjX183 (bases 1408 to 1674 from the *cbp2d* gene in Genbank CP000934.1) was sub-cloned from a previously generated pET-26b construct which incorporated a pelB leader sequence for periplasmic protein export and C-terminal His<sub>6</sub> tag into the pCW-LIC vector (a gift from Cheryl Arrowsmith, Addgene plasmid 26098) using PIPE cloning [76]. The pCW-CjX183 plasmid was used to transform *E. coli* BL21(DE3) cells already transformed with the pEC86 plasmid, encoding the heme maturation system [61]. 1 L cultures of 2TY media (16 g/L tryptone, 10 g/L yeast extract, 6 g/L NaCl) were grown to an A<sub>600</sub> of 0.6 at 37°C shaking at 200 rpm, cooled to 20°C and induced with a final concentration of 0.4 mM IPTG. Cultures were grown for 20 h before harvesting by centrifugation at 5000×g for 20 min at 4°C.

Cells were resuspended in three volumes of ice-cold lysis buffer (50 mM Tris pH 8, 200 mM NaCl, 20% w/v sucrose) to which 40 µl of hen egg white lysozyme was added per gram of cell paste and incubated on ice for 1 h. 60 µl of 1 M MgSO<sub>4</sub> was then added per gram of cell paste and the lysate was incubated for a further 20 min on ice. The suspension was centrifuged at 10 000×g for 20 min at 4°C and supernatant containing the periplasmic fraction was removed. The pellet was then resuspended in three volumes of ice-cold water, left on ice for 1 h and centrifuged as above. The supernatants from the two lysis steps were pooled and protein purified from here on a 5 ml HisTrap FF (GE Healthcare) column equilibrated with buffer A (50 mM Tris pH 8, 200 mM NaCl, 30 mM imidazole). Protein was eluted from the column using a linear gradient of 0 to 100% buffer B (50 mM Tris pH 8, 200 mM NaCl, 300 mM imidazole) over 10 column volumes collecting 1.6 ml fractions. Peak fractions containing CjX183 were pooled and concentrated to 1 ml using a Vivaspin 3000 Da cut off concentrator at 4000×g. Concentrated protein was applied to a HiLoad 16/600 Superdex S75 column (GE Healthcare) equilibrated in GF buffer (50 mM Tris pH 8, 200 mM NaCl) collecting 1.6 ml fractions after the void volume had been eluted. Peak fractions containing the purified CjX183 were combined, concentrated on the same concentrator, and the sample was then quantified using the A<sub>410</sub> extinction coefficient for heme *c* of 106 000 M<sup>-1</sup> cm<sup>-1</sup>.

### Expression and purification of CjAA10BΔCBM and CfAA10

The pET-22-CjAA10BΔCBM construct was generated by deletion of the nucleotide sequence coding for the CBM region of the pET22-CjAA10B construct provided by Gardner *et al.* [57] using the Q5 site-directed mutagenesis kit (New England Biolabs). Following this, CjAA10BΔCBM and CfAA10 were expressed as described by Gardner *et al.* [57]. Periplasmic lysis was performed as described above, in lysis buffer (20 mM sodium phosphate pH 7, 150 mM NaCl, 20% w/v sucrose). Nickel affinity chromatography, using buffer A (20 mM sodium phosphate pH 7, 150 mM NaCl, 30 mM imidazole) and buffer B (20 mM sodium phosphate pH 7,

150 mM NaCl, 300 mM imidazole), and gel filtration using GF buffer (20 mM sodium phosphate pH 7, 150 mM NaCl) were conducted as for *CjX183*. Pure protein was concentrated using a Vivaspin 10 000 Da cut off concentrator at 4000×g and quantified using the  $A_{280}$  extinction coefficient of  $53\,775\text{ M}^{-1}\text{ cm}^{-1}$  for *CjAA10BΔCBM* and  $81\,275\text{ M}^{-1}\text{ cm}^{-1}$  for *CjAA10*.

### Crystallization, X-ray data collection and structure determination for *CjX183*

*CjX183* was crystallized in 0.1 M HEPES pH 7.5, 20% w/v PEG 4000, 10% w/v propan-2-ol via sitting drop vapour-diffusion. Crystals were cryocooled by soaking crystals in mother liquor supplemented with 20% ethylene glycol for 30 s before plunging into liquid nitrogen. Diffraction data was collected at Diamond Light Source on beamline I03 with a wavelength of 0.976 Å. Data were indexed and integrated using XDS [77] with subsequent data processing performed in the CCP4i2 suite [78]. Anomalous diffraction from the heme iron atom was used for phasing using the SHELXC/D/E pipeline [79]. Model building and refinement was performed using iterative cycles of restrained refinement and model building with COOT [80] and REFMAC5 [81].

### Voltammetric electrochemical analysis of *CjX183*

Cyclic voltammetry electrochemical measurements were performed using a standard three electrode set-up consisting of: a working pyrolytic graphite edge electrode attached to an Orgitrod rotator operated in stationary mode; a saturated calomel reference electrode filled with saturated aqueous KCl solution, and a Pt wire counter electrode. For experiments on *CjX183* a 1 μl aliquot of protein was pipetted onto the surface of a freshly abraded working electrode and left to form an adsorbed film for ~1 min. The three electrodes were contained within a custom-built electrochemical cell (constructed by the University of York Department of Chemistry Glass Workshop) surrounded by a thermostat-controlled water jacket which was maintained at 5°C. Calibration of the reference electrode to a standard hydrogen electrode gave a correction factor of +0.243 V which has been applied to the data. The measurements were performed in 50 mM sodium phosphate pH 7, 150 mM NaCl.

### UV-Visible detection of *CjX183* oxidation

Reduced *CjX183* was prepared by complete reduction using 100 mM ascorbate. Ascorbate was then removed by passing the sample through a PD-10 desalting column before *CjX183* was quantified using its  $A_{552}$  and the extinction coefficient for heme  $c$  of  $27\,500\text{ M}^{-1}\text{ cm}^{-1}$ . A batch of 20 μM *CjX183* was prepared, aliquoted and flash frozen in liquid N<sub>2</sub>. UV-Visible spectroscopy experiments were performed in 50 mM sodium phosphate pH 6 in 160 μl volumes in a quartz cuvette on a Cary60 spectrophotometer (Agilent). In the assay, reduced *CjX183* was used at 10 μM and LPMO concentrations were varied between 0 and 50 μM. Reactions were prepared by thawing aliquots of reduced *CjX183* and mixing 80 μl with 80 μl of LPMO at twice the desired concentration. Reactions were monitored at 552 nm with points taken every 0.1 s for 15 min. Data were fitted to a single exponential using the equation for one phase decay on GraphPad, from which rates of decay of the reduced state of *CjX183* could be extracted.

### Activity monitoring using matrix assisted laser desorption ionisation mass spectrometry

Activity assays were set up with phosphoric acid swollen cellulose (PASC). PASC was prepared as described by Wood [82]. Assays using chemically reduced *CjX183* were set up at room temperature with a volume of 1 ml in 5 mM ammonium acetate, pH 6, comprising 10% v/v PASC, 1 μM *CjAA10BΔCBM* or *CjAA10* and 1 mM ascorbate or 150 μM reduced *CjX183*. Assays were performed using either chemically reduced *CjX183* prepared as described above, or electrochemically reduced *CjX183*. For electrochemical reduction, 2 ml of a 235 μM solution of *CjX183* in a pH 7 mixed buffer solution (consisting of 150 mM of NaCl and 50 mM each of acetate, Tris, phosphate and MES) was incubated at 0°C under an ambient atmosphere in the presence of 2.35 μM of the redox mediator methyl viologen and a carbon felt working electrode, held at -400 mV vs SHE for 1180 s. The oxidative state was monitored by the formation of a peak at 552 nm, indicative of the reduced state. Electrochemically reduced *CjX183* was then buffer exchanged using a Zeba spin desalting column into 50 mM sodium acetate pH 6 buffer before use. Activity assays for electrochemically reduced *CjX183* were set up in 50 mM sodium phosphate pH 6 with reaction constituents used at concentrations described above. Reactions were left for 16 h on the bench rotating head over tail at room temperature and analysed using Matrix Assisted Laser Desorption Ionisation Mass Spectrometry (MALDI-MS).

For MALDI-MS experiments, 1  $\mu\text{l}$  of sample was mixed with an equivalent volume of 10 mg/ml 2,5-dihydroxybenzoic acid in 50% acetonitrile, 0.1% trifluoroacetic acid on a Bruker SCOUT-MTP 384 target plate. The spotted samples were then dried in air under a lamp before being analysed by mass spectrometry on a Ultraflex III matrix-assisted laser desorption ionization-time-of-flight/time-of-flight (MALDI-TOF/TOF) instrument (Bruker), as described in Hemsworth *et al.* [13].

### Activity monitoring using the Amplex Red $\text{H}_2\text{O}_2$ assay

Hydrogen peroxide production by *CjAA10* $\Delta$ CBM and *CfAA10* was measured in 96 well plates on a FLUOstar Galaxy (BMG Labtech) plate reader using a coupled assay with Amplex Red [73]. Resorufin formation was detected in fluorescence mode using an excitation wavelength of 570 nm and an emission wavelength of 596 nm. All reactions were performed in triplicate in 50 mM sodium phosphate pH 6.0, at room temperature.  $\text{H}_2\text{O}_2$  production comparison assays used ascorbate at 25  $\mu\text{M}$ , *CjX183* at 50  $\mu\text{M}$  and LPMOs at 2  $\mu\text{M}$ . In assays used to determine the  $K_M$  and  $V_{\text{max}}$  for each enzyme, *CjAA10* $\Delta$ CBM and *CfAA10* were used at 2  $\mu\text{M}$  and concentration of reduced *CjX183* was varied around the  $K_M$  for each enzyme: between 0 and 64  $\mu\text{M}$  for *CjAA10* $\Delta$ CBM and 0–128  $\mu\text{M}$  for *CfAA10* before settling on the final concentrations presented in Figure 4. 50  $\mu\text{M}$  Amplex Red and 0.2  $\text{U ml}^{-1}$  horseradish peroxidase were used for enzyme assays and assays were started with the addition of the electron donor. Reactions were monitored every minute for 120 min. The necessary LPMO free controls were performed alongside enzymatic assays, as *CjX183* and ascorbate will produce peroxide through oxidation in air. Fluorescence readings were converted to  $[\text{H}_2\text{O}_2]$  using a concentration curve performed with 0–10  $\mu\text{M}$   $\text{H}_2\text{O}_2$  and measured at each concentration of *CjX183* to account for any inner filter effects that may arise from light absorption from the heme. Initial rates were measured using data from the first 10 min and fitted to a straight line, rates were plotted against substrate concentration and kinetic parameters were extracted using the Michaelis–Menten function on GraphPad.

### Glycoside hydrolase boosting activity assays

Boosting assays were set up in 500  $\mu\text{l}$  volumes in 50 mM sodium phosphate pH 6 buffer with 10% v/v PASC or 10% w/v Avicel. Endo-1,4- $\beta$ -D-Glucanase from *T. longibrachiatum* (Megazyme) was used at 1  $\text{U ml}^{-1}$  (or  $\sim 250$  nM), *CjAA10* $\Delta$ CBM and *CfAA10* were each used at 2  $\mu\text{M}$ , with ascorbate used at 75  $\mu\text{M}$  and *CjX183* at 150  $\mu\text{M}$ . Reactions were incubated at room temperature, rotating head over tail for 8 h on PASC or 16 h on Avicel. Reactions were stopped by heating at 95°C for 10 min and the insoluble material was pelleted by centrifugation. Soluble material was diluted 10-fold into water, combined in a 1 : 1 ratio with 100  $\mu\text{M}$   $^{13}\text{C}$  labelled cellobiose (Omicron Biochemicals Inc.) and analysed by ESI-MS. Cellobiose production was quantified by comparing the intensity of the  $[\text{cellobiose-Na}]^+$  peak at 365 Da with that of the  $[\text{C}^{13}\text{cellobiose-Na}]^+$  peak at 377 Da.

### Electrospray mass spectrometry to assess protein oxidation

Protein damage experiments were set up in 100  $\mu\text{l}$  volumes in 50 mM sodium phosphate pH 6. 20  $\mu\text{M}$  *CjAA10* $\Delta$ CBM or *CfAA10* were mixed with 200  $\mu\text{M}$  *CjX183* or 100  $\mu\text{M}$  ascorbate. Samples were taken upon mixing, after 1 h, for the *CfAA10* enzyme, after 2 h, for the *CjAA10* $\Delta$ CBM enzyme, and 24 h and analysed by electrospray mass spectrometry.

### Data Availability

The structure and accompanying structure factors for *CjX183* have been deposited with the protein data bank with accession code 7B21. All other data are present in the manuscript or supporting information. For access to raw data, please contact the corresponding author.

### Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

### Funding

G.R.H. is grateful for support from a BBSRC (Biotechnology and Biological Sciences Research Council) David Phillips Fellowship (BB/N019970/1) and internal funding from the University of Leeds, all of which supports the work of J.B. and B.S.R. The Mass Spectrometry Facility instrumentation used for protein intact mass analysis in this work was supported by the BBSRC (BB/M012573/1) and analysis was carried out by Rachel George. We

thank Diamond Light Source for access to beamline I03 (proposal number mx-15378) that contributed to the results presented here.

### CRedit Author Contribution

**Glyn R. Hemsworth:** Conceptualization, Resources, Data curation, Formal analysis, Supervision, Funding acquisition, Validation, Investigation, Visualization, Methodology, Project administration, Writing — review and editing. **Jessie Branch:** Conceptualization, Resources, Data curation, Formal analysis, Supervision, Funding acquisition, Validation, Investigation, Visualization, Methodology, Writing — original draft, Project administration, Writing — review and editing. **Badri S. Rajagopal:** Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing — original draft, Writing — review and editing. **Alessandro Paradisi:** Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing — review and editing. **Nick Yates:** Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing — review and editing. **Peter J. Lindley:** Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing — review and editing. **Jake Smith:** Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing — review and editing. **Kristian Hollingsworth:** Conceptualization, Resources, Data curation, Formal analysis, Supervision, Funding acquisition, Validation, Investigation, Visualization, Methodology, Project administration, Writing — review and editing. **Bruce Turnbull:** Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing — review and editing. **Bernard Henrissat:** Conceptualization, Data curation, Software, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing — review and editing. **Alison Parkin:** Conceptualization, Data curation, Software, Formal analysis, Supervision, Validation, Investigation, Visualization, Methodology, Writing — review and editing. **Alan berry:** Conceptualization, Data curation, Formal analysis, Supervision, Validation, Investigation, Visualization, Methodology, Writing — review and editing.

### Acknowledgements

We gratefully acknowledge Dr Lucy Crouch, Prof Harry Gilbert (Newcastle University) and Dr Jeffrey Gardner (University of Maryland Baltimore) for providing constructs of Cbp2D as templates to begin this work and for useful discussions. We also thank Prof Lars Jeuken and Dr Theodoros Laftoglou for providing the pEC86 plasmid used to help express CjX183. G.R.H would also like to thank Prof Paul Walton for reading drafts of the manuscript, for his support with experiments performed in his laboratory, and for his mentorship which has been invaluable in supporting the work of G.R.H.'s laboratory.

### Abbreviations

CDH, cellobiose dehydrogenase; ESI-MS, electrospray ionisation mass-spectrometry; FN3, fibronectin type III; LPMOs, lytic polysaccharide monoxygenases; MALDI-MS, Matrix Assisted Laser Desorption Ionisation Mass Spectrometry; PASC, phosphoric acid swollen cellulose.

### References

- 1 Harris, P.V., Xu, F., Kreef, N.E., Kang, C. and Fukuyama, S. (2014) New enzyme insights drive advances in commercial ethanol production. *Curr. Opin. Chem. Biol.* **19**, 162–170 <https://doi.org/10.1016/j.cbpa.2014.02.015>
- 2 Himmel, M.E., Ding, S.Y., Johnson, D.K., Adney, W.S., Nimlos, M.R., Brady, J.W. et al. (2007) Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* **315**, 804–807 <https://doi.org/10.1126/science.1137016>
- 3 Singhvi, M.S., Chaudhari, S. and Gokhale, D.V. (2014) Lignocellulose processing: a current challenge. *RSC Adv.* **4**, 8271–8277 <https://doi.org/10.1039/c3ra46112b>
- 4 Cannella, D., Hsieh, C.W., Felby, C. and Jorgensen, H. (2012) Production and effect of aldonic acids during enzymatic hydrolysis of lignocellulose at high dry matter content. *Biotechnol. Biofuels* **5**, 26 <https://doi.org/10.1186/1754-6834-5-26>
- 5 Harris, P.V., Welner, D., McFarland, K.C., Pe, E., Navarro Poulsen, J.C., Brown, K. et al. (2010) Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. *Biochemistry* **49**, 3305–3316 <https://doi.org/10.1021/bi100009p>
- 6 Lo Leggio, L., Simmons, T.J., Poulsen, J.C., Frandsen, K.E., Hemsworth, G.R., Stringer, M.A. et al. (2015) Structure and boosting activity of a starch-degrading lytic polysaccharide monoxygenase. *Nat. Commun.* **6**, 5961 <https://doi.org/10.1038/ncomms6961>
- 7 Quinlan, R.J., Sweeney, M.D., Lo Leggio, L., Otten, H., Poulsen, J.C., Johansen, K.S. et al. (2011) Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad. Sci. U S A.* **108**, 15079–15084 <https://doi.org/10.1073/pnas.1105776108>
- 8 Sabbadin, F., Hemsworth, G.R., Clano, L., Henrissat, B., Dupree, P., Tryfona, T. et al. (2018) An ancient family of lytic polysaccharide monoxygenases with roles in arthropod development and biomass digestion. *Nat. Commun.* **9**, 756 <https://doi.org/10.1038/s41467-018-03142-x>
- 9 Vaaje-Kolstad, G., Westereng, B., Horn, S.J., Liu, Z., Zhai, H., Sorlie, M. et al. (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* **330**, 219–222 <https://doi.org/10.1126/science.1192231>

- 10 Yadav, S.K., Archana, Singh, R., Singh, P.K. and Vasudev, P.G. (2019) Insecticidal fern protein Tma12 is possibly a lytic polysaccharide monoxygenase. *Planta* **249**, 1987–1996 <https://doi.org/10.1007/s00425-019-03135-0>
- 11 Filatrault-Chastel, C., Navarro, D., Haon, M., Grisel, S., Herpoel-Gimbert, I., Chevret, D. et al. (2019) AA16, a new lytic polysaccharide monoxygenase family identified in fungal secretomes. *Biotechnol. Biofuels* **12**, 55 <https://doi.org/10.1186/s13068-019-1394-y>
- 12 Couturier, M., Ladeveze, S., Sulzenbacher, G., Ciano, L., Fanuel, M., Moreau, C. et al. (2018) Lytic xylan oxidases from wood-decay fungi unlock biomass degradation. *Nat. Chem. Biol.* **14**, 306–310 <https://doi.org/10.1038/nchembio.2558>
- 13 Hemsworth, G.R., Henrissat, B., Davies, G.J. and Walton, P.H. (2014) Discovery and characterization of a new family of lytic polysaccharide monoxygenases. *Nat. Chem. Biol.* **10**, 122–126 <https://doi.org/10.1038/nchembio.1417>
- 14 Vu, V.V., Beeson, W.T., Span, E.A., Farquhar, E.R. and Marletta, M.A. (2014) A family of starch-active polysaccharide monoxygenases. *Proc. Natl Acad. Sci. U.S.A.* **111**, 13822–13827 <https://doi.org/10.1073/pnas.1408090111>
- 15 Vaaje-Kolstad, G., Horn, S.J., Sorlie, M. and Eijsink, V.G. (2013) The chitinolytic machinery of *Serratia marcescens*—a model system for enzymatic degradation of recalcitrant polysaccharides. *FEBS J.* **280**, 3028–3049 <https://doi.org/10.1111/febs.12181>
- 16 Morgenstern, I., Powlowski, J. and Tsang, A. (2014) Fungal cellulose degradation by oxidative enzymes: from dysfunctional GH61 family to powerful lytic polysaccharide monoxygenase family. *Brief Funct. Genomics* **13**, 471–481 <https://doi.org/10.1093/bfpg/elu032>
- 17 Levasseur, A., Drula, E., Lombard, V., Coutinho, P.M. and Henrissat, B. (2013) Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol. Biofuels* **6**, 41 <https://doi.org/10.1186/1754-6834-6-41>
- 18 Bey, M., Zhou, S., Poidevin, L., Henrissat, B., Coutinho, P.M., Berrin, J.G. et al. (2013) Cello-oligosaccharide oxidation reveals differences between two lytic polysaccharide monoxygenases (family GH61) from *Podospira anserina*. *Appl. Environ. Microbiol.* **79**, 488–496 <https://doi.org/10.1128/AEM.02942-12>
- 19 Forsberg, Z., Vaaje-Kolstad, G., Westereng, B., Bunaes, A.C., Stenstrom, Y., MacKenzie, A. et al. (2011) Cleavage of cellulose by a CBM33 protein. *Protein Sci.* **20**, 1479–1483 <https://doi.org/10.1002/pro.689>
- 20 Westereng, B., Ishida, T., Vaaje-Kolstad, G., Wu, M., Eijsink, V.G., Igarashi, K. et al. (2011) The putative endoglucanase PcgH61D from *Phanerochaete chrysosporium* is a metal-dependent oxidative enzyme that cleaves cellulose. *PLoS ONE* **6**, e27807 <https://doi.org/10.1371/journal.pone.0027807>
- 21 Eibinger, M., Ganner, T., Bubner, P., Rosker, S., Kracher, D., Haltrich, D. et al. (2014) Cellulose surface degradation by a lytic polysaccharide monoxygenase and its effect on cellulase hydrolytic efficiency. *J. Biol. Chem.* **289**, 35929–35938 <https://doi.org/10.1074/jbc.M114.602227>
- 22 Ciano, L., Davies, G.J., Tolman, W.B. and Walton, P.H. (2018) Bracing copper for the catalytic oxidation of C–H bonds. *Nat. Catal.* **1**, 571–577 <https://doi.org/10.1038/s41929-018-0110-9>
- 23 Bissaro, B., Rohr, A.K., Muller, G., Chylenski, P., Skaugen, M., Forsberg, Z. et al. (2017) Oxidative cleavage of polysaccharides by monocopper enzymes depends on H<sub>2</sub>O<sub>2</sub>. *Nat. Chem. Biol.* **13**, 1123–1128 <https://doi.org/10.1038/nchembio.2470>
- 24 Forsberg, Z., Sorlie, M., Petrovic, D., Courtade, G., Aachmann, F.L., Vaaje-Kolstad, G. et al. (2019) Polysaccharide degradation by lytic polysaccharide monoxygenases. *Curr. Opin. Struct. Biol.* **59**, 54–64 <https://doi.org/10.1016/j.sbi.2019.02.015>
- 25 Walton, P.H. and Davies, G.J. (2016) On the catalytic mechanisms of lytic polysaccharide monoxygenases. *Curr. Opin. Chem. Biol.* **31**, 195–207 <https://doi.org/10.1016/j.cbpa.2016.04.001>
- 26 Hemsworth, G.R., Davies, G.J. and Walton, P.H. (2013) Recent insights into copper-containing lytic polysaccharide mono-oxygenases. *Curr. Opin. Struct. Biol.* **23**, 660–668 <https://doi.org/10.1016/j.sbi.2013.05.006>
- 27 Gudmundsson, M., Kim, S., Wu, M., Ishida, T., Momeni, M.H., Vaaje-Kolstad, G. et al. (2014) Structural and electronic snapshots during the transition from a Cu(II) to Cu(I) metal center of a lytic polysaccharide monoxygenase by X-ray photoreduction. *J. Biol. Chem.* **289**, 18782–18792 <https://doi.org/10.1074/jbc.M114.563494>
- 28 Hemsworth, G.R., Taylor, E.J., Kim, R.O., Gregory, R.C., Lewis, S.J., Turkenburg, J.P. et al. (2013) The copper active site of CBM33 polysaccharide oxygenases. *J. Am. Chem. Soc.* **135**, 6069–6077 <https://doi.org/10.1021/ja402106e>
- 29 Vaaje-Kolstad, G., Forsberg, Z., Loose, J.S., Bissaro, B. and Eijsink, V.G. (2017) Structural diversity of lytic polysaccharide monoxygenases. *Curr. Opin. Struct. Biol.* **44**, 67–76 <https://doi.org/10.1016/j.sbi.2016.12.012>
- 30 Beeson, W.T., Phillips, C.M., Cate, J.H. and Marletta, M.A. (2012) Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monoxygenases. *J. Am. Chem. Soc.* **134**, 890–892 <https://doi.org/10.1021/ja210657t>
- 31 Forsberg, Z., Mackenzie, A.K., Sorlie, M., Rohr, A.K., Helland, R., Arvai, A.S. et al. (2014) Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monoxygenases. *Proc. Natl Acad. Sci. U.S.A.* **111**, 8446–8451 <https://doi.org/10.1073/pnas.1402771111>
- 32 Frandsen, K.E., Simmons, T.J., Dupree, P., Poulsen, J.C., Hemsworth, G.R., Ciano, L. et al. (2016) The molecular basis of polysaccharide cleavage by lytic polysaccharide monoxygenases. *Nat. Chem. Biol.* **12**, 298–303 <https://doi.org/10.1038/nchembio.2029>
- 33 Kim, S., Stahlberg, J., Sandgren, M., Paton, R.S. and Beckham, G.T. (2014) Quantum mechanical calculations suggest that lytic polysaccharide monoxygenases use a copper-oxyl, oxygen-rebound mechanism. *Proc. Natl Acad. Sci. U.S.A.* **111**, 149–154 <https://doi.org/10.1073/pnas.1316609111>
- 34 Kjaergaard, C.H., Qayyum, M.F., Wong, S.D., Xu, F., Hemsworth, G.R., Walton, D.J. et al. (2014) Spectroscopic and computational insight into the activation of O<sub>2</sub> by the mononuclear Cu center in polysaccharide monoxygenases. *Proc. Natl Acad. Sci. U.S.A.* **111**, 8797–8802 <https://doi.org/10.1073/pnas.1408115111>
- 35 Beeson, W.T., Vu, V.V., Span, E.A., Phillips, C.M. and Marletta, M.A. (2015) Cellulose degradation by polysaccharide monoxygenases. *Annu. Rev. Biochem.* **84**, 923–946 <https://doi.org/10.1146/annurev-biochem-060614-034439>
- 36 Hemsworth, G.R., Johnston, E.M., Davies, G.J. and Walton, P.H. (2015) Lytic polysaccharide monoxygenases in biomass conversion. *Trends Biotechnol.* **33**, 747–761 <https://doi.org/10.1016/j.tibtech.2015.09.006>
- 37 Müller, G., Chylenski, P., Bissaro, B., Eijsink, V.G.H. and Horn, S.J. (2018) The impact of hydrogen peroxide supply on LPMO activity and overall saccharification efficiency of a commercial cellulase cocktail. *Biotechnol. Biofuels* **11**, 209 <https://doi.org/10.1186/s13068-018-1199-4>
- 38 Kuusk, S., Kori, R., Kuusk, P., Heering, A., Sorlie, M., Bissaro, B. et al. (2019) Kinetic insights into the role of the reductant in H<sub>2</sub>O<sub>2</sub>-driven degradation of chitin by a bacterial lytic polysaccharide monoxygenase. *J. Biol. Chem.* **294**, 1516–1528 <https://doi.org/10.1074/jbc.RA118.006196>

- 39 Kuusk, S., Bissaro, B., Kuusk, P., Forsberg, Z., Eijsink, V.G.H., Sorlie, M. et al. (2018) Kinetics of H<sub>2</sub>O<sub>2</sub>-driven degradation of chitin by a bacterial lytic polysaccharide monoxygenase. *J. Biol. Chem.* **293**, 523–531 <https://doi.org/10.1074/jbc.M117.817593>
- 40 Bissaro, B., Streit, B., Isaksen, I., Eijsink, V.G.H., Beckham, G.T., DuBois, J.L. et al. (2020) Molecular mechanism of the chitinolytic peroxxygenase reaction. *Proc. Natl Acad. Sci. U.S.A.* **117**, 1504–1513 <https://doi.org/10.1073/pnas.1904889117>
- 41 Calderaru, O., Oksanen, E., Ryde, U. and Hedegard, E.D. (2019) Mechanism of hydrogen peroxide formation by lytic polysaccharide monoxygenase. *Chem. Sci.* **10**, 576–586 <https://doi.org/10.1039/C8SC03980A>
- 42 Wang, B., Johnston, E.M., Li, P., Shaik, S., Davies, G.J., Walton, P.H. et al. (2018) QM/MM studies into the H<sub>2</sub>O<sub>2</sub>-dependent activity of lytic polysaccharide monoxygenases: evidence for the formation of a caged hydroxyl radical intermediate. *ACS Catal.* **8**, 1346–1351 <https://doi.org/10.1021/acscatal.7b03888>
- 43 Wang, B., Walton, P.H. and Rovira, C. (2019) Molecular mechanisms of oxygen activation and hydrogen peroxide formation in lytic polysaccharide monoxygenases. *ACS Catal.* **9**, 4958–4969 <https://doi.org/10.1021/acscatal.9b00778>
- 44 Loose, J.S.M., Arntzen, M.O., Bissaro, B., Ludwig, R., Eijsink, V.G.H. and Vaaje-Kolstad, G. (2018) Multipoint precision binding of substrate protects lytic polysaccharide monoxygenases from self-destructive off-pathway processes. *Biochemistry* **57**, 4114–4124 <https://doi.org/10.1021/acs.biochem.8b00484>
- 45 Paradisi, A., Johnston, E.M., Tovborg, M., Nicoll, C.R., Ciano, L., Dowle, A. et al. (2019) Formation of a copper(II)–tyrosyl complex at the active site of lytic polysaccharide monoxygenases following oxidation by H<sub>2</sub>O<sub>2</sub>. *J. Am. Chem. Soc.* **141**, 18585–18599 <https://doi.org/10.1021/jacs.9b09833>
- 46 Stepanov, A.A., Forsberg, Z., Sorlie, M., Nguyen, G.S., Wentzel, A., Rohr, A.K. et al. (2021) Unraveling the roles of the reductant and free copper ions in LPMO kinetics. *Biotechnol. Biofuels* **14**, 28 <https://doi.org/10.1186/s13068-021-01879-0>
- 47 Sygmund, C., Kracher, D., Scheiblbrandner, S., Zahma, K., Felice, A.K., Harreither, W. et al. (2012) Characterization of the two *Neurospora crassa* cellobiose dehydrogenases and their connection to oxidative cellulose degradation. *Appl. Environ. Microbiol.* **78**, 6161–6171 <https://doi.org/10.1128/AEM.01503-12>
- 48 Langston, J.A., Shaghasi, T., Abbate, E., Xu, F., Vlasenko, E. and Sweeney, M.D. (2011) Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Appl. Environ. Microbiol.* **77**, 7007–7015 <https://doi.org/10.1128/AEM.05815-11>
- 49 Phillips, C.M., Beeson, W.T., Cate, J.H. and Marletta, M.A. (2011) Cellobiose dehydrogenase and a copper-dependent polysaccharide monoxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem. Biol.* **6**, 1399–1406 <https://doi.org/10.1021/cb200351y>
- 50 Kracher, D., Scheiblbrandner, S., Felice, A.K., Breslmayr, E., Preims, M., Ludwicka, K. et al. (2016) Extracellular electron transfer systems fuel cellulose oxidative degradation. *Science* **352**, 1098–1101 <https://doi.org/10.1126/science.1253165>
- 51 Garajova, S., Mathieu, Y., Beccia, M.R., Bennati-Granier, C., Biaso, F., Fanael, M. et al. (2016) Single-domain flavoenzymes trigger lytic polysaccharide monoxygenases for oxidative degradation of cellulose. *Sci. Rep.* **6**, 28276 <https://doi.org/10.1038/srep28276>
- 52 Varnai, A., Urmezawa, K., Yoshida, M. and Eijsink, V.G.H. (2018) The pyroloquinoline-quinone-dependent pyranose dehydrogenase from *Coprinopsis cinerea* drives lytic polysaccharide monoxygenase action. *Appl. Environ. Microbiol.* **84**, e00156–00118 <https://doi.org/10.1128/AEM.00156-18>
- 53 Tan, T.C., Kracher, D., Gandini, R., Sygmund, C., Kittl, R., Haltrich, D. et al. (2015) Structural basis for cellobiose dehydrogenase action during oxidative cellulose degradation. *Nat. Commun.* **6**, 7542 <https://doi.org/10.1038/ncomms8542>
- 54 Courtade, G., Wimmer, R., Rohr, A.K., Preims, M., Felice, A.K., Dimarogona, M. et al. (2016) Interactions of a fungal lytic polysaccharide monoxygenase with beta-glucan substrates and cellobiose dehydrogenase. *Proc. Natl Acad. Sci. U.S.A.* **113**, 5922–5927 <https://doi.org/10.1073/pnas.1602568113>
- 55 Book, A.J., Yennamalli, R.M., Takasuka, T.E., Currie, C.R., Phillips, Jr, G.N. and Fox, B.G. (2014) Evolution of substrate specificity in bacterial AA10 lytic polysaccharide monoxygenases. *Biotechnol. Biofuels* **7**, 109 <https://doi.org/10.1186/1754-6834-7-109>
- 56 Li, X., Beeson, W.T., Phillips, C.M., Marletta, M.A. and Cate, J.H. (2012) Structural basis for substrate targeting and catalysis by fungal polysaccharide monoxygenases. *Structure* **20**, 1051–1061 <https://doi.org/10.1016/j.str.2012.04.002>
- 57 Gardner, J.G., Crouch, L., Labourel, A., Forsberg, Z., Bukhman, Y.V., Vaaje-Kolstad, G. et al. (2014) Systems biology defines the biological significance of redox-active proteins during cellulose degradation in an aerobic bacterium. *Mol. Microbiol.* **4**, 1121–1133 <https://doi.org/10.1111/mmi.12821>
- 58 Tuveng, T.R., Arntzen, M.O., Bengtsson, O., Gardner, J.G., Vaaje-Kolstad, G. and Eijsink, V.G. (2016) Proteomic investigation of the secretome of *Cellvibrio japonicus* during growth on chitin. *Proteomics* **16**, 1904–1914 <https://doi.org/10.1002/pmic.201500419>
- 59 Vincent, F., Molin, D.D., Weiner, R.M., Bourne, Y. and Henrissat, B. (2010) Structure of a polyisoprenoid binding domain from *Saccharophagus degradans* implicated in plant cell wall breakdown. *FEBS Lett.* **584**, 1577–1584 <https://doi.org/10.1016/j.febslet.2010.03.015>
- 60 Kurth, J.M., Brito, J.A., Reuter, J., Flegler, A., Koch, T., Franke, T. et al. (2016) Electron accepting units of the diheme cytochrome *c* TsdA, a bifunctional thiosulfate dehydrogenase/tetrathionate reductase. *J. Biol. Chem.* **291**, 24804–24818 <https://doi.org/10.1074/jbc.M116.753863>
- 61 Arslan, E., Schulz, H., Zufferey, R., Kunzler, P. and Thony-Meyer, L. (1998) Overproduction of the *Bradyrhizobium japonicum* c-type cytochrome subunits of the cbb3 oxidase in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **251**, 744–747 <https://doi.org/10.1006/bbrc.1998.9549>
- 62 Moore, G. and Pettigrew, G. (1990) *Cytochromes c: Evolutionary, Structural and Physicochemical Aspects*, Springer-Verlag, Berlin Heidelberg
- 63 Holm, L. (2019) Benchmarking fold detection by DALI Lite v.5. *Bioinformatics* **35**, 5326–5327 <https://doi.org/10.1093/bioinformatics/btz536>
- 64 Hearing, H.A., Weiner, J.H. and Armstrong, F.A. (1997) Direct detection and measurement of electron relays in a multicentered enzyme: voltammetry of electrode-surface films of *E. coli* fumarate reductase, an iron–sulfur flavoprotein. *J. Am. Chem. Soc.* **119**, 11628–11638 <https://doi.org/10.1021/ja9723242>
- 65 Yates, N.D., Dowsett, M.R., Bentley, P., Dickenson-Fogg, J.A., Pratt, A., Blanford, C.F. et al. (2020) Aldehyde-mediated protein-to-surface tethering via controlled diazonium electrode functionalization using protected hydroxylamines. *Langmuir* **36**, 5654–5664 <https://doi.org/10.1021/acs.langmuir.9b01254>
- 66 Achmann, F.L., Sorlie, M., Skjak-Braek, G., Eijsink, V.G. and Vaaje-Kolstad, G. (2012) NMR structure of a lytic polysaccharide monoxygenase provides insight into copper binding, protein dynamics, and substrate interactions. *Proc. Natl Acad. Sci. U.S.A.* **109**, 18779–18784 <https://doi.org/10.1073/pnas.1208822109>
- 67 Borisova, A.S., Isaksen, T., Dimarogona, M., Kognole, A.A., Mathiesen, G., Varnai, A. et al. (2015) Structural and functional characterization of a lytic polysaccharide monoxygenase with broad substrate specificity. *J. Biol. Chem.* **290**, 22955–22969 <https://doi.org/10.1074/jbc.M115.660183>

- 68 Zouraris, D., Dimarogona, M., Karnaouri, A., Topakas, E. and Karantonis, A. (2018) Direct electron transfer of lytic polysaccharide monoxygenases (LPMOs) and determination of their formal potentials by large amplitude Fourier transform alternating current cyclic voltammetry. *Bioelectrochemistry* **124**, 149–155 <https://doi.org/10.1016/j.bioelechem.2018.07.009>
- 69 Crouch, L.J., Labourel, A., Walton, P.H., Davies, G.J. and Gilbert, H.J. (2016) The contribution of non-catalytic carbohydrate binding modules to the activity of lytic polysaccharide monoxygenases. *J. Biol. Chem.* **291**, 7439–7449 <https://doi.org/10.1074/jbc.M115.702365>
- 70 Loose, J.S., Forsberg, Z., Kracher, D., Scheiblbrandner, S., Ludwig, R., Eijlsink, V.G. et al. (2016) Activation of bacterial lytic polysaccharide monoxygenases with cellobiose dehydrogenase. *Protein Sci.* **25**, 2175–2186 <https://doi.org/10.1002/pro.3043>
- 71 Bennati-Granier, C., Garajova, S., Champion, C., Grisel, S., Haon, M., Zhou, S. et al. (2015) Substrate specificity and regioselectivity of fungal AAg lytic polysaccharide monoxygenases secreted by *Podospora anserina*. *Biotechnol. Biofuels* **8**, 90 <https://doi.org/10.1186/s13068-015-0274-3>
- 72 Kittl, R., Kracher, D., Burgstaller, D., Haltrich, D. and Ludwig, R. (2012) Production of four *Neurospora crassa* lytic polysaccharide monoxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol. Biofuels* **5**, 79 <https://doi.org/10.1186/1754-6834-5-79>
- 73 Zhou, M., Diwu, Z., Panchuk-Voloshina, N. and Haugland, R.P. (1997) A stable nonfluorescent derivative of resorufin for the fluorometric determination of trace hydrogen peroxide: applications in detecting the activity of phagocyte NADPH oxidase and other oxidases. *Anal. Biochem.* **253**, 162–168 <https://doi.org/10.1006/abio.1997.2391>
- 74 Petrovic, D.M., Varnai, A., Dimarogona, M., Mathiesen, G., Sandgren, M., Westereng, B. et al. (2019) Comparison of three seemingly similar lytic polysaccharide monoxygenases from *Neurospora crassa* suggests different roles in plant biomass degradation. *J. Biol. Chem.* **294**, 15068–15081 <https://doi.org/10.1074/jbc.RA119.008196>
- 75 Kracher, D., Forsberg, Z., Bissaro, B., Gangl, S., Preims, M., Sygmund, C. et al. (2020) Polysaccharide oxidation by lytic polysaccharide monoxygenase is enhanced by engineered cellobiose dehydrogenase. *FEBS J.* **287**, 897–908 <https://doi.org/10.1111/febs.15067>
- 76 Klock, H.E. and Lesley, S.A. (2009) The polymerase incomplete primer extension (PIPE) method applied to high-throughput cloning and site-directed mutagenesis. *Methods Mol. Biol.* **498**, 91–103 [https://doi.org/10.1007/978-1-59745-196-3\\_6](https://doi.org/10.1007/978-1-59745-196-3_6)
- 77 Kabsch, W. (2010) Xds. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 <https://doi.org/10.1107/S0907444909047337>
- 78 Potterton, L., Agirre, J., Ballard, C., Cowtan, K., Dodson, E., Evans, P.R. et al. (2018) CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr. D Struct. Biol.* **74**, 68–84 <https://doi.org/10.1107/S2059798317016035>
- 79 Sheldrick, G.M. (2008) A short history of SHELX. *Acta Crystallogr. A* **64**, 112–122 <https://doi.org/10.1107/S0108767307043930>
- 80 Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 <https://doi.org/10.1107/S0907444910007493>
- 81 Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 <https://doi.org/10.1107/S0907444996012255>
- 82 Wood, T.M. (1988) Preparation of crystalline, amorphous, and dyed cellulase substrates. *Methods Enzymol.* **160**, 19–25 [https://doi.org/10.1016/0076-6879\(88\)60103-0](https://doi.org/10.1016/0076-6879(88)60103-0)