Production and Spectroscopic Characterisation of Lytic Polysaccharide Monooxygenases

Glyn R. Hemsworth[1,2†], Luisa Ciano[3†], Gideon J. Davies[3], Paul H. Walton[3*]

[1]School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, United Kingdom
[2]Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds LS2 9JT, United Kingdom
[3]Department of Chemistry, University of York, York YO10 5DD, United Kingdom

*to whom correspondence should be addressed: paul.walton@york.ac.uk
†these authors contributed equally to this chapter.

Contents

Abstract

Lytic Polysaccharide Monooxygenases (LPMOs, also known as PMOs) are a recently discovered family of enzymes that play a key role in the breakdown of polysaccharide substrates. The ability of LPMOs to introduce chain breaks, using an oxidative mechanism, has particularly attracted attention as the world seeks more cost-effective and environmentally friendly ways of producing second generation biofuels for the future. LPMOs are mononuclear copper dependent enzymes and have an unusual active site which includes the N-terminal residue of the protein in the copper coordination sphere. This N-terminal histidine sidechain is also methylated in fungal enzymes, the molecular reason for which is still a debated topic. The production of these enzymes poses several challenges if we are to understand their chemical mechanisms. Here we describe the methods that have been used in the field to produce LPMOs and provide information on the workflows that we use for our Electron Paramagnetic Resonance (EPR) Spectroscopy experiments. EPR has been a particularly powerful tool in the study of these enzymes and our objective with this chapter is to provide some helpful information for researchers for whom this technique might be daunting or theoretically difficult to access.

Keywords: LPMO, biofuel, bioinorganic chemistry, EPR, protein production

1. Introduction

The production of second generation biofuels from lignocellulosic feedstocks has been hotly pursued for the last decade as a more sustainable fuel source in comparison to oil and first-generation biofuels (several reviews are available, e.g. (Merino and Cherry, 2007; Naik *et al*., 2010; Saini *et al*., 2015)). One of the major challenges towards the realisation of the cost-effective production of these fuels, however, has been the nature of the substrate, cellulose, and its resistance to degradation (Himmel *et al*., 2007). The study of enzymes involved in the breakdown of this substrate was pioneered by Reese and Mandels through the 1950s (Mandels and Reese, 1957; Reese *et al*., 1950), 60s (Mandels and Reese, 1960; Mandels *et al*., 1962) and 70s (Mandels *et al*., 1971, 1974), and we continue to make new discoveries in this area to this day (See recent reviews by Bischof *et al*. (2016) and Payne *et al*. (2015)). What is now viewed as the "classical" model of enzymatic cellulose deconstruction requires the combined action of essentially three distinct enzymatic activities to allow the liberation of single glucose molecules from this challenging substrate (Reviewed in Payne *et al*. (2015)). The glycoside hydrolases involved include endoglucanases, cellobiohydrolases and β-glucosidases, which synergise with one another to bring about this molecular conversion for metabolism in nature. These naturally occurring enzymes, however, despite their efficiency in degrading their substrates, have arguably been too slow to make the enzymatic conversion of biomass cost effective in the biorefinery. As such, new enzymatic activities have been sought after (as reviewed by Harris *et al*. (2014) and Horn *et al.* (2012)). In this context, Reese, in his original 1950 publication, suggested that there was likely an additional activity required to disrupt the crystalline structure of cellulose to allow glycoside hydrolases access to the substrate (Reese *et al*., 1950). This activity is now thought to be provided by a family of enzymes known as Lytic Polysaccharide Monooxygenases (LPMOs) (See Beeson *et al*. (2015), Ciano *et al*. (2018) and Hemsworth *et al*. (2015) for recent reviews).

Prior to the demonstration of LPMO activity in 2010 (Vaaje-Kolstad *et al*., 2010), LPMOs were defined in two families in the carbohydrate active enzymes database (www.cazy.org) (Lombard *et al*., 2014) – Glycoside hydrolase-61 (GH61) and carbohydrate binding module-33 (CBM33) (See Morgenstern *et al*. (2014) for a review). In 2010, Harris *et al*. (2010) had demonstrated that proteins classified as GH61s from *Thielavia terrestris* and *Thermoascus aurantiacus* were able to reduce the enzyme load required for the depolymerisation of cellulose 2-fold in a metal dependent manner. The biochemical basis for this became clear when Vaaje-Kolstaad *et al.* showed that the structurally related protein, CBP21 from *Serratia marcescens* (classified as a CBM33 in CAZy), was able to oxidatively induce chain breaks into chitin in a reducing agent and O_2 dependent manner, thereby establishing CBM33s as LPMOs for the first time (Vaaje-Kolstad *et al*., 2010). Subsequent work demonstrated that GH61s were also oxidative enzymes and that they required a copper cofactor to function (Phillips *et al*., 2011; Quinlan *et al*., 2011). LPMOs have now been discovered from diverse organisms across the tree of life and have been implicated in the deconstruction of a range of polysaccharides beyond cellulose and chitin (Agger *et al*., 2014; Bennati-Granier *et al*., 2015; Borisova *et al*., 2015; Couturier *et al*., 2018; Frommhagen *et al*., 2015; Lo Leggio *et al*., 2015; Simmons *et al*., 2017; Vu *et al*., 2014a). Due to their redox function, LPMOs are now defined in CAZy as auxiliary activities (AA) (Levasseur *et al*., 2013) with six families currently classified - AA9, AA10, AA11, AA13, AA14 and AA15.

LPMOs are unusual in being monooxygenases which utilise only a single copper ion in their active site (see Solomon *et al*. (2014) for a useful review on copper dependent enzymes). Indeed, there have been recent suggestions that LPMOs may, under certain conditions, also function as peroxygenases, but further work is still required to determine the true physiological co-substrate for LPMOs (Bissaro *et al*., 2017; Kuusk *et al*., 2018). LPMOs bind their single copper cofactor in a motif that has been dubbed the "histidine brace" (Quinlan *et al*., 2011), an arrangement of two histidine residues, one of which is the N-terminal residue, which coordinate the copper ion in a T-shaped geometry (Figure 1) (see reviews by Ciano *et al*. (2018), Hemsworth *et al*. (2013a), Vaaje-Kolstad *et al*. (2017) and Vu & Ngo (2018)). In addition, for fungal enzymes at least, there is a modification to the N-terminal histidine sidechain in the form of $\boldsymbol{\tau}$-N-methylation which is only detected when proteins are expressed in fungal hosts (Frandsen *et al*., 2016; Lo Leggio *et al*., 2015; Li *et al*., 2012; Quinlan *et al*., 2011; Vu *et al*., 2014a, 2014b). The importance of this modification is still debated, but it is clear that LPMOs remain active in its absence when expressed in *e.g*. *Pichia pastoris* (Bennati-Granier *et al*., 2015; Bey *et al*., 2013; Couturier *et al*., 2018; Wu *et al*., 2013). The requirement for the N-terminal histidine in the mature protein has dictated the strategies that have been used to produce these enzymes which we will discuss in this chapter. We will also describe our workflows for conducting Electron Paramagnetic Resonance (EPR) spectroscopy on these enzymes. EPR has been a powerful tool in contributing towards unravelling the steps along the reaction coordinate catalysed by LPMOs. EPR is a conceptually challenging technique, however, which requires expertise to both successfully simulate the data and to interpret the resulting spin Hamiltonian parameters in a meaningful way to gain insights into the electronic aspects of the copper active site. We hope that the topics covered in this chapter will, therefore, prove useful not only for those working on LPMOs, but more broadly for researchers interested in studying diverse copper dependent enzymes.

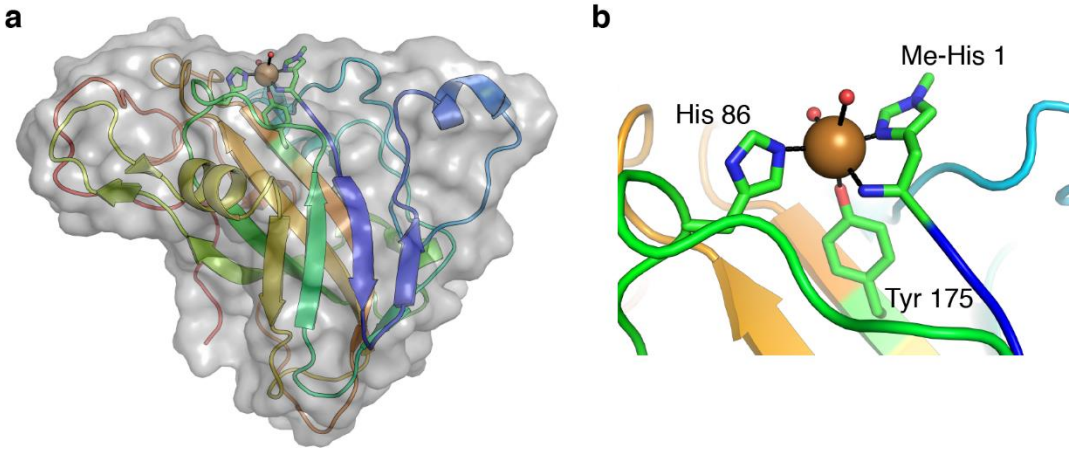

Figure 1. Example structure of an AA9 LPMO. **a,** overall structure of *Thermoascus aurantiacus* AA9 LPMO shown in cartoon coloured from N- (blue) to C-terminus (red). **b,** Close up view of the active site copper ion (orange sphere) coordinated by the histidine brace which includes the N-terminal methylated histidine of the protein.

2. Expression and Purification of LPMOs

LPMOs have been identified from diverse organisms including bacteria (Vaaje-Kolstad *et al*., 2005a, 2010), viruses (Chiu *et al*., 2015), fungi (Karkehabadi *et al*., 2008; Langston *et al*., 2011;

Phillips *et al*., 2011; Quinlan *et al*., 2011; Saloheimo *et al*., 1997) and more recently higher eukaryotes (Sabbadin *et al*., 2018). This can pose challenges towards the choice of expression strategy in order to produce pure proteins to be studied. This choice is compounded by the fact that these proteins are typically secreted in their native hosts, leading to the generation of the N-terminal histidine integral to the active site (for reviews please see Hemsworth *et al*. (2013a) and Vaaje-Kolstad *et al*. (2017)). Endogenously, LPMOs will often also contain disulfide bonds, and eukaryotic family members can be further modified by N- and O-glycosylation, as well as methylation of the N-terminal histidine. The choice of expression system must therefore be tailored towards the target enzyme and can have a significant impact on the protein yield, and potentially on the activity of the enzyme though this has not been investigated in great detail.

2.1 Recombinant expression and purification from the *E. coli* periplasm

Escherichia coli is the workhorse of the academic lab when it comes to protein production. Bacterial LPMOs of the AA10 class are therefore most commonly purified from this organism. The most common approach has been to secrete the resultant protein to the periplasm (Courtade *et al*., 2017; Forsberg *et al*., 2014a; Hemsworth *et al*., 2013b; Vaaje-Kolstad *et al*., 2005a, 2010). It is also possible, however, to express these enzymes in the cytoplasm using specialised *E. coli* strains (Forsberg *et al.*, 2011, 2014b; Gregory *et al*., 2016). These have largely been our methods of choice for LPMO production and though fungal enzymes are rarely produced using these strategies, we were able to express and purify an AA11 from *Aspergillus oryzae* using a periplasmic secretion system in *E. coli* (Hemsworth *et al*., 2014).

In order to target the protein to the periplasm, we carefully design our constructs for cloning into a pelB containing vector and typically use HiFi DNA Assembly (New England Biolabs), In-Fusion (Takara) or PIPE cloning (Klock and Lesley, 2009). This strategy allows the target gene to be placed directly after the signal sequence without the addition of any residues from use of restriction enzymes and negates the need to use site directed mutagenesis later to remove any undesired residues. Others have found success using the CBP21 leader peptide as the signal sequence (Crouch *et al*., 2016; Gardner *et al*., 2014), and one should always consider screening other potential leader peptides in order to improve the chances of obtaining good yields of soluble protein in the periplasm. In addition, careful thought needs to be put towards the use of affinity tags using the periplasmic secretion strategy. An affinity tag would typically be placed at the C-terminus in this case. We have tended to avoid using His-tags for fear that they might interfere with copper binding and produce unwanted copper binding sites during our EPR analyses and also, potentially, reactivity studies. If a His-tag is used to purify the protein and not removed prior to such an analysis, then careful controls must be performed as Cu binding to the tag cannot be excluded if the metal ion is added in excess. EPR can offer a quick and powerful way to check the sample for the presence of any additional/adventitious binding sites (see below). Other affinity tags can also be used of course, with a strep-tag recently used for the purification of the latest LPMO discovered in family AA15 providing a prominent example (Sabbadin *et al*., 2018).

Once the desired constructs are obtained and confirmed as correct by Sanger sequencing, plasmids are transformed typically into BL21(DE3) *E. coli* cells as a first port of call. Other *E. coli* strains should also be considered. Transformants are plated out on selective LB agar plates and allowed to grow at 37 °C overnight. A 5 mL LB liquid culture is then inoculated using

a single colony the next day and the cells are allowed to grow at 37 °C in a shaker at 180 rpm throughout the day. Once the culture is visibly turbid, 500 µL is used to inoculate a 50 mL culture which is left to grow overnight in the same conditions. The next day, 2 L baffled flasks each containing 500 mL of LB are then inoculated using 5 mL each from the overnight culture and these are incubated at 37 °C, shaking at 180 rpm. The OD_{600} is monitored closely and once this has reached 0.6 the temperature is then typically cooled to 16 °C for about an hour. Protein expression can then be induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and the cultures are left overnight. The next day cells are harvested by centrifugation at 8,000 x g, the spent media discarded and cell pellets can be frozen at -80 °C until ready for subsequent use.

We use a modified periplasmic preparation protocol rather than the cold osmotic shock protocol (Neu and Heppel, 1965) that is often used as we have found that we recovered better yields of protein using this approach. Typically, cell paste is thoroughly resuspended in 3x volumes of 20% sucrose (20 mM Tris can also be included to stabilise pH). To this, 40 µL of 10 mg/mL lysozyme is added for every gram of cell paste being used and the suspension is then left on ice for 1 hour. 60 µL of 1 M $MgSO_4$ is then added for every gram of cell paste being used and the cells are left on ice for an additional 20 minutes. Cell debris is then harvested by centrifugation at 10,000 x g for 20 minutes and the supernatant is removed as Fraction I to a fresh container and kept on ice. The pellet is re-suspended in ice cold water and this is also left on ice for a further 1 hour (some DNAse can be added at this point to try and reduce the viscosity of the solution produced at the end of the process). The cell debris is again removed by centrifugation at 10,000 x g for 20 minutes and the supernatant is removed as Fraction II. Fraction I and II can be checked at this point to see whether the LPMO is abundantly present in one fraction over the other, however, usually both fractions are combined for subsequent purification.

If using an affinity tag, standard chromatographic approaches can be used for the purification of LPMO. We typically perform purifications using FPLC instruments using an affinity chromatography step for the initial purification and gel filtration for a final “polishing” step. In the absence of a tag we would typically aim to purify via ion exchange chromatography followed by gel filtration (see Figure 2a and b for example gels). We would recommend using Good’s buffers throughout any purification, these represent biologically compatible buffers which should not interact significantly with biologically relevant cations such as copper (Good *et al*., 1966). If attempting to measure copper binding to LPMO it is imperative that all buffers are thoroughly de-metallated prior to the final step of the purification. This is typically achieved by cycling the buffer through a 10 mL chelex column (Bio-Rad) for 24 hours before being used in gel filtration. The protein should also be thoroughly treated with EDTA (typically at 10 mM concentration or higher for at least an hour) to remove any metal that might already be bound since LPMOs have very high affinity for copper. Indeed, it is probably this high affinity for copper that resulted in low level activity, which could have been mistaken for hydrolytic activity, during early studies on GH61s (Karlsson *et al*., 2001). The high affinity of the enzyme for copper allows it to scavenge its cofactor even when present at very low concentration and could also have contributed to the initial confusion over the identity of the correct metal co-factor following the initial structure determinations for these enzymes (Karkehabadi *et al*., 2008; Vaaje-Kolstad *et al*., 2005b, 2010).

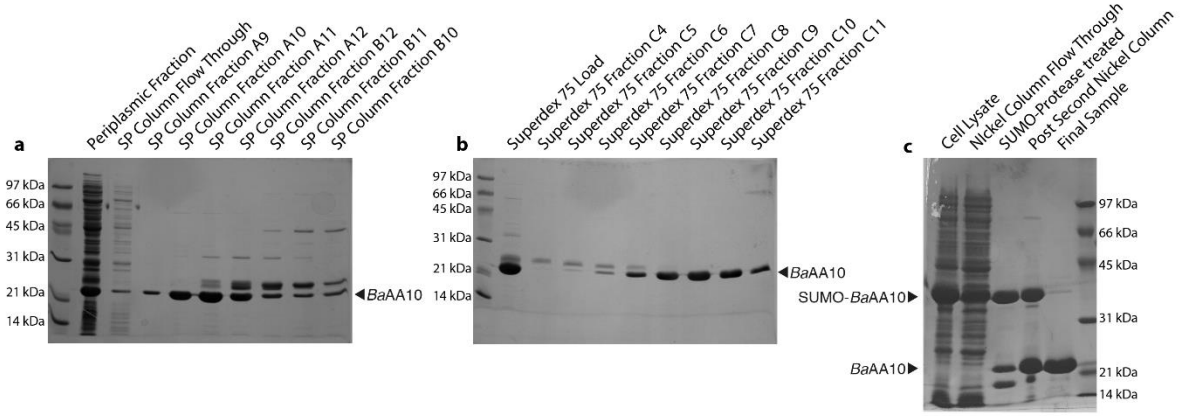

Figure 2. Example gels for LPMO Purification strategies. **a,** gel showing initial purification of *Ba*AA10 from the periplasm using an SP ion exchange column. **b,** subsequent purification of BaAA10 using a Superdex 75 size exclusion column to yield the final sample *(*Hemsworth *et al*., 2013b). **c,** SDS-PAGE gel showing stages from the purification of *Ba*AA10 using a SUMO-tagged construct (Gregory *et al*., 2016).

If an apo-enzyme is not desired, but rather a copper loaded sample is to be used for study, prior to the final gel filtration step, $CuCl_2$ can be added in a slight excess (typically 1.5x) in order to load the protein with copper prior to gel filtration. The sample is then typically passed through a 16/600 Superdex 75 column (GE Healthcare) to remove any excess copper, giving a highly purified, copper loaded sample ready for study. In order to store protein, we typically concentrate to 10 mg/mL or more using a 3, 5 or 10 kDa Vivaspin column depending on the size of the LPMO (avoid cellulose-based membranes concentrators). The sample is then flash frozen in liquid nitrogen for storage at -80 °C.

Before discussing other strategies for LPMO production in *E. coli* and other organisms, there is another option for protein secretion to obtain mature LPMOs - to secrete the protein from gram positive bacteria into the culture medium (Nakagawa *et al*., 2015). We have not used this approach ourselves, but this could be an excellent strategy for large scale LPMO production with industry often using gram-positive expression systems for the production of many stable and industrially-useful enzymes.

2.2 Recombinant expression and purification from the *E. coli* cytoplasm

Given that LPMOs are typically extracellular enzymes that contain disulfide bonds, the periplasmic expression protocol described above often represents the first port of call for protein production (Courtade *et al*., 2017; Forsberg *et al*., 2014a; Hemsworth *et al*., 2013b; Vaaje-Kolstad *et al*., 2005a, 2010). This is not the only option however, and there are reports of LPMOs being expressed intracellularly using cleavable tags at the N-terminus to aid purification (Forsberg *et al*., 2011, 2014b; Gregory *et al*., 2016). In order to achieve this, the protease to be used to remove that tag must be carefully selected in order to leave the native N-terminal histidine available for copper binding. Factor Xa (Forsberg *et al*., 2014b) and enterokinase are two such proteases. We prefer the use of a SUMO-tag (Champion-pET-SUMO, Invitrogen) which can be removed using SUMO-protease (Invitrogen) (Gregory *et al*., 2016). This method has the advantage that the SUMO-tag can act as a solubility enhancing tag in addition to providing an N-terminal histidine tag to aid in the purification of the protein. SUMO-protease recognises the tertiary structure of the SUMO domain and cleaves directly

after a di-glycine motif which can be placed directly upstream of the N-terminal His coding region, so that cleavage results in the release of the native protein. Constructs can be carefully designed using the cloning strategies discussed above in order to achieve this without the use of restriction enzymes. This strategy was used to improve the yield of the AA10 from *Bacillus amyloliquefaciens* as much as 3-fold (Gregory *et al*., 2016). While this protein lacked disulfide bonds which would have aided in the correct folding of the protein intracellularly, we have since used this strategy to express disulfide containing LPMOs intracellularly using specialised cells such as the NEB Express T7 cells (New England Biolabs) (Unpublished data).

The specific protocol for protein expression in *E. coli* is much the same as that discussed above with cells grown in shake flasks and the temperature typically being lowered to 16 or 20 °C upon IPTG induction to ensure correct protein folding. Cells are harvested the same way by centrifugation at 8,000 x g, with spent media being discarded and cell pellets frozen at -80 °C until ready for lysis.

Since the protein is now present in the cytoplasm, there is no need to perform the laborious periplasmic prep protocol in order to release the protein from the cells. Cells are simply resuspended in Buffer A (typically 50 mM HEPES pH 7, 250 mM NaCl, 30 mM imidazole) and can be lysed by one's preferred method (typically sonication or French press is used in our laboratories). The cell lysate can then be loaded directly onto a nickel resin to which the His_6-SUMO-tagged LPMO will bind (Figure 2c). Protein is then released by the application of a gradient from 0 to 100% buffer B (Buffer A + 300 mM imidazole) and 1.8 mL fractions are collected across the gradient. Peak fractions containing His_6-SUMO-LPMO are combined ready for tag removal. In order to prepare for tag cleavage, it is important to remove excess imidazole and NaCl from the sample which can be detrimental to SUMO-protease function. This can be achieved by buffer exchange on a 10 kDa MW cut-off concentrator, using a desalting column or by dialysis. The protease buffer should also contain a reducing agent to ensure optimal SUMO-protease activity since it is a cysteine protease. We, therefore, typically buffer exchange our samples into 20 mM Tris pH 8, 5 mM β-mercaptoethanol. Small scale tag-cleavage reactions should be set up the first time that this strategy is used to try to obtain optimal ratio of protease to SUMO-LPMO which can then be scaled up to the whole sample. We typically perform tag removal at a SUMO-LPMO concentration of 1-2 mg/mL overnight at 4 °C.

Once the cleavage has been performed, the released tag, any uncleaved protein and the SUMO-protease are removed from the sample by passing it back down a nickel column ensuring that the flow through, which will contain the mature LPMO, is collected (Figure 2c). The differences in the LPMO's affinity for copper over nickel ensure that there is not transmetallation of the enzyme. The enzyme sample can then be concentrated, copper loaded or thoroughly EDTA treated as described above, before a final polishing step using a 16/600 Superdex 75 column (GE Healthcare) to yield the final sample (Figure 2c).

2.3 Recombinant expression and purification from yeast

For many LPMOs, expression in *E. coli* has not been possible and so laboratories have turned to other expression systems for challenging targets. *Pichia pastoris* is generally the most common yeast expression system used for protein production in academic laboratories and has been used for the production of several LPMOs (Kittl, Kracher, Burgstaller, Haltrich, &

Ludwig, 2012; Bennati-Granier *et al*., 2015; Tanghe *et al*., 2015). While *P. pastoris* lacks the machinery required to perform the methylation of the N-terminal histidine found in fungal LPMOs, it maintains the ability to glycosylate proteins which may be important to ensure the correct folding of LPMOs from eukaryotic organisms.

The most common vector used for LPMO production in *P. pastoris* is the commercially available pICZα (Invitrogen) (Bennati-Granier *et al*., 2015; Couturier *et al*., 2018; Kittl *et al*., 2012), although other vectors such as the pPpT4 vector can also be used (Tanghe *et al*., 2015). These are shuttle vectors which allow all of the cloning to be performed in *E. coli* for speed and ease before the vector is introduced into the *P. pastoris* host for protein expression. As for expression in *E. coli*, similar factors need to be considered during construct design in *P. pastoris*. The vectors contain signal sequences to direct the protein for secretion and so the LPMO sequence must be carefully cloned such that the codon for the N-terminal histidine is immediately placed after the cleavage site for the signal peptide. Alternatively, if the protein is from a fungal source the native leader sequence of the LPMO gene can also be used rather than the signal peptide available in the vector (Kittl *et al*., 2012; Tanghe *et al*., 2015).

The target gene is typically placed under the control of the AOX1 promoter in these vectors, allowing induction of protein expression using methanol (Bennati-Granier *et al*., 2015; Couturier *et al*., 2018; Kittl *et al*., 2012). Cells are typically grown in defined media based on the protocol provided by Invitrogen so we will not re-describe this here. Thanks to the use of the signal sequence the protein is directed out of the cell into the media and so the protein needs to be directly purified from here. Affinity tags can be incorporated at the C-terminus of the protein to assist in the purification, but in many cases the protein has been purified without the use of an affinity tag, for example Kittle *et al* (2012) used a multi-step purification process involving hydrophobic chromatography, ion-exchange chromatography and finally size-exclusion chromatography.

Copper loading procedures are much the same as those described above for LPMOs expressed in *E. coli*, with the addition of a small excess of copper prior to the final gel-filtration step in the purification. If the protein is to be used for crystallography, it is often necessary to deglycosylate it. In our laboratories this step is typically achieved using Endoglycosidase H (New England Biolabs) following the manufacturer's protocols. If, however, the protein is to be used in assays, we typically do not deglycosylate the protein as this can affect both protein stability and, potentially, protein-substrate interactions.

2.4 Recombinant expression and purification from fungi

The majority of LPMOs that have been studied to date are found in fungi and so there are clear advantages to expressing these proteins in fungal hosts. These include correct insertion of methylation on the N-terminal histidine sidechain as well as protein secretion into the extracellular milieu following correct glycosylation and protein folding. Fungal recombinant expression systems, however, are far less well established in academic laboratories. We have, therefore, benefited greatly from our long-term collaboration with Novozymes who use a highly specialised *Aspergillus oryzae* expression system for LPMO production on a large scale (Frandsen *et al*., 2016; Lo Leggio *et al*., 2015; Quinlan *et al*., 2011). The high yields offered by this system are able to meet the high sample demands for EPR (see below). These systems are not readily available for use by others, however, and so examples of recombinant LPMO

expression from fungal sources are harder to find in the literature. Phillips *et al* (2011) early on were able to purify *Neurospora crassa* LPMOs directly from secretomes when the fungus was grown on cellulose. Since then, the Marletta lab and their colleagues have been able to use a *N. crassa* expression system to produce both *N. crassa* LPMOs (Vu *et al*., 2014b) and LPMOs from other fungi (Span *et al*., 2017). Using this approach, a knock-in system is used to insert the LPMO gene downstream of a glyceraldehyde-3-phosphate dehydrogenase promoter (GAP) (Bardiya and Shiu, 2007). The DNA is then incorporated into the *N. crassa* genome by recombination and colonies in which the gene has been successfully incorporated need to be screened for LPMO production. As for *Pichia pastoris*, the LPMO should be secreted into the growth media allowing direct purification from the cleared media. Affinity tags can be incorporated at the C-terminus of the protein to assist in this respect if they do not interfere with subsequent planned experiments.

3. Electron Paramagnetic Resonance Spectroscopy

Once pure protein is obtained, a key technique that has been used in the field to gain insight into the copper environment in LPMOs is Electron Paramagnetic Resonance (EPR) spectroscopy. The fact that copper in its oxidised (+2) state (S=½, Kramers' doublet) has a single unpaired electron makes it EPR active. Therefore, by collecting EPR spectra from pure LPMOs it has been possible to gain additional insight into the electronic state of the copper centre in the resting protein. In addition, upon the inclusion of polysaccharide substrates in the samples, it is sometimes possible to observe marked changes in the EPR spectra (Figure 3) (Bissaro *et al*., 2018; Borisova *et al*., 2015; Frandsen *et al*., 2016; Simmons *et al*., 2017), suggesting significant changes occur at the copper active site in the presence of substrate, which might be of key importance to activating oxygen and hence catalysis. In order to interpret these spectra and understand what the changes in the EPR mean, it is necessary to simulate the data. While there are algorithms that will attempt to simulate the data automatically, these often fall short of a truly meaningful simulation. Optimal EPR simulation is, therefore, still best performed by an expert EPR spectroscopist and requires considerable patience and knowhow in order to obtain the best possible outcome. Here we describe the steps taken to prepare samples for EPR analysis and provide information on how we interpret our EPR spectra.

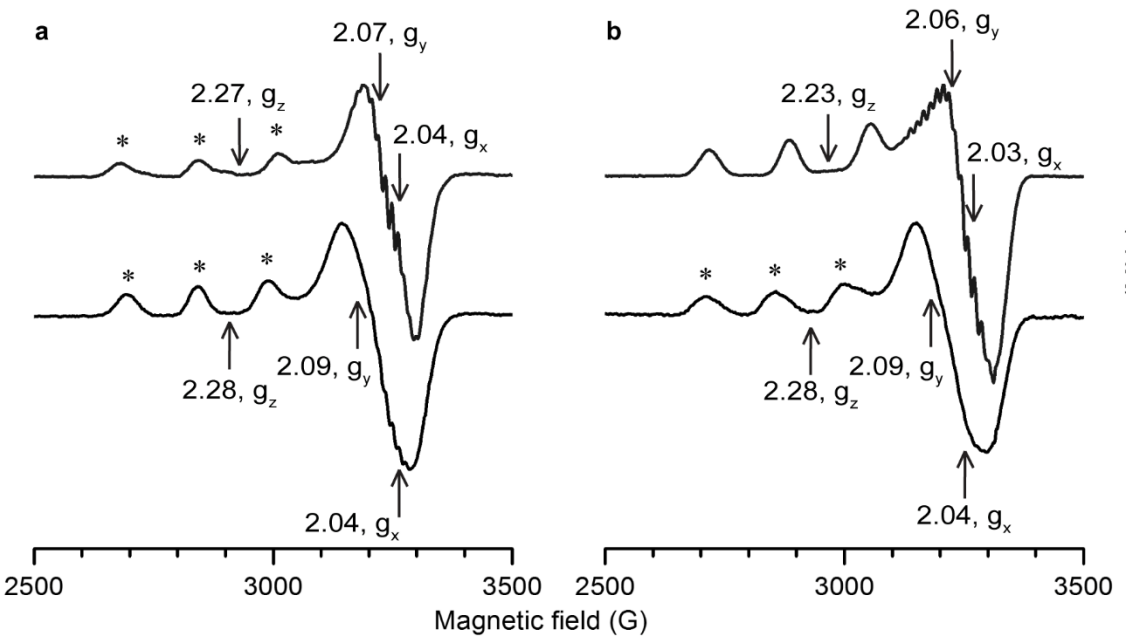

Figure 3: Example of changes observed in the X-band CW EPR spectra of an AA9 upon addition of substrate. **a,** *Ls*AA9A–G6 (top) and *Ls*AA9A with no cellohexaose (bottom) in low chloride conditions. **b,** *Ls*AA9A–G6 (top) and *Ls*AA9A with no cellohexaose (bottom) in high chloride conditions. * indicates signals from copper site where water is the exogenous ligand, other signals from species with chloride as exogenous ligand. Figure adapted from Frandsen *et al* (2016).

3.1 EPR sample preparation for LPMOs

Most EPR spectroscopy on LPMOs has been performed in frozen solution at X-band frequencies (ca 9.3 GHz). At these frequencies, sample holders are typically quartz tubes with roughly 3 to 5 mm internal diameters. The height of the resonator cavity will be of the order of 10-15 mm, meaning that ideally a sample volume of ca 300-600 µL is required. In reality, by optimising the collection parameters, good EPR spectra can be obtained with much smaller volumes, in the order of 80 to 120 µL with a concentration of at least 0.15 mM Cu. Albeit often seen by biochemists as a highly demanding technique in terms of protein consumption, EPR is non-destructive, and often samples can be fully recovered and re-used after the experiments. In our hands, the repeated cooling/warming cycles that a sample typically experiences during collection of EPR spectra do not lead to protein degradation. Obtaining a signal with good signal-to-noise is a function of several experimental variables including the capabilities of the EPR spectrometer, the temperature at which the spectra are collected, the number of scans and the experimental settings. However, good spectra with excellent signal-to-noise ratios can be obtained with sample concentrations and volumes reported above at 100 to 170 K with acquisition settings reported in Table 1. Of particular note is the power setting which must be at a level which does not cause saturation of the absorbance, especially in concentrated samples. This level should be such that relaxation rates are greater than absorption rates. For copper spectroscopy, relatively high power settings can be used because of the high relaxation rates associated with a metal centre. For detailed spectra, then higher sample concentrations may be needed, possibly up to 1-2 mM concentration. Great caution is needed at these concentrations however since face-to-face dimerization of LPMOs may occur with the result that two nearby Cu centres will change their coordination geometries or will lead to dipolar relaxation effects that perturb spin Hamiltonian values. Our

experience in this regard is that concentrations of ca 2 mM lead to spectral changes that could be associated with protein dimerization.

Table 1: Standard acquisition parameters for frozen solution X-band EPR spectra of LPMOs.

Centre field	3000 G
Sweep	2300 – 3700 G
Modulation amplitude	4 G
Modulation frequency	100 kHz
Time constant	163.8 msec
Sweep time	92 sec
Conversion time	90 msec
Power	5 – 10 mW
Scans	1 – 8
Temperature	150 – 170 K

EPR spectra collected in aqueous solution often require the use of a glassing agent such as glycerol or sucrose to minimise the formation of ice crystals. Such crystals can increase microwave scattering, reduce the microwave cavity quality factor, cause changes in protein structure, pH or ionic strength, or even shatter the microwave tube! For smaller tubes which are used with high frequency EPR this can be a particular problem. Tube breakages for X-band EPR tubes are less common, especially if thick wall tubes are used. For work with LPMO solutions, glycerol is typically the glassing agent of choice (10-50% *v/v*), although sucrose can also be used (2 mg/µL). When combined with very rapid freezing then the resulting sample usually forms a good glass. The cooling method needs to be carefully considered however in that very rapid freezing is required to prevent entropic driven processes from occurring at the metal centre. For instance, at room temperature the copper coordination may include the three atoms of the ligating histidine brace along with a fourth exogenous ligand. Cooling slowly to typical EPR collection temperatures may drive the formation of a five coordinate Cu site in which a further water molecule is recruited into the coordination sphere. The result would be that the EPR spectrum collected at low temperature is not representative of the solution phase structure. Therefore, very rapid freezing is required to 'trap' the room temperature species. This is best achieved by plunging the EPR sample tube into a cold liquid (e.g. methanol in dry ice (Chaplin *et al.*, 2016)) rather than into liquid nitrogen where—despite the very low temperature of the liquid—the resulting bubbles of N_2 insulate the tube from the liquid to prevent very rapid freezing. Notwithstanding the advantages of rapid cooling of samples, the experimenter needs to be aware that any heterogeneity in protein structure, where the interconversion is slower than the cooling timescale, may be trapped by the cooling process, such that sample heterogeneity is increased, the result of which is a severe loss of spectral resolution. It is not unusual for rapidly-cooled samples to have broader spectral peaks that those samples cooled less rapidly (Guzzi *et al.*, 2001). Necessarily, therefore, different cooling techniques need to be explored. In our experience, a happy compromise can be found for X-band LPMO samples when a sample at room temperature is plunged into the stream of cooling N_2 gas which cools the resonator. Also, on occasion for high concentration samples, it has been possible to collect spectra of LPMO solutions at room temperature. The slow tumbling rate of the protein in solution means that the spectrum is

still anisotropic, from which it is possible to link the spin Hamiltonian parameters at room temperature with those at low temperature (unpublished data).

The type of glassing agent that is employed in these studies is an important consideration in experimental design. The principal consideration is that glassing agents have a 'poly-ol' structure which can mimic that of the natural polysaccharide substrate of the enzyme and possibly interfere with the Cu active site by binding close to it. Accordingly, the dependency of the spectrum on the concentration of glassing agent should be established in any experiment. Unfortunately, an examination of the LPMO literature shows that such dependency experiments are relatively rare and only appear to have been done in a handful of cases (Sabbadin *et al*., 2018). When there is evidence to suggest that the glassing agent is interfering with the Cu centre, then it is recommended changing the agent to another (*e.g.* DMSO, Ficoll) or that the spectrum is collected without any glassing agent at all. The latter method appears to work surprisingly well for LPMO samples *(*Bissaro *et al.,* 2018; Borisova *et al*., 2015; Chaplin *et al*., 2016; Forsberg *et al*., 2014b, 2014a, 2016).

One of the great advantages of EPR is that spectra can be obtained on a sample in any state of matter: liquid, solid, slurry, frozen glass or even a gas. For LPMOs, of course, since the substrates are usually insoluble polysaccharides, the interaction of the LPMO with its natural substrate can be studied directly. Moreover, introducing a solid phase sample into a solution to create a slurry, essentially removes the need for a glassing agent. Necessarily, the nature of the solid phase means that protein binding may be slow or incomplete and it is advisable to leave the slurry sample for some time before then acquiring the EPR spectra. In the majority of cases, at least part of the protein binds very rapidly (Frandsen *et al*., 2016; Simmons *et al*., 2017), but for some samples we have noticed that full binding can take up to several hours and/or cannot be achieved, so would recommend that samples are left incubating overnight in the absence of oxygen and/or a reducing agent. It is also important that the introduced sample is isotropically distributed within the EPR tube, otherwise alignment effects could complicate the spectrum. Beyond these considerations, it appears as if solid state slurries of LPMO solutions with their insoluble substrates are perfectly good samples for both CW and pulsed EPR studies (Bissaro *et al*., 2018; Borisova *et al*., 2015; Frandsen *et al*., 2016; Simmons *et al*., 2017).

3.2 Multi-frequency collection

EPR spectra of LPMOs in the Cu(II) states give spin Hamiltonian parameters in which there is either a typical 'type 2' Cu centre with large $|A_z|$ values (ca 400-600 MHz) (Borisova *et al*., 2015; Chaplin *et al*., 2016; Couturier *et al*., 2018; Forsberg *et al*., 2014a, 2016; Frandsen *et al*., 2016; Hansson *et al*., 2017; Hemsworth *et al*., 2014; Quinlan *et al*., 2011; Sabbadin *et al*., 2018; Simmons *et al*., 2017; Span *et al.*, 2017) or a more rhombic Cu(II) centre that has a reduced $|A_z|$ value but large $|A_x|$ and $|A_y|$ values (Bissaro *et al*., 2018; Forsberg *et al*., 2014b; Gregory *et al*., 2016; Hemsworth *et al*., 2013b). In both cases the perpendicular region of the spectrum cannot be unambiguously simulated from the X-band data alone. For this reason, if an in-depth analysis is required, collecting spectra at different frequencies is paramount. Depending on the kind of information desired, one can decide to use lower (*e.g*. S-band) or higher (*e.g*. Q-band) frequency than X-band, with the latter being usually the preferred method. A couple of examples of this approach are present in the LPMO literature (Couturier *et al*., 2018; Hansson *et al*., 2017), but it is easy to imagine that more and more will appear as

the fine details of the electronics at the Cu active site are investigated. For Q-band spectra, quartz tubes of ca 2 mm in diameter are employed, while the samples are typically 15-30 µL in volume at a concentration of copper of 1 mM or higher. The same, if not greater, caution in optimising the acquisition settings as for X-band should be exercised when collecting Q-band spectra. In our experience, obtaining good quality Q-band spectra, especially if protein concentration is an issue due to protein aggregation at high concentration or limited amount of protein available, is a non-trivial problem, but it can be sometimes overcome with patience and trials of different settings. Typical settings used for our experiments include centre field at 1150 mT, sweep width 150 mT, modulation width 0.6 – 1 mT, time constant 0.3 s, scan time 1 – 2 min, power 0.5 – 1 mW.

Collecting CW Q-band spectra has a further benefit in that EPR spectra of LPMOs can sometimes contain high field features (Bissaro *et al*., 2018). These features under simulation (see below) are very hard to assign unambiguously as they could arise either from large hyperfine coupling constants or from low principal *g* values. There are indeed cases in the literature where such features have been simulated with low *g* values. It is, however, well known in the wider EPR community that the high field signals may arise from a spectral artefact known as an overshoot feature, associated with large $|A_z|$ hyperfine coupling values (see figure 4). By collecting CW EPR spectra at different frequencies, the normal choice being Q-band, this misinterpretation can be avoided.

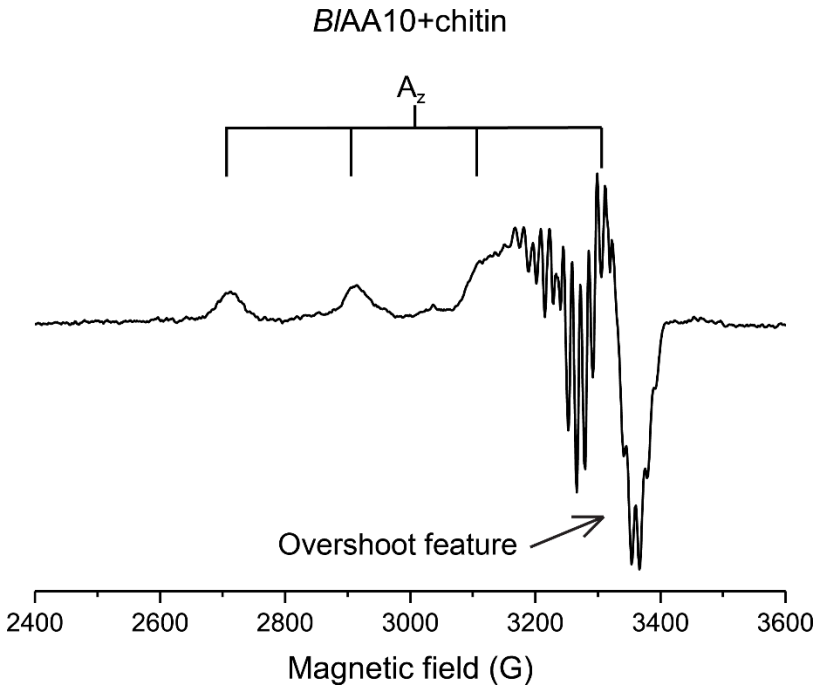

Figure 4: X-band CW EPR spectrum of an AA10 from *Bacillus licheniformis*, *Bl*AA10, after addition of squid pen chitin, showing the overshoot feature associated with the large $|A_z|$ value (unpublished data).

3.3 Simulation procedures

Once the CW spectra have been recorded, simulations have to be performed in order to gain information about the copper active site. Several simulation packages are available to users, some of them developed by EPR instruments manufacturers (*e.g.* the WinEPR and SimFonia

suites (Bruker)), others made available by researchers in the field (*e.g.* SpinCount (The Hendrich Metalloprotein Group) or EasySpin (Stoll and Schweiger, 2006)). Regardless of the programme used to carry out the simulations, some general considerations can be done on how to approach such a task, in particular for those who might not have encountered EPR before. Determining the spin Hamiltonian parameters for the parallel region of the spectrum is in general straightforward, as at least 3 of the 4 hyperfine peaks to the Cu can be clearly observed between 2500 and 3100 G in the X-band spectra of LPMOs. Careful analysis of this region of the spectrum can also reveal the presence of multiple species or adventitious binding sites, if a second set of hyperfine peaks is present either as individual peaks or as shoulders of the main component of the spectrum. Specifically, we would recommend particular care if the LPMO carries an affinity tag, as highlighted in section 2.1. In these circumstances, EPR provides a very powerful tool for the identification of second binding sites in amounts as low as 10% of the main species. As mentioned above, most LPMOs show a type 2 copper active site according to the Peisach-Blumberg (P-B) classification (Peisach and Blumberg, 1974), with SOMO mostly $d(x^2-y^2)$ in character and close to square planar geometry. While the P-B classification is a useful 'first-step' guide to the general structure of the copper ion in the active sites of LPMOs, we would caution against its in-depth use for the assignment of coordinating atoms (*e.g*. determining between CuN_2O_2 and CuN_3O coordination spheres), due to the multiple contributing factors to both g_z and $|A_z|$ values, the most problematic of which is the unknown degree of covalency in the metal-ligand bonds. In this regard, it is indeed possible to notice the high degree of variability in the published EPR parameters of LPMOs in a P-B plot (Figure 5 and Table 2) even for proteins that, from a structural point of view, show the same coordination environment (recently reviewed by Ciano *et al.* (2018) and Vu and Ngo (2018)). Finally, in the context of the broad interpretation of spin Hamiltonian parameters of LPMOs, there are reports of compressed trigonal bipyramidal structures with $d(z^2)$ SOMOs, but this assignment is not in accord with the spectral envelopes observed in all LPMO EPR spectra (Vu and Ngo, 2018).

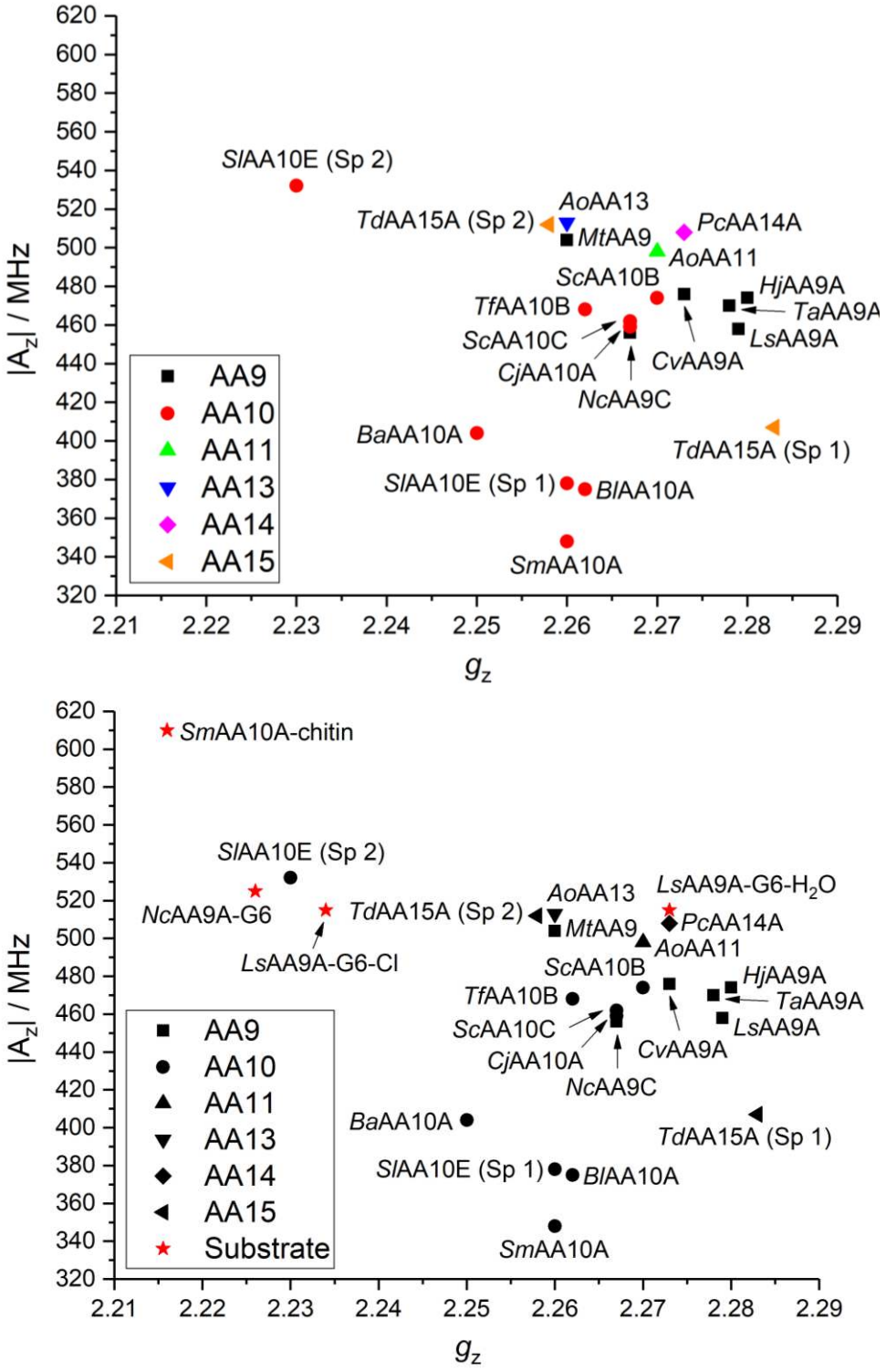

Figure 5: Peisach-Blumberg plots of published EPR data for LPMOs (data and references reported in Table 2). For clarity and consistency, all enzymes have been named using the CAZy abbreviation, where two letters (in italics) define the organism the protein originates from, followed by the Auxiliary Activity family and the specific enzyme, if reported and more than one are present in the same organism (*e.g. Ls*AA9A). Top: P-B plot of LPMOs divided by family, according to the legend on the plot. Bottom: P-B plot as above, with the parameters for substrate-bound enzymes included; "G6" is used for cellohexaose.

Table 2: Table of g_z and $|A_z|$ LPMO EPR values published to date, used for the plots in Figure 5.

LPMO	Reference	g_z	$\|A_z\|$ MHz
*Nc*AA9C	Borisova *et al*., 2015	2.267	456
*Ta*AA9A	Quinlan *et al.*, 2011	2.278	470
*Hj*AA9A	Hansson *et al.*, 2017	2.280	474
*Mt*AA9	Span *et al.*, 2017	2.260	504
*Cv*AA9A	Simmons *et al.* 2017	2.273	476
*Ls*AA9A	Frandsen *et al.*, 2016	2.279	458
*Ba*AA10A	Hemsworth *et al.*, 2013b	2.25	404
*Cj*AA10A	Forsberg *et al.*, 2016	2.267	462
*Sm*AA10A	Forsberg *et al.*, 2014b	2.260	348
*Sc*AA10C	Forsberg *et al.*, 2014b	2.267	459
*Bl*AA10	Forsberg *et al.*, 2014b	2.262	375
*Tf*AA10B	Forsberg *et al.*, 2014b	2.262	468
*Sc*AA10B	Forsberg *et al.*, 2014a	2.270	474
*Sl*AA10E	Chaplin *et al.,* 2016	2.260 (Sp 1) 2.230 (Sp 2)	378 (Sp 1) 532 (Sp 2)
*Ao*AA11	Hemsworth *et al*., 2014	2.27	498
*Ao*AA13	Lo Leggio *et al.*, 2015	2.26	513
*Pc*AA14A	Couturier *et al*., 2018	2.273	508
*Td*AA15A	Sabbadin *et al.*, 2018	2.283 (Sp 1) 2.258 (Sp 2)	407 (Sp 1) 512 (Sp 2)
*Nc*AA9C-G6	Borisova *et al.*, 2015	2.226	525
*Ls*AA9A-G6-H_2O	Frandsen *et al.*, 2016	2.273	515
*Ls*AA9A-G6-Cl	Frandsen *et al.*, 2016	2.234	515
*Sm*AA10A-chitin	Bissaro *et al*., 2018	2.216	610

The perpendicular region of the spectrum, instead, is not as straightforward to interpret and simulate, especially from X-band data alone. Here, the partial overlap of the g_x/g_y and $|A_x|/|A_y|$ values and the usually broad appearance of the spectrum do not allow to unambiguously determine the spin Hamiltonian parameters of the copper ion. Great care should therefore be used in these circumstances. The problem can be overcome by simultaneous simulating both X- and Q-band data, when multi-frequency collection of the spectra is a possibility. Simultaneous fits of the data from two different frequencies significantly narrows the sets of possible spin Hamiltonian values, with the further advantage that artefacts arising from overshoot features, whether present, can be all but eliminated.

Analysis of the superhyperfine (SHF) coupling, when visible, can provide insights into the coordination environment of the copper active site and possible changes triggered by the binding of substrate (Frandsen *et al*., 2016). Simulation of the SHF coupling can be a very time consuming and, at times, frustrating challenge, but it can provide key information about the active site. In our experience, a reliable simulation of superhyperfine coupling patterns can take many days to perform accurately. Moreover, it is important to point out that the most direct and reliable method to obtain coupling values to coupled nuclei is the use of pulsed

EPR techniques such as ENDOR or HYSCORE (not discussed herein), in which the orientation-selective feature of the techniques allows some separation of the individual coupling components of each coordinating atom. However, as pulsed techniques are not easily accessible and can require high amount of protein, careful analysis of CW spectra can give surprisingly accurate information about the type and number of nuclei coupled to the unpaired electron and their coupling values. The risk of over-parameterising the problem should be taken in account, and we would recommend to proceed in steps and add further coupled nuclei only if the experimental data cannot be satisfactory simulated otherwise. Simultaneous simulations of both first and second derivative spectra can be of great help in establishing the coupling values, although usually only the largest component of the coupling, when anisotropic, can be determined. If possible, isotopic labelling (*e.g.* ^{15}N) should be performed in order to restrict the set of possible values which would give good fit to the experimental data.

Finally, some authors have used DFT calculations to augment their interpretations of EPR data (Bissaro *et al*., 2018). While useful in a broad context, there are well-known problems in using DFT to give accurate representations of spin Hamiltonian parameters (Neese, 2009, 2017). Indeed, these problems appear to be evident in DFT calculations performed on LPMOs. The key difficulties stem from the issues associated with the inaccurate modelling of the Fermi contact term (which, in turn, arises from configuration interaction of the SOMO with s-orbitals and polarisation of core electrons at the copper ion), the relatively poor modelling of medium-distance electron-electron interactions, the important role of dispersion forces, and the tendency of DFT calculations to overly apply covalency to metal-ligand bonding. In partial mitigation of these problems, we would recommend using a basis set in which core functions are enhanced (*e.g*. CP(PPP) (Neese, 2002)). For the latter issue, a functional should be selected in which the degree of Hartree-Fock (HF) exchange has been manually increased, *e.g*. B3LYP-D functional has 20% HF exchange set as default, but should be increased to ~38% with the final value being determined by trial and error. However, as is evident in the preceding sentence, this is necessarily an empirical procedure and our experience in this regard is that the spin Hamiltonian parameters obtained from DFT calculations for LPMOs can be significantly different from the experimental values. As such, great caution must be exercised in using DFT-determined spin Hamiltonian parameters for in-depth interpretation of the active site electronics, even when there appears to be good agreement with the experimental values. Indeed, whether the results of DFT calculations on spin Hamiltonian parameters can give useful absolute results at all in such a complicated system is debateable. Our advice in this regard therefore is that trends in calculated parameters may be meaningful (when the same basis sets, functionals and methods are used between systems), as might the absolute signs of calculated hyperfine coupling values, but that further values should *not* be used as a basis for in-depth analysis of the electronic structure of the SOMO without a critical evaluation of the underlying issues. In this context, our further advice is that a ligand field analysis in which many of the 'problematic' values are parameterised offers a more reliable and suitably-caveated analysis of LPMO EPR spectra, from which appropriately framed conclusions can be drawn.

6. Acknowledgements

GRH gratefully acknowledges support from a BBSRC (Biotechnology and Biological Sciences Research Council) David Phillips Fellowship (grant # BB/N019970/1). LC, GJD and PHW also gratefully acknowledge the support of the BBSRC under grant BB/L001926/1.

7. References

Agger, J.W., Isaksen, T., Várnai, A., Vidal-Melgosa, S., Willats, W.G.T., Ludwig, R., Horn, S.J., Eijsink, V.G.H., and Westereng, B. (2014). Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. Proc. Natl. Acad. Sci. U. S. A. *111*, 6287–6292.

Bardiya, N., and Shiu, P.K.T. (2007). Cyclosporin A-resistance based gene placement system for Neurospora crassa. Fungal Genet. Biol. *44*, 307–314.

Beeson, W.T., Vu, V. V, Span, E.A., Phillips, C.M., and Marletta, M.A. (2015). Cellulose degradation by polysaccharide monooxygenases. Annu. Rev. Biochem. *84*, 923–946.

Bennati-Granier, C., Garajova, S., Champion, C., Grisel, S., Haon, M., Zhou, S., Fanuel, M., Ropartz, D., Rogniaux, H., Gimbert, I., et al. (2015). Substrate specificity and regioselectivity of fungal AA9 lytic polysaccharide monooxygenases secreted by Podospora anserina. Biotechnol. Biofuels *8*, 90.

Bey, M., Zhou, S., Poidevin, L., Henrissat, B., Coutinho, P.M., Berrin, J.-G., and Sigoillot, J.-C. (2013). Cello-oligosaccharide oxidation reveals differences between two lytic polysaccharide monooxygenases (family GH61) from Podospora anserina. Appl. Environ. Microbiol. *79*, 488–496.

Bischof, R.H., Ramoni, J., and Seiboth, B. (2016). Cellulases and beyond: the first 70 years of the enzyme producer Trichoderma reesei. Microb. Cell Fact. *15*, 106.

Bissaro, B., Røhr, A.K., Müller, G., Chylenski, P., Skaugen, M., Forsberg, Z., Horn, S.J., Vaaje-Kolstad, G., and Eijsink, V.G.H. (2017). Oxidative cleavage of polysaccharides by monocopper enzymes depends on H2O2. Nat. Chem. Biol. *851*, 1.

Bissaro, B., Isaksen, I., Vaaje-Kolstad, G., Eijsink, V.G.H., and Røhr, Å.K. (2018). How a Lytic Polysaccharide Monooxygenase Binds Crystalline Chitin. Biochemistry *57*, 1893–1906.

Borisova, A.S., Isaksen, T., Dimarogona, M., Kognole, A.A., Mathiesen, G., Várnai, A., Røhr, Å.K., Payne, C.M., Sørlie, M., Sandgren, M., et al. (2015). Structural and Functional Characterization of a Lytic Polysaccharide Monooxygenase with Broad Substrate Specificity. J. Biol. Chem. *290*, 22955–22969.

Bruker EPR Simulation Suites. https://www.bruker.com/products/mr/epr/epr-software/simulation-suites/overview.html

Chaplin, A.K., Wilson, M.T., Hough, M.A., Svistunenko, D.A., Hemsworth, G.R., Walton, P.H., Vijgenboom, E., and Worrall, J.A.R. (2016). Heterogeneity in the Histidine-brace Copper Coordination Sphere in Auxiliary Activity Family 10 (AA10) Lytic Polysaccharide Monooxygenases. J. Biol. Chem. *291*, 12838–12850.

Chiu, E., Hijnen, M., Bunker, R.D., Boudes, M., Rajendran, C., Aizel, K., Oliéric, V., Schulze-Briese, C., Mitsuhashi, W., Young, V., et al. (2015). Structural basis for the enhancement of virulence by viral spindles and their in vivo crystallization. Proc. Natl. Acad. Sci. U. S. A. *112*,

3973–3978.

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.

Courtade, G., Wimmer, R., Røhr, A.K., Preims, M., Felice, A.K.G., Dimarogona, M., Vaaje-Kolstad, G., Sørlie, M., Sandgren, M., Ludwig, R., et al. (2016). Interactions of a fungal lytic polysaccharide monooxygenase with β-glucan substrates and cellobiose dehydrogenase. Proc. Natl. Acad. Sci. U. S. A. *113*, 5922–5927.

Courtade, G., Le, S.B., Sætrom, G.I., Brautaset, T., and Aachmann, F.L. (2017). A novel expression system for lytic polysaccharide monooxygenases. Carbohydr. Res.

Couturier, M., Ladevèze, S., Sulzenbacher, G., Ciano, L., Fanuel, M., Moreau, C., Villares, A., Cathala, B., Chaspoul, F., Frandsen, K.E., et al. (2018). Lytic xylan oxidases from wood-decay fungi unlock biomass degradation. Nat. Chem. Biol. *14*, 306–310.

Crouch, L.I., 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 Monooxygenases. J. Biol. Chem. *291*, 7439–7449.

Forsberg, Z., Vaaje-Kolstad, G., Westereng, B., Bunæs, A.C., Stenstrøm, Y., MacKenzie, A., Sørlie, M., Horn, S.J., and Eijsink, V.G.H. (2011). Cleavage of cellulose by a CBM33 protein. Protein Sci. *20*, 1479–1483.

Forsberg, Z., Mackenzie, A.K., Sørlie, M., Røhr, Å.K., Helland, R., Arvai, A.S., Vaaje-Kolstad, G., and Eijsink, V.G.H. (2014a). Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. Proc. Natl. Acad. Sci. U. S. A. *111*, 8446–8451.

Forsberg, Z., Røhr, Å.K., Mekasha, S., Andersson, K.K., Eijsink, V.G.H., Vaaje-Kolstad, G., and Sørlie, M. (2014b). Comparative study of two chitin-active and two cellulose-active AA10-type lytic polysaccharide monooxygenases. Biochemistry *53*, 1647–1656.

Forsberg, Z., Nelson, C.E., Dalhus, B., Mekasha, S., Loose, J.S.M., Crouch, L.I., Røhr, A.K., Gardner, J.G., Eijsink, V.G.H., and Vaaje-Kolstad, G. (2016). Structural and Functional Analysis of a Lytic Polysaccharide Monooxygenase Important for Efficient Utilization of Chitin in Cellvibrio japonicus. J. Biol. Chem. *291*, 7300–7312.

Frandsen, K.E.H., Simmons, T.J., Dupree, P., Poulsen, J.-C.N., Hemsworth, G.R., Ciano, L., Johnston, E.M., Tovborg, M., Johansen, K.S., von Freiesleben, P., et al. (2016). The molecular basis of polysaccharide cleavage by lytic polysaccharide monooxygenases. Nat. Chem. Biol. *12*, 298–303.

Frommhagen, M., Sforza, S., Westphal, A.H., Visser, J., Hinz, S.W.A., Koetsier, M.J., van Berkel, W.J.H., Gruppen, H., and Kabel, M.A. (2015). Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase. Biotechnol. Biofuels *8*, 101.

Gardner, J.G., Crouch, L., Labourel, A., Forsberg, Z., Bukhman, Y. V, Vaaje-Kolstad, G., Gilbert, H.J., and Keating, D.H. (2014). Systems biology defines the biological significance of redox-active proteins during cellulose degradation in an aerobic bacterium. Mol. Microbiol. *94*, 1121–1133.

Good, N.E., Winget, G.D., Winter, W., Connolly, T.N., Izawa, S., and Singh, R.M. (1966).

Hydrogen ion buffers for biological research. Biochemistry *5*, 467–477.

Gregory, R.C., Hemsworth, G.R., Turkenburg, J.P., Hart, S.J., Walton, P.H., and Davies, G.J. (2016). 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.

Guzzi, R., Stirpe, A., Verbeet, M., and Sportelli, L. (2001). Structural heterogeneity of blue copper proteins: an EPR study of amicyanin and of wild-type and Cys3Ala/Cys26Ala mutant azurin. Eur. Biophys. J. *30*, 171–178.

Hansson, H., Karkehabadi, S., Mikkelsen, N., Douglas, N.R., Kim, S., Lam, A., Kaper, T., Kelemen, B., Meier, K.K., Jones, S.M., et al. (2017). High-resolution structure of a lytic polysaccharide monooxygenase from Hypocrea jecorina reveals a predicted linker as an integral part of the catalytic domain. J. Biol. Chem.

Harris, P. V, Welner, D., McFarland, K.C., Re, E., Navarro Poulsen, J.-C., Brown, K., Salbo, R., Ding, H., Vlasenko, E., Merino, S., 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.

Harris, P. V, Xu, F., Kreel, N.E., Kang, C., and Fukuyama, S. (2014). New enzyme insights drive advances in commercial ethanol production. Curr. Opin. Chem. Biol. *19*, 162–170.

Hemsworth, G.R., Davies, G.J., and Walton, P.H. (2013a). Recent insights into copper-containing lytic polysaccharide mono-oxygenases. Curr. Opin. Struct. Biol. *23*, 660–668.

Hemsworth, G.R., Taylor, E.J., Kim, R.Q., Gregory, R.C., Lewis, S.J., Turkenburg, J.P., Parkin, A., Davies, G.J., and Walton, P.H. (2013b). The copper active site of CBM33 polysaccharide oxygenases. J. Am. Chem. Soc. *135*, 6069–6077.

Hemsworth, G.R., Henrissat, B., Davies, G.J., and Walton, P.H. (2014). Discovery and characterization of a new family of lytic polysaccharide monooxygenases. Nat. Chem. Biol. *10*, 122–126.

Hemsworth, G.R., Johnston, E.M., Davies, G.J., and Walton, P.H. (2015). Lytic Polysaccharide Monooxygenases in Biomass Conversion. Trends Biotechnol. *33*, 747–761.

Himmel, M.E., Ding, S.-Y., Johnson, D.K., Adney, W.S., Nimlos, M.R., Brady, J.W., and Foust, T.D. (2007). Biomass recalcitrance: engineering plants and enzymes for biofuels production. Sci. (New York, NY) *315*, 804–807.

Horn, S.J., Vaaje-Kolstad, G., Westereng, B., and Eijsink, V.G. (2012). Novel enzymes for the degradation of cellulose. Biotechnol. Biofuels *5*, 45.

Karkehabadi, S., Hansson, H., Kim, S., Piens, K., Mitchinson, C., and Sandgren, M. (2008). The first structure of a glycoside hydrolase family 61 member, Cel61B from Hypocrea jecorina, at 1.6 A resolution. *383*, 144–154.

Karlsson, J., Saloheimo, M., Siika-Aho, M., Tenkanen, M., Penttilä, M., and Tjerneld, F. (2001). Homologous expression and characterization of Cel61A (EG IV) of Trichoderma reesei. Eur. J. Biochem. *268*, 6498–6507.

Kittl, R., Kracher, D., Burgstaller, D., Haltrich, D., and Ludwig, R. (2012). Production of four Neurospora crassa lytic polysaccharide monooxygenases in Pichia pastoris monitored by a

fluorimetric assay. Biotechnol. Biofuels *5*, 79.

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.

Kuusk, S., Bissaro, B., Kuusk, P., Forsberg, Z., Eijsink, V.G.H., Sørlie, M., and Väljamäe, P. (2018). Kinetics of H2O2-driven degradation of chitin by a bacterial lytic polysaccharide monooxygenase. J. Biol. Chem. *293*, 523–531.

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.

Lo Leggio, L., Simmons, T.J., Poulsen, J.-C.N., Frandsen, K.E.H., Hemsworth, G.R., Stringer, M.A., von Freiesleben, P., Tovborg, M., Johansen, K.S., De Maria, L., et al. (2015). Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase. Nat. Commun. *6*, 5961.

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.

Li, X., Beeson, W.T., Phillips, C.M., Marletta, M.A., and Cate, J.H.D. (2012). Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. Structure *20*, 1051–1061.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. *42*, D490-5.

Mandels, M., and Reese, E.T. (1957). Induction of cellulase in Trichoderma viride as influenced by carbon sources and metals. J. Bacteriol. *73*, 269–278.

Mandels, M., and Reese, E.T. (1960). Induction of cellulase in fungi by cellobiose. J. Bacteriol. *79*, 816–826.

Mandels, M., Parrish, F.W., and Reese, E.T. (1962). Sophorose as an inducer of cellulase in Trichoderma viride. J. Bacteriol. *83*, 400–408.

Mandels, M., Weber, J., and Parizek, R. (1971). Enhanced cellulase production by a mutant of Trichoderma viride. Appl. Microbiol. *21*, 152–154.

Mandels, M., Hontz, L., Nystrom, J., and Lee R Lynd, I.B. (1974). Enzymatic hydrolysis of waste cellulose. Biotechnol. Bioeng. *16*, 1471–1493.

Merino, S.T., and Cherry, J. (2007). Progress and challenges in enzyme development for biomass utilization. Adv. Biochem. Eng. Biotechnol. *108*, 95–120.

Morgenstern, I., Powlowski, J., and Tsang, A. (2014). Fungal cellulose degradation by oxidative enzymes: from dysfunctional GH61 family to powerful lytic polysaccharide monooxygenase family. Brief. Funct. Genomics *13*, 471–481.

Naik, S.N., Goud, V. V, Rout, P.K., and Dalai, A.K. (2010). Production of first and second generation biofuels: a comprehensive review. Renew. Sustain. Energy Rev. *14*, 578–597.

Nakagawa, Y.S., Kudo, M., Loose, J.S.M., Ishikawa, T., Totani, K., Eijsink, V.G.H., and Vaaje-Kolstad, G. (2015). A small lytic polysaccharide monooxygenase from Streptomyces griseus

786 targeting α- and β-chitin. FEBS J. *282*, 1065–1079.

787 Neese, F. (2002). Prediction and interpretation of the57Fe isomer shift in Mössbauer spectra
788 by density functional theory. Inorganica Chim. Acta *337*, 181–192.

789 Neese, F. (2009). Prediction of molecular properties and molecular spectroscopy with
790 density functional theory: From fundamental theory to exchange-coupling. Coord. Chem.
791 Rev. *253*, 526–563.

792 Neese, F. (2017). Quantum Chemistry and EPR Parameters. In EMagRes, (Chichester, UK:
793 John Wiley & Sons, Ltd), pp. 1–22.

794 Neu, H.C., and Heppel, L.A. (1965). The release of enzymes from Escherichia coli by osmotic
795 shock and during the formation of spheroplasts. J. Biol. Chem. *240*, 3685–3692.

796 Payne, C.M., Knott, B.C., Mayes, H.B., Hansson, H., Himmel, M.E., Sandgren, M., Ståhlberg,
797 J., and Beckham, G.T. (2015). Fungal cellulases. Chem. Rev. *115*, 1308–1448.

798 Peisach, J., and Blumberg, W.E. (1974). Structural implications derived from the analysis of
799 electron paramagnetic resonance spectra of natural and artificial copper proteins. Arch.
800 Biochem. Biophys. *165*, 691–708.

801 Phillips, C.M., Beeson, W.T., Cate, J.H., and Marletta, M.A. (2011). Cellobiose dehydrogenase
802 and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation
803 by Neurospora crassa. ACS Chem. Biol. *6*, 1399–1406.

804 Quinlan, R.J., Sweeney, M.D., Lo Leggio, L., Otten, H., Poulsen, J.-C.N., Johansen, K.S., Krogh,
805 K.B.R.M., Jørgensen, C.I., Tovborg, M., Anthonsen, A., et al. (2011). Insights into the
806 oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass
807 components. Proc. Natl. Acad. Sci. U. S. A. *108*, 15079–15084.

808 Reese, E.T., Siu, R.G.H., and Levinson, H.S. (1950). The biological degradation of soluble
809 cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis. J.
810 Bacteriol. *59*, 485–497.

811 Sabbadin, F., Hemsworth, G.R., Ciano, L., Henrissat, B., Dupree, P., Tryfona, T., Marques,
812 R.D.S., Sweeney, S.T., Besser, K., Elias, L., et al. (2018). An ancient family of lytic
813 polysaccharide monooxygenases with roles in arthropod development and biomass
814 digestion. Nat. Commun. *9*, 756.

815 Saini, J.K., Saini, R., and Tewari, L. (2015). Lignocellulosic agriculture wastes as biomass
816 feedstocks for second-generation bioethanol production: concepts and recent
817 developments. 3 Biotech *5*, 337–353.

818 Saloheimo, M., Nakari-Setälä, T., Tenkanen, M., and Penttilä, M. (1997). cDNA cloning of a
819 Trichoderma reesei cellulase and demonstration of endoglucanase activity by expression in
820 yeast. Eur. J. Biochem. *249*, 584–591.

821 Simmons, T.J., Frandsen, K.E.H., Ciano, L., Tryfona, T., Lenfant, N., Poulsen, J.C., Wilson,
822 L.F.L., Tandrup, T., Tovborg, M., Schnorr, K., et al. (2017). Structural and electronic
823 determinants of lytic polysaccharide monooxygenase reactivity on polysaccharide
824 substrates. Nat. Commun. *8*, 1064.

825 Solomon, E.I., Heppner, D.E., Johnston, E.M., Ginsbach, J.W., Cirera, J., Qayyum, M., Kieber-
826 Emmons, M.T., Kjaergaard, C.H., Hadt, R.G., and Tian, L. (2014). Copper active sites in

827 biology. Chem. Rev. *114*, 3659–3853.

828 Span, E.A., Suess, D.L.M., Deller, M.C., Britt, R.D., and Marletta, M.A. (2017). The Role of the
829 Secondary Coordination Sphere in a Fungal Polysaccharide Monooxygenase. ACS Chem. Biol.
830 *12*, 1095–1103.

831 Stoll, S., and Schweiger, A. (2006). EasySpin, a comprehensive software package for spectral
832 simulation and analysis in EPR. J. Magn. Reson. *178*, 42–55.

833 Tanghe, M., Danneels, B., Camattari, A., Glieder, A., Vandenberghe, I., Devreese, B., Stals, I.,
834 and Desmet, T. (2015). Recombinant Expression of Trichoderma reesei Cel61A in Pichia
835 pastoris: Optimizing Yield and N-terminal Processing. Mol. Biotechnol.

836 The Hendrich Metalloprotein Group Spincount.
837 http://www.chem.cmu.edu/groups/hendrich/facilities/index.html

838 Vaaje-Kolstad, G., Horn, S.J., van Aalten, D.M.F., Synstad, B., and Eijsink, V.G.H. (2005a). The
839 non-catalytic chitin-binding protein CBP21 from Serratia marcescens is essential for chitin
840 degradation. J. Biol. Chem. *280*, 28492–28497.

841 Vaaje-Kolstad, G., Houston, D.R., Riemen, A.H.K., Eijsink, V.G.H., and van Aalten, D.M.F.
842 (2005b). Crystal structure and binding properties of the Serratia marcescens chitin-binding
843 protein CBP21. J. Biol. Chem. *280*, 11313–11319.

844 Vaaje-Kolstad, G., Westereng, B., Horn, S.J., Liu, Z., Zhai, H., Sørlie, M., and Eijsink, V.G.H.
845 (2010). An oxidative enzyme boosting the enzymatic conversion of recalcitrant
846 polysaccharides. Sci. (New York, NY) *330*, 219–222.

847 Vaaje-Kolstad, G., Forsberg, Z., Loose, J.S., Bissaro, B., and Eijsink, V.G. (2017). Structural
848 diversity of lytic polysaccharide monooxygenases. Curr. Opin. Struct. Biol. *44*, 67–76.

849 Vu, V. V., and Ngo, S.T. (2018). Copper active site in polysaccharide monooxygenases. Coord.
850 Chem. Rev. *368*, 134–157.

851 Vu, V. V, Beeson, W.T., Span, E.A., Farquhar, E.R., and Marletta, M.A. (2014a). A family of
852 starch-active polysaccharide monooxygenases. Proc. Natl. Acad. Sci. U. S. A. *111*, 13822–
853 13827.

854 Vu, V. V, Beeson, W.T., Phillips, C.M., Cate, J.H.D., and Marletta, M.A. (2014b). Determinants
855 of regioselective hydroxylation in the fungal polysaccharide monooxygenases. J. Am. Chem.
856 Soc. *136*, 562–565.

857 Wu, M., Beckham, G.T., Larsson, A.M., Ishida, T., Kim, S., Payne, C.M., Himmel, M.E.,
858 Crowley, M.F., Horn, S.J., Westereng, B., et al. (2013). Crystal structure and computational
859 characterization of the lytic polysaccharide monooxygenase GH61D from the Basidiomycota
860 fungus Phanerochaete chrysosporium. J. Biol. Chem. *288*, 12828–12839.
861