

ABSTRACT

SCHRÖDER, GABRIELA CHARLOTTE INGEBORG. Dissecting Metalloprotein Chemistry: Characterization of a Copper Containing Lytic Polysaccharide Monooxygenase and a Heme Cytochrome P450 Using X-ray and Neutron Scattering. (Under the direction of Dr. Flora Meilleur).

Metalloproteins display a broad of chemical functionality facilitated by their metal cofactors. In particular the copper and iron cofactors of lytic polysaccharide monooxygenases (LPMOs) and cytochrome P450s (CYP450s), respectively, are central to their oxidoreductase function. Presented here are the structural and mechanistic insights on LPMOs and CYP450s gained from X-ray and neutron scattering studies.

LPMOs are copper-containing enzymes that perform oxidative cleavage of the glycosidic bond in crystalline cellulose. During LPMO catalysis, the reaction is initiated by reduction of the copper active site and addition of molecular oxygen to ultimately form an unknown activated copper–oxygen species responsible for polysaccharide substrate hydrogen atom abstraction. Neutron protein crystallography, which allows direct visualization of hydrogen atoms, was used to investigate the active site and intermediate protonation states in the initial stages of oxygen activation directly following active site copper reduction by ascorbate in *Neurospora crassa* LPMO9D. A stabilized activated dioxygen intermediate in a mixture of superoxo and hydroperoxo states was cryo-trapped and the conserved second coordination shell residue His157 was found to be the proton donor. Mining minima free energy calculations indicated that this second shell His157 displays conformational flexibility in solution which may play a role in early *Nc*LPMO9D activation of molecular oxygen. Furthermore, binding-face to binding-face crystal packing present in the *Nc*LPMO9D crystal is analogous to substrate binding and ensures active site bound intermediate stabilization.

CYP450s are a class of heme center proteins ubiquitous in most life forms that catalyze the insertion of oxygen into a non-activated C-H bond with the addition of molecular oxygen, two protons and two electrons obtained from reduced cofactors. A high-valent iron-oxo species termed Compound I is responsible for oxygen insertion during catalysis. Formation of Compound I requires cleavage of the O-O bond in molecular oxygen by protonation of the leaving oxygen group. A proton delivery pathway involving

active site residues and water molecules is likely to be involved, but the exact nature of the pathway is still debated. In CYP450cam from *Pseudomonas putida*, which catalyzes the hydroxylation of D-camphor, the acid-alcohol pair composed of Asp251 and Thr252 are thought to be involved in proton delivery to active site reaction intermediates. Crystallization trials were performed on CYP450cam to obtain crystals sufficiently large for analysis by neutron protein crystallography to visualize critical protonation states in the active site and gain insight into the proton transfer pathway.

CYP450cam undergoes a number of conformational changes that are dependent on substrate camphor binding and interactions with the redox partner protein Putidaredoxin (Pdx). The substrate-dependent conformational changes of CYP450cam, a Leu358Pro mutant and a covalent CYP450cam-Pdx complex were investigated using solution small angle X-ray scattering (SAXS). CYP450cam was found to undergo camphor-dependent conformational changes supported by a change in the radius of gyration (R_g) and the maximum particle dimension (D_{max}). The presence of a second camphor binding allosteric site was supported by changes in conformation indicated by an increase in R_g and D_{max} at 43 μ M camphor. The conformational flexibility of CYP450cam was modelled using normal mode analysis (NMA) which indicated that the β 2-sheet as well as the H-helix, G-helix, D-helix and C-helix, which are associated with the allosteric site, show increased flexibility when compared to the crystal structure.

© Copyright 2021 by Gabriela Charlotte Ingeborg Schröder

All Rights Reserved

Dissecting Metalloprotein Chemistry: Characterization of a Copper Containing Lytic Polysaccharide
Monooxygenase and a Heme Cytochrome P450 Using X-ray and Neutron Scattering

by
Gabriela Charlotte Ingeborg Schröder

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Biochemistry

Raleigh, North Carolina

2021

APPROVED BY:

Dr. Flora Meilleur
Committee Chair

Dr. Reza A. Ghiladi

Dr. Michael B. Goshe

Dr. Guozhou Xu

DEDICATION

To Mum and Dad for your endless love and support every step of the way.

To Carlyle Deane, the best Cuzzy in the world.

Thank you for always believing in me.

I miss you more than words can say.

BIOGRAPHY

Gabriela Charlotte Ingeborg Schröder was born in Cape Town in the Western Cape, South Africa, to Michelle and Joachim Schröder and has a younger sister, Christianna (Christie). She spent her early formative years in the Mother City at the foot of the iconic Table Mountain before her family settled in Harrismith in the Free State at the foot of the arguably equally iconic albeit much smaller Platberg. Gabriela attended Harriston Primary and High School. Her father is a professor of mathematics and her mother is a language teacher and their combined tutelage helped her excel in both the Sciences and Arts, being awarded the academic Dux student award in her final year. During her time at school, Gabriela participated in numerous Science Expo competitions which promoted her interest in science.

At the beginning of 2011, Gabriela commenced a Bachelor of Science (BSc) degree at the University of the Free State in Bloemfontein, South Africa with an emphasis in Chemistry and Biochemistry. She graduated with distinction in 2014 and received the several honors including the second and third year Chemistry and Biochemistry Best Student Award as well as the Dean's Medal and the Senate Medal.

Following her undergraduate studies, Gabriela joined the Biocatalysis Group at the Department of Microbial, Biochemical and Food Biotechnology at the University of the Free State to pursue an Honours degree in Biochemistry. Gabriela's research project encompassed work on cytochrome P450s (CYP450s) for biocatalytic applications and it is during this time that her passion for metalloproteins and elucidating enzyme structure and chemistry was born. She graduated with a BScHons degree with distinction in 2015.

During her undergraduate studies, Gabriela had the opportunity to participate in exchange programs to the University of Vermont and Stanford University. The outstanding research facilities and academic environment in the United States made an impression on her, and she applied for a Fulbright Scholarship. Gabriela was awarded a Fulbright to pursue a PhD in Biochemistry at Department of Molecular and Structural Biochemistry at North Carolina State University in Raleigh in the Fall of 2015. She joined the research group of Dr. Flora Meilleur, inspired by the research conducted by Dr. Meilleur on CYP450s and

lytic polysaccharide monoxygenases (LPMOs) using neutron protein crystallography. Gabriela has conducted her thesis research at Oak Ridge National Laboratory in Oak Ridge, Tennessee where she had the opportunity to perform neutron protein crystallography diffraction experiments at the neutron scattering beamlines IMAGINE and MaNDi at the High Flux Isotope Reactor (HFIR) and Spallation Neutron Source (SNS), respectively.

Gabriela is expected to graduate with her PhD in Biochemistry in December 2021 once she meets the degree requirements and hopes to continue her research on metalloprotein structure, chemistry and their biocatalytic applications.

ACKNOWLEDGMENTS

I would like to begin by thanking Dr. Flora Meilleur for her expert guidance and opening the door to the spectacular (and intricate) world of structural biology to a young scientist. I am incredibly grateful that you imparted me with your vast knowledge of protein crystallography, enzyme chemistry and neutron scattering. Through your mentorship I learnt the fundamentals of scientific inquiry and the freedom you gave me to pursue my own scientific questions have helped me to become a more independent researcher. In a field that centers on locating the smallest of the elements, you taught me to step back to see the forest for the trees, or indeed the protein for the hydrogens, and I will carry forward those valuable lessons on perspective. Your enthusiasm and inquisitive nature have been an inspiration. Thank you for this incredible opportunity.

Thank you to my committee members Dr. Mike Goshe, Dr. Reza Ghiladi and Dr. Guozhou Xu for your advice and valuable input throughout my graduate career. I would also like to thank Dr. Paul Swartz and Dr. Robert Rose for teaching me the fundamentals of protein crystallography – your mentorship helped to lay a strong foundation for my subsequent neutron structural studies. I would also like to acknowledge all the wonderful members of the Department of Molecular and Structural Biochemistry at NCSU.

To the previous members of the Meilleur Group Dr. Annette Bodenheimer and Dr. Brad O'Dell: thank you for being wonderful friends and great hands-on mentors. Moving to a new country and then relocating again a short time later can be intimidating but your support both in the lab and in settling in to a new home were indispensable. Thank you also to Brad for the groundwork you laid for the LPMO project. Your diligence and scientific insight ensured the success of the project when I later inherited it. I would like to acknowledge Dr. Pratul Aggarwal (Oklahoma State University) and Dr. Simon Webb (VeraChem) for their computational contributions to the LPMO project which have been invaluable.

At NCSU I would like to thank Dr. Sayan Chakraborty for taking me under his wing as a new graduate student. Sayan, you are a born teacher, and I learnt many valuable practical skills from you (as well as the best places to eat in Raleigh). At ORNL, I would especially like to thank Dr. Kevin Weiss for taking technical support to the next level. Your problem solving skills, guidance and ideas were fundamental to my research. I would like to thank Dr. Wellington Leite for instructing me on performing SAXS and analyzing data when I was entirely new to the technique. To all the friends and colleagues I made along the way at ORNL: I am fortunate to have met you all. Thank you for enriching my life both personally and academically. I have learnt valuable lessons about being a good scientist and human being from all of you.

To my family, friends and loved ones: my gratitude truly cannot be expressed in words. I thank my parents Michelle and Joachim Schröder for their love and support and for providing me with so many incredible opportunities. Mum, you taught me an excellent work ethic and to dream big! I will always remember our long conversations and plans for the future on picnic blankets dappled with sunlight. Those moments opened new horizons I could never have imagined! Dad, your analytical mind and scientific ethos molded me into the scientist I am today, and I thank you for that. To my precious little sister Christie, thank you for understanding that I had to travel far away to pursue my dreams. You bring me so much joy my Noodle! To Gabriele, my Snoopy, thank you for your infinite love and support. You encouraged me when I felt down, were always ready to lend a helping hand and most importantly always make me smile. I am so lucky to have someone as radiant as you in my life! Thank you also to Faith and Shane for all the encouragement and care packages. And to Lorena and Walter: *Grazie mille per tutto. Sono fortunata a conoscermi!*

I would finally like to thank the National Research Foundation (NRF) in South Africa and the Graduate Opportunities (GO!) Program at ORNL for funding. I would also like to thank the Fulbright Scholarship and their mission to promote global connections for this incredible opportunity.

TABLE OF CONTENTS

LIST OF TABLES	xii
LIST OF FIGURES.....	xiii
Chapter 1 : Characterization of metalloproteins using neutron protein crystallography	1
1.1 Metalloprotein catalysis: structural and mechanistic insights into oxidoreductases from neutron protein crystallography	1
1.1.1 Synopsis.....	1
1.1.2 Abstract.....	1
1.1.3 Keywords.....	2
1.1.4 Introduction.....	2
1.1.4.1 Metals in biology.....	2
1.1.4.2 Structural studies and radiation damage.....	3
1.1.4.3 Neutron protein crystallography.....	4
1.1.5 Neutron protein crystallography of transition state metal oxidoreductases.....	6
1.1.5.1 Copper nitrite reductase.....	8
1.1.5.2 Copper amine oxidase.....	12
1.1.5.3 Heme peroxidases	15
1.1.5.4 Lytic polysaccharide monooxygenase.....	18
1.1.5.5 Manganese superoxide dismutase.....	22
1.1.5.6 Chlorite dismutase.....	25
1.1.5.7 Amicyanin.....	28
1.1.5.8 Manganese catalase	31
1.1.6 Perspective	34
1.1.7 Conflict of interest	34
1.1.8 Funding Information.....	34
1.1.9 References.....	35
1.2 IMAGINE: neutrons reveal enzyme chemistry	47
1.2.1 Synopsis.....	47
1.2.2 Abstract.....	47
1.2.3 Keywords.....	47
1.2.4 Neutron protein crystallography	47
1.2.5 Overview of the IMAGINE beamline	49
1.2.6 Scientific highlights	52
1.2.6.1 HIV-1 protease	55
1.2.6.2 <i>H. pylori</i> 5-methylthioadenosine nucleosidase (HpMTAN).....	56
1.2.6.3 <i>Neurospora crassa</i> lytic polysaccharide monooxygenase 9D (NcLPMO9D).....	57
1.2.6.4 Cholesterol oxidase.....	59
1.2.6.5 Dihydrofolate reductase.....	60
1.2.6.6 H-RAS GTPase.....	61
1.2.7 Neutron protein crystallography at cryogenic temperature	62
1.2.8 Instrument development	64
1.2.9 Conclusions.....	65
1.2.10 Acknowledgements.....	66
1.2.11 References.....	67
1.3 A Guide to neutron crystallography data collection and processing for modelling hydrogen atoms in protein structures.....	69
1.3.1 Keywords.....	69
1.3.2 Summary.....	69
1.3.3 Abstract.....	69
1.3.4 Introduction.....	70

1.3.5 Representative results	78
1.3.6 Discussion	89
1.3.7 Acknowledgements.....	94
1.3.8 Disclosures.....	94
1.3.9 References.....	95
1.4 Overview of research.....	99
Chapter 2 : Elucidating the reaction mechanism of lytic polysaccharide monooxygenases with neutron protein crystallography	101
2.1 Introduction to catalysis by lytic polysaccharide monooxygenases	101
2.1.1 Lignocellulosic biomass	101
2.1.2 Fungal enzymatic cellulose breakdown.....	104
2.1.3 Lytic polysaccharide monooxygenases	108
2.1.3.1 Classification of LPMOs	108
2.1.3.2 Structure of LPMOs	109
2.1.3.3 Active site	112
2.1.3.4 Mechanism of LPMOs.....	113
2.1.3.5 H ₂ O ₂ as the LPMO co-substrate	116
2.1.4 Proposed hydrogen atom abstraction (HAA) species.....	117
2.1.4.1 Superoxo Cu ²⁺ -(O ₂ ^{•-}).....	117
2.1.4.3 Hydroperoxo Cu ²⁺ -(OOH) or Cu ³⁺ -(OOH)	118
2.1.4.4 Hydroxyl Cu ³⁺ -(OH).....	118
2.1.4.5 Oxyl Cu ²⁺ -(O [•]) and oxo Cu ³⁺ -(O)	119
2.1.4.6 Oxyl Cu ²⁺ -(O [•]) derived from H ₂ O ₂	120
2.1.5 Electron donors.....	121
2.1.5.1 Small molecule electron donors	121
2.1.5.2 Cellobiose dehydrogenase (CDH)	121
2.1.5.3 Pyrroloquinoline quinone (PQQ)-dependent pyranose dehydrogenase (PDH)	122
2.1.5.4 Glucose-methanol choline (GMC) Oxidases and Polyphenol Oxidases.....	123
2.1.5.5 Light-driven reduction.....	123
2.1.6 Structural studies on oxygen activation in <i>Neurospora crassa</i> LPMO9D	123
2.1.7 Aim of research.....	127
2.1.8 References.....	128
2.2 Preliminary results on neutron and X-ray diffraction data collection of a lytic polysaccharide monooxygenase under reduced and acidic conditions.....	137
2.2.1 Synopsis.....	137
2.2.2 Abstract.....	137
2.2.3 Keywords.....	138
2.2.4 Introduction.....	138
2.2.5 Methods and Materials	141
2.2.5.1 Protein production and crystallization	141
2.2.5.2 Data Collection	141
2.2.6 Results and Discussion	144
2.2.7 Acknowledgements.....	147
2.2.8 Funding Information.....	147
2.2.9 References.....	148
2.3 Neutron cryo-crystallography captures activated dioxygen intermediates at the copper-active site of a lytic polysaccharide monooxygenase	151
2.3.1 Abstract.....	151
2.3.2 Introduction.....	152
2.3.3 Results and discussion	154
2.3.3.1 Crystallographic features of the NcLPMO9D neutron structures.....	154
2.3.3.2 Ascorbate-reduced copper active site	156

2.3.3.3 Protonation state of His157	159
2.3.3.4 N-terminal protonation following initial dioxygen activation	163
2.3.3.5 Histidine protonation at acidic conditions.....	164
2.3.3.6 Conformational flexibility of His157.....	166
2.3.4 Conclusion	169
2.3.5 Experimental section	170
2.3.5.1 Protein expression and crystallization.....	170
2.3.5.2 Crystal ascorbate soak and data collection	170
2.3.5.3 Crystal low pH vapor exchange data collection	170
2.3.5.4 Crystal low pH direct soak and data collection.....	171
2.3.5.5 Model refinement.....	171
2.3.5.6 Electronic structure (density functional theory) calculations.....	171
2.3.5.7 VM2 calculations.....	172
2.3.6 Conflicts of interest.....	174
2.3.7 Acknowledgements.....	174
2.3.8 References.....	175

Chapter 3 : Optimization of expression and crystallization of CYP450cam..... 179

3.1 Introduction to the structure and mechanism of cytochrome P450s..... 179

3.1.1 Cytochrome P450s in the different kingdoms of life.....	179
3.1.1.1 Prokaryotic CYP450s.....	179
3.1.1.2 Viral CYP450s.....	180
3.1.1.3 Fungal CYP450s	180
3.1.1.4 Plant CYP450s.....	180
3.1.1.5 Mammalian CYP450s.....	181
3.1.2 Nomenclature of CYP450s.....	181
3.1.3 Classes of CYP450s.....	182
3.1.4 Structure of CYP450s.....	184
3.1.5 The Cytochrome P450 Catalytic cycle	186
3.1.6 Protonation pathway	188
3.1.7 CYP101A1 (CYP450cam).....	191
3.1.8 Neutron Studies of CYP450cam.....	193
3.1.9 Large crystal growth for neutron protein crystallography	194
3.1.10 General principles of protein crystallization.....	195
3.1.10.1 Vapor diffusion.....	198
3.1.10.2 Dialysis.....	198
3.1.10.3 Batch.....	199
3.1.10.4 Capillary counter-diffusion.....	199
3.1.10.5 Strategies for neutron protein crystallography.....	199

3.2 Aim of research..... 201

3.3 Materials and methods..... 202

3.3.1 Reagents.....	202
3.3.2 Protein Expression	202
3.3.3 Protein purification	202
3.3.3.1 Ni-affinity purification.....	202
3.3.3.2 Anion exchange purification	203
3.3.3.3 Size exclusion chromatography.....	203
3.3.4 Hemin reconstitution	203
3.3.5 Characterization.....	204
3.3.5.1 SDS-PAGE	204
3.3.5.2 Native SDS-PAGE	204
3.3.5.3 Absorption spectrum	204
3.3.5.4 CO-difference spectrum	204
3.3.6 Crystallization.....	205

3.3.6.1 Vapor diffusion.....	205
3.3.6.2 Microbatch.....	205
3.3.6.3 Microdialysis.....	206
3.3.6.4 Screens.....	206
3.3.6.5 Microseeding.....	206
3.3.7 X-ray diffraction.....	207
3.4 Results and discussion.....	208
3.4.1 Protein expression and purification.....	208
3.4.1.1 His-tag dependent purification.....	209
3.4.1.2 Anion exchange chromatography.....	210
3.4.1.3 DEAE column.....	211
3.4.1.4 Q-column purification.....	211
3.4.2 Heme reconstitution.....	212
3.4.3 Native SDS-PAGE.....	215
3.4.4 Crystallization.....	216
3.4.4.1 Vapor diffusion.....	217
3.4.4.2 Microbatch.....	218
3.4.4.3 Microdialysis.....	219
3.4.4.4 Screens.....	220
3.4.4.5 Microseeding.....	221
3.4.5 X-ray diffraction.....	221
3.5 Conclusions.....	226
3.6 References.....	227
Chapter 4 : Conformational dynamics of CYP450cam investigated by small angle X-ray scattering.....	235
4.1 Introduction to cytochrome P450cam conformational dynamics.....	235
4.1.1 Catalysis by the CYP450cam-Pdx Couple.....	235
4.1.2 Substrate-dependent CYP450cam conformational changes.....	236
4.1.3 Structural characteristics of the CYP450cam-Pdx complex.....	238
4.1.4 The L358P mutation.....	241
4.1.5 An allosteric site for CYP450cam.....	242
4.1.6 Small angle X-ray scattering studies of CYP450cam.....	245
4.2. Aim of research.....	246
4.3 Materials and methods.....	247
4.3.1 Reagents.....	247
4.3.2 Mutant constructs.....	247
4.3.2.1 CYP450cam C334A.....	247
4.3.2.2 CYP450cam L358P.....	247
4.3.2.3 CYP450cam crosslink mutant.....	247
4.3.2.4 Pdx crosslink mutant.....	247
4.3.3 Protein Expression.....	248
4.3.4 Purification.....	248
4.3.4.1 CYP450cam.....	248
4.3.4.2 Putidaredoxin.....	249
4.3.5 Complex formation.....	249
4.3.6 Characterization.....	250
4.3.6.1 SDS-PAGE.....	250
4.3.6.2 CO-difference spectrum.....	250
4.3.7 Small angle X-ray scattering sample preparation.....	251
4.3.7.1 Concentration series.....	251
4.3.7.2 Camphor binding.....	251

4.3.8 Small angle X-ray scattering data collection	251
4.4 Results and discussion	253
4.4.1 Protein expression and purification	253
4.4.2 Complex formation.....	255
4.4.3 SAXS data evaluation.....	256
4.4.4 Effect of camphor concentration on CYP450cam	259
4.4.5 FOXS crystal structure and data comparison	267
4.4.6 SREFLEX conformational distribution calculations	270
4.5 Conclusions	277
4.6 References.....	279
Chapter 5 : Conclusions and perspectives	283
5.1 The role of active site and second shell residues in oxygen activation by <i>Nc</i>LPMO9D	283
5.2 Large crystal growth of CYP450cam.....	285
5.3 Conformational dynamics of CYP450cam	286
5.4 References.....	289

LIST OF TABLES

Table 1.1	Neutron protein crystallography structures of transition state metal oxidoreductases	7
Table 1.2	Neutron scattering lengths and X-ray scattering factors for atoms in biological macromolecules	48
Table 1.3	A comprehensive list of neutron and joint X-ray/neutron structures of proteins deposited in the Protein Database from the IMAGINE beamline	53
Table 1.4	Mini kappa specifications	65
Table 1.5	Neutron scattering lengths and incoherent scattering values.....	74
Table 2.1	Classification of LPMO families	109
Table 2.2	Data Collection and Processing	145
Table 2.3	Structurally characterized dioxygen species in LPMOs	159
Table 2.4	Stability and electronic state of DFT models.....	161
Table 3.1	Data collection and refinement statistics	223
Table 4.1	Size analysis of CYP450cam(Crosslink) as a function of camphor concentration	261
Table 4.2	Size analysis of CYP450cam(Crosslink) as a function of camphor concentration	264
Table 4.3	Comparison of the experimental and calculated dimensions using the FOXS server	268
Table 4.4	Comparison of the experimental and calculated dimensions using SREFLEX normal mode analysis	271
Table 4.5	Structural elements with increased mobility after introduction of structural flexibility using SREFLEX normal mode analysis	273

LIST OF FIGURES

Figure 1.1	Incoherent neutron scattering cross sections and coherent neutron scattering lengths for selected elements	4
Figure 1.2	The active site of copper nitrite reductase	11
Figure 1.3	The active site of copper amine oxidase	14
Figure 1.4	The active site of the heme-center ascorbate peroxidase (APX)	17
Figure 1.5	The active site of LPMO	21
Figure 1.6	The active site of manganese superoxide dismutase	24
Figure 1.7	The active site of chlorite dismutase	27
Figure 1.8	The active site of amicyanin and strong hydrogen bond with an unexchanged tryptophan.....	29
Figure 1.9	The active site and protonation states of manganese catalase.....	33
Figure 1.10	IMAGINE beamline layout.....	50
Figure 1.11	IMAGINE neutron spectra at the sample position.....	51
Figure 1.12	IMAGINE installed on CG-4D	52
Figure 1.13	HIV-1 protease in complex with clinical drug darunavir.....	55
Figure 1.14	<i>Hp</i> MTAN in complex with the late transition state mimic <i>p</i> -CIPh-Thio-DADMe-ImmA	57
Figure 1.15	Active site of <i>Nc</i> LPMO9D in the enzymatic resting state	58
Figure 1.16	The cholesterol oxidase active site and its vicinity	60
Figure 1.17	DHFR D27–folate hydrogen bonding.....	61
Figure 1.18	H-RAS in complex with GTP analogue GppNHp	62
Figure 1.19	An automated lift moves the CCR up and down.....	63
Figure 1.20	T4 Lysozyme “cavity 1” as observed at 80 K.....	64
Figure 1.21	The IMAGINE Instrument at the High Flux Isotope Reactor.....	72
Figure 1.22	The MaNDi Instrument at the Spallation Neutron Source.....	72
Figure 1.23	Flow chart of neutron protein crystallography workflow	77
Figure 1.24	Structure of the lytic polysaccharide monooxygenase <i>Nc</i> LPMO9D	78
Figure 1.25	Crystal with sufficient volume in sitting drop crystallization tray.....	79
Figure 1.26	Additional information from neutron SLD maps.....	81
Figure 1.27	Water molecule positioning	81
Figure 1.28	Neutron SLD F_o-F_c omit maps.....	82
Figure 1.29	Multiple protonation states displayed with neutron SLD maps	83
Figure 1.30	Appearance of water molecules in neutron SLD maps	84
Figure 1.31	Information about amino acid orientation and protonation provided by neutron SLD maps.....	86
Figure 1.32	Discontinuous neutron SLD maps.....	88
Figure 2.1	The hierarchical structure of lignocellulose.....	102
Figure 2.2	Enzymes involved in the degradation of cellulose.....	105
Figure 2.3	Structure of <i>Tt</i> GH61E with Zn^{2+} modelled in the active site	106
Figure 2.4	Common architecture of the LPMO conserved copper-coordinating histidine brace across families	111
Figure 2.5	Tetragonal and trigonal geometry of AA9 and AA10.....	113
Figure 2.6	Overview of the oxidative reaction catalyzed by LPMOs	114
Figure 2.7	The active site of the ascorbate-reduced <i>Nc</i> LPMO9D.....	125
Figure 2.8	Neutron structure of the resting state active site of <i>Nc</i> LPMO9D	126
Figure 2.9	The structure of the LPMO active site	140
Figure 2.10	Neutron diffraction images of LPMO9D	142
Figure 2.11	LPMO9D crystal for data collection	143
Figure 2.12	Crystal packing of the Molecule A and B in the LPMO9D crystals.....	155

Figure 2.13	Solvent accessible surface in the crystal packing of Molecule A and Molecule B of the asymmetric unit of LPMO9D	156
Figure 2.14	Neutron diffraction structure of the ascorbate reduced LPMO9D crystal	158
Figure 2.15	Mixed protonation state His157 in ascorbate reduced LPMO9D	162
Figure 2.16	Protonation state of the N-terminal His157 amino group in the ascorbate reduced LPMO9D crystal	164
Figure 2.17	Protonation state of second-shell His157 at acidic conditions	166
Figure 2.18	Alignment of a second shell residue His157 in the “inward” and “outward” conformation	168
Figure 2.19	Second shell His157 conformations from VM2 free energy minima calculations	168
Figure 3.1	Nomenclature of cytochrome P450s (CYP450s)	182
Figure 3.2	CYP450 Class I electron transport system	183
Figure 3.3	CYP450 Class II electron transport system	183
Figure 3.4	Representation of the structural elements of CYP450s	185
Figure 3.5	The overall CYP450 catalytic cycle	187
Figure 3.6	Residues involved in the CYP450 protonation pathway	189
Figure 3.7	Overview of the CYP450cam structure	191
Figure 3.8	The catalytic cycle of CYP450cam	192
Figure 3.9	Crystallization phase diagram for proteins	197
Figure 3.10	Illustration of the conventional crystallization methods	198
Figure 3.11	SDS-PAGE of protein expression	208
Figure 3.12	Representative whole cell CO-difference spectrum of CYP450cam after expression	209
Figure 3.13	SDS-PAGE CYP450cam(C334A)-His ₆ after affinity purification	210
Figure 3.14	SDS-PAGE CYP450cam(C334A) after anion exchange purification	211
Figure 3.15	SDS-PAGE CYP450cam(C334A)	212
Figure 3.16	Representative SDS-PAGE of pure CYP450cam used for crystallization shown in a concentration range	212
Figure 3.17	Absorbance spectrum of CYP450cam to determine sample purity	213
Figure 3.18	Absorbance spectrum of CYP450cam to determine sample purity and heme incorporation after hemin addition	214
Figure 3.19	Representative whole cell CO-difference spectrum of CYP450cam after addition of hemin	214
Figure 3.20	Absorbance spectrum of CYP450cam to determine sample purity and heme incorporation after the Lee <i>et al.</i> protocol was used for heme incorporation	215
Figure 3.21	Representative CO-difference spectrum of CYP450cam after the Lee <i>et al.</i> protocol was used for heme incorporation	215
Figure 3.22	Blue-native PAGE of CYP450cam(C334A) shown in a concentration range	216
Figure 3.23	Crystallization trials using sitting drop vapor diffusion	217
Figure 3.24	Crystallization trials using large scale sitting drop vapor diffusion	218
Figure 3.25	Crystallization trials using microbatch under paraffin oil	219
Figure 3.26	Crystallization trials using microdialysis buttons	220
Figure 3.27	Crystallization screens in sitting drop vapor diffusion	220
Figure 3.28	Crystallization screens in sitting drop vapor diffusion	221
Figure 3.29	Active site of CYP450cam with camphor bound	224
Figure 3.30	CYP450cam Asp251 salt bridge interactions	225
Figure 3.31	Interactions of the B'-helix in CYP450cam	225
Figure 4.1	The catalytic cycle of CYP450cam	236
Figure 4.2	The open (PDB 3L61) and closed conformation (PDB 2CPP) of CYP450cam	237
Figure 4.3	The CYP450cam-Pdx complex	239
Figure 4.4	Position of the CYP450cam L358P mutation	242
Figure 4.5	Proposed CYP450cam allosteric binding site	243

Figure 4.6	SDS-PAGE of CYP450cam mutant expression.....	253
Figure 4.7	Representative SDS-PAGE of pure CYP450cam(Crosslink) samples	254
Figure 4.8	Representative CO-difference spectrum of CYP450cam(Crosslink).....	254
Figure 4.9	SDS-PAGE of purified Pdx(Crosslink) shown in a concentration range.....	255
Figure 4.10	Structure of the covalent CYP450cam-Pdx complex.....	255
Figure 4.11	SDS-PAGE of components of the crosslink reaction mixture and the formed CYP450cam-Pdx complex	256
Figure 4.12	SAXS analysis of CYP450cam at 1000 μ M camphor	260
Figure 4.13	The radius of gyration (R_g) of CYP450cam as a function of camphor concentration	261
Figure 4.14	Pair distribution function $P(r)$ of CYP450cam	262
Figure 4.15	The radius of gyration (R_g) of CYP450cam(L358P) as a function of concentration	264
Figure 4.16	Pair distribution function $P(r)$ of CYP450cam(L358P)	265
Figure 4.17	Comparison SAXS data from CYP450cam and the CYP450cam-Pdx complex.....	266
Figure 4.18	The radius of gyration (R_g) of the CYP450cam-Pdx complex as a function of camphor concentration	267
Figure 4.19	Pair distribution function $P(r)$ of the CYP450cam-Pdx complex	267
Figure 4.20	Overlay of calculated and experimental SAXS curves determined by FOXS and SREFLEX	269
Figure 4.21	Structural elements in CYP450cam (PDB 1DZ4) with increased mobility found by SREFLEX normal mode analysis	271
Figure 4.22	Structural alignment of SREFLEX model with original CYP450cam crystal structure.....	274
Figure 4.23	Structural alignment of SREFLEX model with original CYP450cam(L358P) crystal structure.....	275
Figure 4.24	Structural alignment of SREFLEX model with original CYP450cam-Pdx complex crystal structure	276
Figure 5.1	Sitting drop crystallization of CYP450cam with needle-like crystals and precipitate formation	288

Chapter 1 : Characterization of metalloproteins using neutron protein crystallography

1.1 Metalloprotein catalysis: structural and mechanistic insights into oxidoreductases from neutron protein crystallography

The following work was reprinted with permission from: Schröder, G. C. & Meilleur, F. Metalloprotein catalysis: structural and mechanistic insights into oxidoreductases from neutron protein crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **D77**, (2021).

1.1.1 Synopsis

Neutron protein crystallography provides insight into the structure and reaction mechanism of transition state metal oxidoreductases without resulting in radiation damage-induced artefacts.

1.1.2 Abstract

Metalloproteins catalyze a range of reactions, with enhanced chemical functionality due to their metal cofactor. The reaction mechanism of metalloproteins has been experimentally characterized by spectroscopy, macromolecular crystallography and cryo-electron microscopy. An important caveat in structural studies of metalloproteins remains the artefacts that can be introduced by radiation damage. Photoreduction, radiolysis and ionization deriving from the electromagnetic beam used to probe the structure complicate structural and mechanistic interpretation. Neutron protein diffraction remains the only structural probe that leaves protein samples devoid of radiation damage, even when data is collected at room temperature. Additionally, neutron protein crystallography provides information on the position of light atoms such as hydrogen and deuterium allowing characterization of protonation states and hydrogen bonding networks. Neutron protein crystallography has been used in conjunction further experimental and computational techniques to gain insight into the structure and reaction mechanism of several transition state metal oxidoreductases with iron, copper and manganese cofactors. Here we review the contribution of neutron protein crystallography toward elucidating the reaction mechanism of metalloproteins.

1.1.3 Keywords

Neutron protein crystallography, X-ray diffraction, metalloproteins, enzymatic mechanism, protonation, radiation damage.

1.1.4 Introduction

Metals play a central role in biology, particularly in association with the proteome where they have catalytic, electron transfer, structural and storage roles ¹. Metal cofactors provide proteins with enriched functional range by expanding their physiochemical properties and playing a role in activation and stabilization during catalysis ^{2,3}. The reaction mechanism of metalloproteins has been explored with a plethora of techniques including structural, spectroscopic and computational studies ⁴. To gain a more complete understanding of protein interactions and catalysis, a complete, all-atom structure is a requisite. Neutron protein crystallography remains the sole structural technique to determine hydrogen atom positions without radiation induced damage at room temperature and cryo-conditions making it particularly valuable for the study of metalloproteins and their reaction mechanism ⁵⁻⁸.

1.1.4.1 Metals in biology

Metalloproteins comprise approximately a third of structures deposited in the Protein Data Bank (PDB) ⁹. Within metalloenzymes, metal cofactors are coordinated to protein residues in the primary coordination sphere where liganded amino acids finely tune the metal reactivity and redox potentials for optimal catalysis ^{10,11}. The chemistry afforded by the metal cofactor is determined primarily by the identity of the metal, however the secondary coordination sphere has a considerable influence on enzyme reactivity ¹². The secondary coordination sphere is composed of the residues proximal but not directly bonded to the metal ion, plays an important role in reactivity, while also ensuring substrate specificity and regioselectivity ¹³. Of particular interest are metal-cofactor containing oxidoreductases that are involved in the critical processes of electron exchange between donor and acceptor molecules in reactions as varied as oxygen activation and insertion, electron transfer, hydride transfer and hydrogen abstraction ¹⁴. Oxidoreductases comprise approximately 44% of the Enzyme Commission classes (EC Class 1) and are coordinated to a range of metals including iron, copper, manganese, nickel and zinc ¹⁵. Gaining a detailed structural

understanding of the primary and secondary coordination spheres is central to elucidating the reaction mechanism of these metalloproteins which will help inform improvements for pharmaceutical and biotechnological applications¹⁶.

1.1.4.2 Structural studies and radiation damage

An essential aspect of elucidating the reaction mechanism of metalloproteins is the ability to visualize the position of hydrogen atoms in amino acid residues, water molecules and reaction intermediates since hydrogen atoms are central to enzyme chemistry¹⁷⁻²¹. Several structural techniques have been implemented to break this subatomic resolution barrier and thereby inform protein chemistry including macromolecular crystallography and more recently cryo-electron microscopy (cryo-EM)^{22,23}. While X-ray crystallography with the advancement of synchrotron source's high-brilliance beams can provide ultra-high resolution structural data, very few hydrogen atoms are visible and geometric structure analysis as well as novel computational techniques such as quantum crystallography have to be implemented²⁴⁻²⁶. Micro electron diffraction (MicroED) which uses crystals smaller than 5 μm shows promise for hydrogen atom visualization since hydrogen atoms scatter electrons more strongly when compared to X-rays^{27,28}. MicroED structures with atomic resolution have been collected^{29,30}, however clear hydrogen atom visualization is thus far limited to small-molecule structures³¹. Cryo-EM, a structural technique not requiring a crystalline sample, has made stunning advances in hydrogen atom visualization with structures of apoferritin and $\beta 3$ GABA_A clearly indicating hydrogen atom positions in omit maps^{32,33}. X-ray crystallography, microED and cryo-EM all suffer from beam induced radiation damage³⁴⁻³⁶. The photosensitive cofactors and metal centers in metalloproteins are particularly sensitive to photo-induced damage^{37,38}. In multiple structures collected by X-ray radiation, photoreduction is observed which introduces artifacts that complicate structural and mechanistic interpretation^{16,39-42}. Indeed, metal-center photo-induced damage can be attributed not only to high energy radiation sources, since it has been found that UV/Vis radiation may also induce photoinactivation⁴³. Photo-induced damage is often accompanied by changes in bond lengths to the metal center, a loss/gain in ligands or radiolysis⁴⁴⁻⁴⁷. X-ray induced damage can be circumvented by use of serial femtosecond crystallography (SFX) or serial femtosecond rotation crystallography (SF-ROX) with

a X-ray free-electron laser (XFEL) that provides a brief, intense X-ray pulse to take multiple partial diffraction “snapshots” from many small crystals or one large crystal before significant beam-induced damage^{48,49}. A similar serial data collection approach can be applied to mitigate damage in electron microscopy and diffraction techniques^{30,50,51}.

1.1.4.3 Neutron protein crystallography

In contrast to X-ray diffraction, where the scattering intensity is proportional to the number of electrons making light elements like hydrogen poorly visible, the coherent scattering lengths of hydrogen and its isotope deuterium are comparable in magnitude to the backbone carbon (C), nitrogen (N), oxygen (O) atoms as well as metal atoms such as iron (Fe) and copper (Cu) (Figure 1.1)^{52,53}. As a consequence, hydrogen/deuterium atoms are visible in neutron scattering length density (NSLD) maps at moderate resolutions of up to 2.5 Å making assignment of protonation states and water molecule orientations possible^{54,55}.

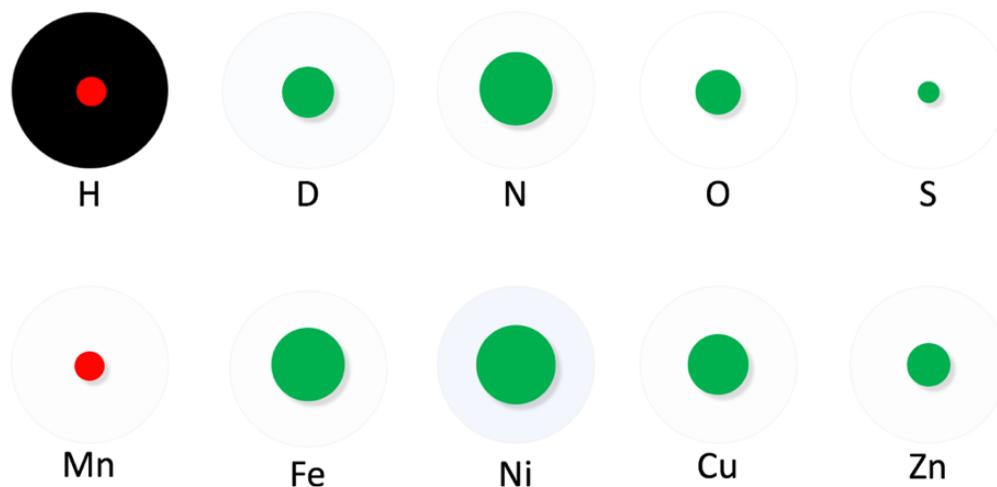


Figure 1.1 Incoherent neutron scattering cross sections and coherent neutron scattering lengths for selected elements. Relative incoherent scattering cross sections are represented by the disk coloured on a grey scale (dark: high incoherent cross section, light: low incoherent cross section), and relative coherent scattering lengths are represented by the red and green disks. The red disk for hydrogen and manganese indicates the negative sign of their scattering length while those shown in green are positive.

In order to successfully perform a neutron protein diffraction experiment, there are technique-inherent limitations that must be taken into consideration. Firstly, since hydrogen has a negative coherent

neutron scattering length and has a high incoherent cross section, it is necessary to exchange the hydrogen atoms present in the protein crystal with its isotope deuterium (D/H^2)^{56,57}. This can be accomplished by exchange of H with D at titratable sites, while the non-exchangeable sites remain hydrogenated. H/D exchange can be performed by vapor exchange or soaking of the hydrogenated protein crystal with deuterated crystallization buffer or by directly crystallizing the hydrogenated protein in deuterated buffer^{58,59}. An alternative to performing H/D exchange is perdeuteration, in which the non-exchangeable, carbon-bound, H atoms are also exchanged to D by protein expression in deuterated media which provides a fully deuterated protein sample⁶⁰. Substitution of H with D results in a decrease of incoherent scattering cross section from 80.26 b to 2.05 b, which results in an improved signal and a reduction of background noise during data collection (Figure 1.1). Furthermore, the negative coherent neutron scattering length of H with a value of -3.74 fm in contrast to the positive scattering lengths for D, C, N and O (Figure 1.1) leads to density cancellation in NSLD maps which can complicate interpretation. While H/D exchange partially alleviates NSLD map cancellations, cancellation can be fully circumvented by use of a perdeuterated protein samples which result in continuous NSLD maps, thereby aiding modelling of side-chains. However, it should be noted that protein perdeuteration results in lower yields and may impede crystallization.

Secondly, in metalloproteins, metal with weak positive or negative scattering lengths, such as zinc (5.68 fm) or manganese (-3.73 fm) are poorly visible in NSLD maps. It is therefore advantageous to perform a joint refinement in which the X-ray data can be used to determine the metal atom position and identity. X-ray diffraction data is typically collected following neutron data collection using the same crystal or using a different crystal but grown under identical conditions. The two datasets are subsequently used for a joint refinement in which the heavy atoms of the protein backbone and side-chain coordinates are refined against the X-ray data, while the neutron data is used to determine the H/D positions⁶¹. X-ray datasets collected on metalloproteins should however be collected using the lowest X-ray dose possible to limit radiation-induced damage. It is additionally advisable to perform an X-ray radiation dose exposure series to establish specific effects that X-ray beam exposure may have on the metalcenter. This will assist in

identification of any X-ray induced artefacts during joint refinement. Thirdly, since the flux of neutron sources is low when compared to high brilliance X-ray sources or even home-source X-ray generators, large protein crystals with a minimum volume of 0.1 mm³ are required and data collection requires several days^{62,63}. The interested reader is referred to extensive descriptions of the technique by O'Dell *et al.* and Schröder *et al.* for practical application^{56,64}.

1.1.5 Neutron protein crystallography of transition state metal oxidoreductases

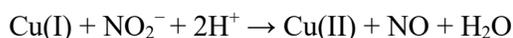
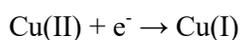
The sensitivity of neutron protein crystallography to the position of hydrogen atoms while not inducing radiation damage makes it an ideal technique to probe the structure of metalloproteins. Presented here is an overview of the structural and mechanistic insight gained from neutron protein diffraction of various transition-metal containing proteins in the oxidoreductase (EC 1) family (Table 1.1). While these examples demonstrate the power of neutron protein diffraction to characterize the active site and reaction intermediates, they also illustrate the value of neutron protein crystallography as a complementary tool. Neutron protein crystallography provides an all-atom model of the metalloprotein active site and intermediates which can inform quantum chemistry calculations of the electronic rearrangements and bond breakage and formation in the redox metalloprotein mechanism. Metalloproteins can be modelled using density functional theory (DFT) calculations to determine the energetics of a limited number of atoms, often confined to active site residues and intermediates^{65,66}. Quantum mechanics/molecular mechanics (QM/MM) modelling of metalloprotein uses QM to model the electronic structure of the active site and MM, described by a force field, to model the effects of the surrounding protein environment⁶⁷. Both DFT and QM/MM calculations are computationally expensive, making complete knowledge of the active site, particularly protonation states and hydrogen bond orientations, important to inform the calculations and minimize the number of input models. Use of additional techniques such as XFEL, spectroscopy, mutagenesis and kinetic studies provide a complete picture of the underlying enzyme chemistry as will be discussed herein.

Table 1.1 Neutron protein crystallography structures of transition state metal oxidoreductases.

Protein	PDB ID	Sequence length	Metal cofactor	Conditions	Deuteration level	Temperature (K)	Neutron resolution (Å)	X-ray resolution (Å)	Space group	Unit cell a, b, c (Å)
Copper nitrite reductase (<i>Achromobacter cycloclastes</i>)	6GTJ	340	Cu ²⁺	Resting state, pD 5.4	Perdeuterated	293	1.8	N/A	P2 ₁ 3	97.98, 97.98, 97.98
Copper nitrite reductase (<i>Geobacillus thermodenitrificans</i>)	6L46	323	Cu ²⁺	Resting state, pD. 5.3	Hydrogenated (H/D exchange)	100	1.50	1.30	H3	114.21, 114.21, 83.69
Copper amine oxidase (<i>Arthrobacter glibiformis</i>)	6L9C	621	Cu ²⁺	Cofactor TPQ-bound	Hydrogenated (H/D exchange)	100	1.72	1.14	C121	157.55, 61.78, 92.33
Ferric Cytochrome C Peroxidase (<i>Saccharomyces cerevisiae</i>)	4CVI	294	Fe ³⁺	Ferric heme	Hydrogenated (H/D exchange)	290	2.41	2.10	P2 ₁ 2 ₁ 2 ₁	51.70, 76.80, 107.60
Ferric Cytochrome C Peroxidase (<i>Saccharomyces cerevisiae</i>)	4CVJ	294	Fe ⁴⁺	Compound I heme	Hydrogenated (H/D exchange)	100	2.50	2.18	P2 ₁ 2 ₁ 2 ₁	51.19, 75.83, 107.59
Ferric Ascorbate Peroxidase (<i>Glycine max</i>)	5JPR	261	Fe ⁴⁺	Compound II heme	Hydrogenated (H/D exchange)	100	2.20	1.81	P4 ₂ 2 ₁ 2	82.10, 82.10, 75.16
Ferric Ascorbate Peroxidase (<i>Glycine max</i>)	6TAE	261	Fe ³⁺	Ferric heme	Perdeuterated	100	2.22	1.90	P4 ₂ 2 ₁ 2	82.90, 82.90, 75.87
Ferric Ascorbate Peroxidase (<i>Glycine max</i>)	6XV4	261	Fe ³⁺	Co-substrate ascorbic acid-bound	Perdeuterated	100	2.09	1.90	P4 ₂ 2 ₁ 2	81.86, 81.86, 74.97
Lytic polysaccharide monooxygenase (<i>Neurospora crassa</i>)	5TKI	223	Cu ²⁺	Resting state	Hydrogenated (H/D exchange)	298	2.12	1.50	P12 ₁ 1	68.12, 42.23, 70.29
Lytic polysaccharide monooxygenase (<i>Jonesia denitrificans</i>)	5VG1	142	Cu ²⁺	Dioxygen species	Hydrogenated (H/D exchange)	295	2.10	N/A	P2 ₁ 2 ₁ 2 ₁	32.50, 76.40, 122.10
Super Oxide Dismutase (<i>Homo sapiens</i>)	7KKS, 7KKU	199	Mn ³⁺	Oxidized	Perdeuterated	296	2.20	2.02	P6 ₁ 22	81.30, 81.30, 241.84
Super Oxide Dismutase (<i>Homo sapiens</i>)	7KKW, 7KLB	199	Mn ²⁺	Reduced	Perdeuterated	296	2.30	2.16	P6 ₁ 22	81.33, 81.33, 242.88
Chlorite dismutase (<i>Cyanothece PCC7425</i>)	5NKU	188	Fe ³⁺	Ferric heme, pD 9.4	Hydrogenated (H/D exchange)	293	2.35	2.00	P1	52.43, 53.02, 55.34
Amicyanin (<i>Paracoccus denitrificans</i>)	3L45	105	Cu ²⁺	Resting state	Hydrogenated (H/D exchange)	293	1.80	1.50	P12 ₁ 1	27.54, 56.58, 28.86
Manganese Catalase (<i>Thermus thermophilus</i> HB27)	6KK8	302	Mn ³⁺	Resting state	Hydrogenated (H/D exchange)	293	2.35	1.37	P2 ₁ 3	133.40, 133.40, 133.40

1.1.5.1 Copper nitrite reductase

Copper nitrite reductases (CuNiRs) (E.C. 1.7.2.1) form part of the anaerobic respiratory pathway of denitrification⁶⁸. During denitrification, microbes use nitrate (NO_3^-) as electron acceptors in a four step reductase-dependent process composed of sequential formation of NO_2^- , NO, N_2O and final release of N_2 ⁶⁹. CuNiRs are periplasmic enzymes that catalyze the reduction of NO_2^- by Cu(I) with the subsequent addition of two protons to form water and NO as follows⁷⁰:



The structure of CuNiRs has been well-characterized consisting of a homotrimer which contains two cupredoxin-type domains: a type I (T1Cu) domain involved in electron transfer from an electron donor partner protein to the type II (T2Cu) catalytic domain^{71,72}. The copper of the T1Cu domain is coordinated by two histidine residues, a cysteine and a methionine in a tetrahedral geometry⁷³. The catalytic T2Cu center is located ~12 Å from T1Cu and the two sites are connected via an electron transfer bridge composed of cysteine and histidine. The T2Cu copper is coordinated by three histidine residues in the equatorial plane. Of particular catalytic importance are the active site residues Asp_{CAT} and His_{CAT} that play a role in proton transfer and substrate binding and are connected by a bridging water molecule via hydrogen bonds^{74,75}. Extensive structural studies have been performed on CuNiRs to determine their catalytic mechanism, however X-ray induced reduction results in structural changes and unintended redox reactions⁷⁶⁻⁸². To circumvent the X-ray induced radiation damage, structures were obtained using X-ray free-electron laser (XFEL) crystallography^{71,83-85}. To further characterize the mechanism of CuNiR by investigating active site protonation states and substrate binding in the absence of radiation damage, Halsted *et al.* collected serial femtosecond rotational crystallography (SF-ROX) data on the oxidized, reduced and substrate-bound forms of *Achromobacter cycloclastes* CuNiR (*AcNiR*) as well as a room temperature neutron protein diffraction structure on the oxidized form⁸⁶. SF-ROX is an approach that utilizes XFEL

data collection by performing consecutive exposures following stepwise rotation and translation along a large crystal with known orientation, in contrast to the single-use exposure of randomly oriented micro crystals⁴⁸. Asp_{CAT} is known to be present in two conformations in CuNiRs, a “proximal” orientation facing T2Cu and hydrogen bonded with His_{CAT} via the bridging water and a “gatekeeper” orientation facing away from T2Cu and hydrogen bonding to the copper water ligand⁷⁹. Analysis of the SF-ROX data of the oxidized form revealed a T2Cu active site with an Asp_{CAT} in the proximal conformation with two orientations as opposed to the single proximal orientation observed in synchrotron data. In contrast, the neutron data collected at 1.9 Å resolution indicated a proximal Asp_{CAT} in only one orientation. Unexpectedly, several previously determined XFEL structures have noted the absence of the T2Cu axial water ligand^{83–85}, however the SF-ROX data presented by Halsted *et al.* indicated that a water molecule is coordinated to the T2Cu in the axial position which is confirmed by the analysis of the NSLD maps (Figure 1.2 A). The NSLD maps further indicate that Asp_{CAT} is deprotonated while His_{CAT} is N^{δ2} protonated at the catalytically optimal pD of 5.4 in contrast to previous spectroscopic and computational studies that proposed that the resting state His_{CAT} is doubly protonated⁸⁷. The NSLD additionally support this His_{CAT} single protonation state by confirming the positioning of the bridging water as a hydrogen bond donor to both Asp_{CAT} and the N^{ε1} of His_{CAT}. The authors conclude that the observed protonation states of the *AcNiR* active site support the hypothesis that upon binding of the nitrite substrate and displacement of the axial water ligand, protonation of Asp_{CAT} is triggered via the bridging water. This protonation of Asp_{CAT} initiates proton coupled electron transfer from T1Cu and subsequent catalysis^{87,88}

An additional CuNiR neutron protein crystallography structure was subsequently solved from *Geobacillus thermodenitrificans* (*GtNiR*) by Fukuda *et al.* at 1.5 Å resolution⁸⁹. Contrary to the finding on *AcNiR* by Halsted *et al.*, they observed two ligands coordinated to T2Cu. One of these ligands is an equatorial water while the second was fitted as a two-atom axial hydroxide ligand according to the NSLD maps (Figure 1.2 B). The Asp_{CAT} residue was in the “proximal” conformation in one orientation only and

was found to be deprotonated in agreement with previous findings, however the His_{CAT} was found to be doubly protonated at the measured pD of 5.3. In this *GtNiR* structure, at an improved resolution, the bridging water acts as a hydrogen bond donor to Asp_{CAT}, while it is a hydrogen bond acceptor to the His_{CAT} N^{ε1}. The authors reason that the difference in protonation may be due to a proton exchange between His_{CAT} and the axial water observed in *AcNiR*. The presence of a stable hydroxide ligand in the neutron structure is similar to the hydroxide ligand observed bound to T2Cu in computational studies suggesting direct release of NO without forming a copper-nitrosyl intermediate^{87,90,91}. The authors therefore conclude that their structure evidences that the nitrite reduction proceeds without formation of a copper nitrosyl intermediate, although this is in disagreement with a NO-soaked and multiple serial structures from one crystal (MSOX) studies that indicate a Cu-nitrosyl species^{78,81,92,93}. The level of H/D exchange was further analyzed to draw conclusions on the dynamics of the structure, particularly as it pertains to the electron transfer pathway from T1Cu to T2Cu. The bridging histidine that links the two copper sites shows comparatively low H/D exchange on its N^{δ2} site (low D atom occupancy) which suggests that electron transfer proceeds across this rigid hydrogen bond linking the cysteine backbone carbonyl and the histidine N^{δ2} hydrogen. Such an electron jump across a rigid hydrogen bond has been proposed by DFT calculations, and these neutron diffraction structural findings represent the first experimental evidence thereof⁹⁴.

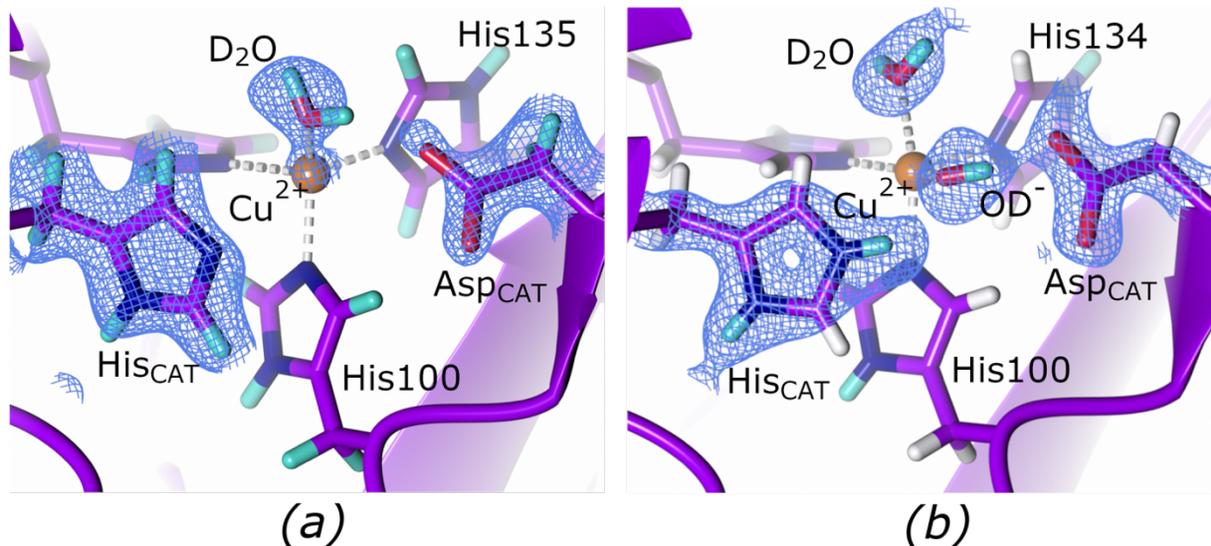


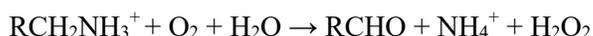
Figure 1.2 The active site of copper nitrite reductase (PDB code 6GTJ, perdeuterated ⁸⁶) **(A)** A D₂O is bound to the active site with a neutral His256 (His_{SCAT}). 2F_o-F_c NSLD map ($\sigma = 1.00$) is displayed as blue mesh; H and D atoms displayed in white and turquoise respectively. **(B)** An OD⁻ is bound to the active site with a positively charged His 211 (His_{SCAT}) (PDB code 6L46, H/D exchanged ⁸⁹). 2F_o-F_c NSLD map ($\sigma = 1.50$) is displayed as blue mesh; H and D atoms displayed in white and turquoise respectively.

Neutron protein diffraction studies provide valuable insight into protein chemistry; however, it is also useful to use further complementary techniques to obtain a more complete understanding of the system. Such is the case for CuNiR, in which recently solved atomic resolution damage free XFEL structures revealed new mechanistic insights which provide a new perspective on the neutron diffraction findings ⁹⁵. Rose *et al.* used advanced synchrotron radiation and XFEL data to obtain substrate-free, substrate-bound and product-bound structures of a *Bradyrhizobium* copper nitrite reductase (*Br*^{2D}NiR). Using unrestrained refinement of their high-resolution structures they show that His_{SCAT} is single N^{δ2} protonated in agreement with Halsted *et al.* ⁸⁶. Rose *et al.* also find that Asp_{CAT} is in the “proximal” conformation with two orientations as seen in *Ac*NiR. To generate the *Br*^{2D}NiR enzyme-product complex, crystals were soaked with nitrite, reduced with dithionite after which XFEL data was collected. Analysis revealed the presence of a Cu-nitrosyl species, refuting the hypothesis that NO is immediately released following proton coupled electron transfer leaving a hydroxide bound copper intermediate ⁸⁹. The presence of the hydroxide intermediate in the *Gt*NiR is suggested to be an artefact which can be attributed to the additional artificial coppers present in the neutron structure. These damage-free XFEL and neutron diffraction data findings

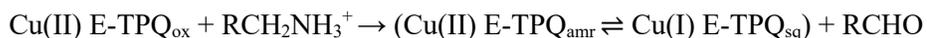
highlight the strengths of multiple techniques to gain insight into the CuNiR reaction mechanism while also allowing previous experimental findings to be critically evaluated.

1.1.5.2 Copper amine oxidase

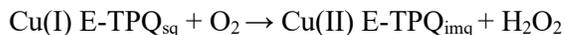
Copper amine oxidases (CAOs) (E.C. 1.4.3.6) are a class of redox enzymes present in prokaryotes and eukaryotes involved in the oxidative deamination of primary amines ⁹⁶. The overall reaction of CAOs results in the formation of an aldehyde from its corresponding primary amine after which molecular oxygen is reduced to hydrogen peroxide and ammonia is released to restore the enzyme resting state according to the general equation given below ⁹⁷.



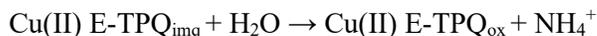
CAOs were initially described as “Pink Enzymes” due to the 2,4,5-trihydroxyphenylalanine quinone (TPQ) cofactor derived from a post-translationally modified tyrosine present in the active site with a visible absorbance peak at 480 nm ⁹⁸. The structure of CAOs has been structurally characterized as a mushroom-shaped homodimer with a 70 – 95 kDa molecular mass with the active site located in a C-terminal β -sandwich domain ⁹⁹. CAO catalysis is classified as a ping-pong bi-bi mechanism composed of an oxidative and reductive half-reaction ^{100,101}. In the reductive half-reaction, the resting state TPQ cofactor present in its oxidative enolate form, TPQ_{ox}, is nucleophilically attacked by the substrate amine ¹⁰². Following formation of the substrate- and product-Schiff bases (TPQ_{ssb} and TPQ_{psb}) the aldehyde product is released resulting in the formation of the aminoresorcinol (TPQ_{amr}) in equilibrium with the semiquinone radical (TPQ_{sq}) ¹⁰³.



The oxidative half-reaction proceeds with oxidation of TPQ_{sq} to the iminoquinone (TPQ_{imq}) by two electron reduction and protonation of molecular oxygen to form hydrogen peroxide.



Hydrolysis of TPQ_{imq} results in formation of the resting state TPQ_{ox} with the concomitant release of ammonia.



The active-site copper is believed to play no role in the reductive half-reaction¹⁰⁴. During the oxidative half-reaction, the copper is involved in mediating electron transfer between TPQ_{amr} and dioxygen and binding the reduced dioxygen species¹⁰⁵. The CAO reaction mechanism involves several intermediates in which proton transfer and protonation states play an important role^{106,107}.

To further characterize the active site of CAO from *Arthrobacter globiformis* (AGAO), *Murakawa et al.* collected a 1.72 Å resolution neutron data set of a resting state AGAO together with a 1.14 Å X-ray diffraction dataset for joint refinement¹⁰⁸. Analysis of the neutron scattering length density (NSLD) maps revealed that the resting state enolate (TPQ_{ox}) was in equilibrium with its keto form (Figure 1.3). This was evident from the ellipsoidal positive NSLD difference density peak at the C3 position of TPQ, indicating a mixed -CD and CD_2 state characteristic of the enolate and ketone form, respectively. The high-resolution X-ray diffraction data indicated that the TPQ quinone ring had a bent conformation which was further supported by DFT calculations which indicated a 59% keto and 41% enolate occupancy. Aspartic acid has been proposed to function as a catalytic base, by abstracting a proton from the substrate amine, the substrate Schiff base (TPQ_{ssb}) as well as during TPQ_{imq} hydrolysis to TPQ_{ox} ^{106,109}. In their neutron diffraction structure, *Murakawa et al.* emphasize the capability on neutron diffraction to reveal new and exciting conformations by presenting evidence for a triply shared proton between TPQ and the conserved aspartic acid residue Asp298 in a trifurcated hydrogen bond identified by a positive NSLD difference density peak. Albeit an uncommon protonation state, a triply shared proton is not unprecedented, having been observed

in a 5'-methylthioadenosine nucleosidase ¹¹⁰. However, a subsequent QM/MM study investigating the observed protonation states of Asp298 and TPQ found that a triply shared proton was energetically unfavorable, with bonding to one of the two carboxyl oxygen atoms representing a more stable conformation ¹¹¹. Although computational studies found this triply shared proton to be unstable, neutron diffraction data corroborates the role of Asp298 as a catalytic base with a highly reactive proton.

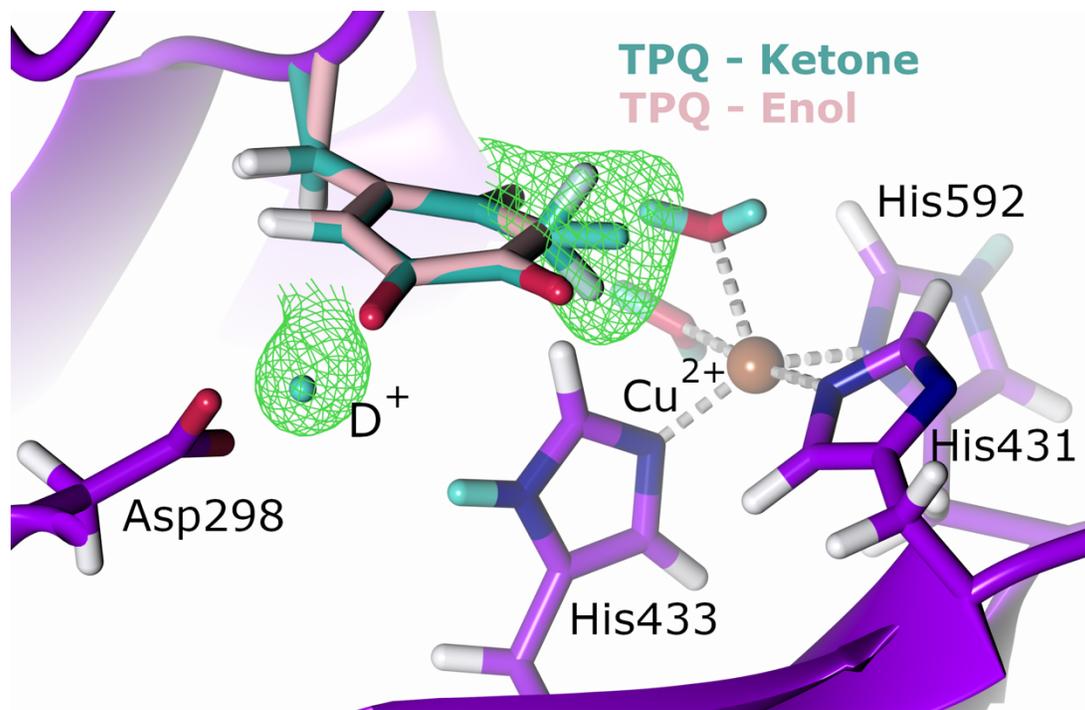


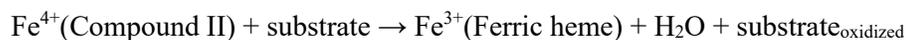
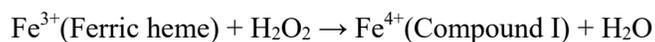
Figure 1.3 The active site of copper amine oxidase. The cofactor TPQ is present in its ketone and enol form and a proton is shared between TPQ and the active site Asp298 (PDB code 6L9C, H/D exchanged ¹⁰⁸). $F_o - F_c$ NSLD omit map ($\sigma = 3.00$) is displayed as green mesh; H and D atoms displayed in white and turquoise respectively.

The AGAO active site copper is coordinated by two water molecules (W_{ax} and W_{eq}) and three histidine residues (His431, His 433 and His592) with square-pyramidal geometry ¹¹². *Murakawa et al.* assigned His431 as a negatively charged imidazolate anion, potentially due to the ability of positively charged metal ions to significantly lower the apparent pKa of imidazole rings thereby resulting in deprotonation ¹¹³. Similarly, the coordinating W_{ax} was modelled in its deprotonated state, as a hydroxide anion, which potentially donates a proton to the enolate form of TPQ to form the observed mixed-state keto form. Additional QM/MM calculations suggest that His431 is potentially present as a imidazolate ion

dependent on the surrounding hydrogen bonding network, demonstrating the importance of these networks and protonation pathways in reaction mechanism¹¹¹. The QM/MM calculations could not definitively show a stabilized W_{ax} deprotonated as a hydroxide anion. Despite requiring further experimental and computational data to fully elucidate the AGAO reaction mechanism and unique protonation states observed in the AGAO neutron diffraction structure, these results provide an informative perspective on copper-containing amine oxidases and will help to inform on enzymes displaying similar chemistry.

1.1.5.3 Heme peroxidases

Heme peroxidases catalyze the hydrogen peroxide-dependent oxidation of a variety of substrates¹¹⁴. Oxygen activation during catalysis involves formation of two oxidized ferryl (Fe^{IV}) heme intermediates termed Compound I and Compound II^{115,116}.



Compound I forms first in the reaction cycle and contains either a porphyrin π -cation radical or a protein radical, while further one-electron reduction results in Compound II formation¹¹⁷. The chemical nature of the ferryl species has been an area of extensive research, with debate on whether the species should be characterized as a deprotonated iron(IV)-oxo ($Fe(IV)=O$) or a protonated iron(IV)-hydroxide ($Fe(IV)-OH$)¹¹⁸⁻¹²¹. The identity of the ferryl intermediate has been probed using multiple spectroscopic methods including resonance Raman, EXAFS and Mössbauer, however none were definitive^{122,123}. Structural studies using X-ray crystallography sought to determine the identity by differentiating between the shorter $Fe(IV)=O$ and longer $Fe(IV)-OH$ iron oxygen atom distances, however in-beam photoreduction results in increased bond length which complicates species assignment^{124,125}. An important step toward unravelling the identity of the ferryl species came with neutron protein diffraction studies on the Class I heme peroxidases cytochrome *c* peroxidase (CcP) (E.C. 1.11.1.5) and ascorbate peroxidase (APX) (E.C.

1.11.1.11) ¹²⁶. CcP and APX are intracellular proteins that contain a heme iron coordinated to a proximal histidine residue ^{127,128}. CcP is present in the mitochondria and plays a role in the electron transport chain by catalyzing the reduction of H₂O₂ to water by receiving reducing equivalents from cytochrome *c* (Cc) ¹²⁹. Upon reaction of the ferric CcP with H₂O₂, Compound I is formed with the radical located on a tryptophan residue adjacent to the heme ¹³⁰. Casadei *et al.* successfully used neutron protein diffraction to characterize the ferric form of CcP at 2.4 Å resolution at room temperature and were able to identify the protonation state of the ferryl Compound I species under cryo-conditions with a structure at 2.5 Å resolution ¹³¹. Analysis of the NSLD maps revealed that the ferric form contained a water coordinated to the heme on the distal face as well as a neutral distal histidine. Joint refinement of cryo-trapped Compound I established the identity of the ferryl species as the deprotonated Fe(IV)=O and indicated a positively charged, doubly protonated distal histidine – a thus far unprecedented finding since it had been assumed that both protons of H₂O₂ are used in formation of the water molecule. Although the iron oxygen bond length could not be accurately determined at the resolution of the neutron structures, NSLD omit maps clearly indicate a deprotonated Compound I species. This calls into question the previously established role of the distal histidine as a peroxide deprotonating base catalyst and subsequent acid catalyst for water formation to form Compound I ¹³². Casadei *et al.* posit that an additional proton is needed for Compound I formation, similar to Compound I formation in cytochrome P450s ¹³³. Following the identification of Compound I, Kwon *et al.* were able to trap Compound II of APX in a further cryo-neutron diffraction study ¹³⁴. APX catalyzes the H₂O₂-dependent oxidation of ascorbate resulting in formation of water and monodehydroascorbate ¹³⁵. The formation of Compound I is rapid, however the subsequent decay into Compound II is stable for a sufficient timespan to allow for cryo-trapping. Analysis of the APX Compound II 2.2 Å resolution NSLD maps is consistent with the presence of a protonated Fe(IV)-OH species (Figure 1.4 A). The distal histidine residue was also found to be doubly protonated and positively charged. The presence of an Fe(IV)-OH species, although supported by convincing NSLD maps, is suggested to be unlikely because an electron donating proximal ligand such as the thiolate in cytochrome P450s is necessary as opposed to the histidine proximal

ligand in heme peroxidases¹³⁶. This is supported by a recent study that utilized Mössbauer and X-ray absorption spectroscopy to show that the APX Compound II species is an unprotonated Fe(IV)=O¹³⁷.

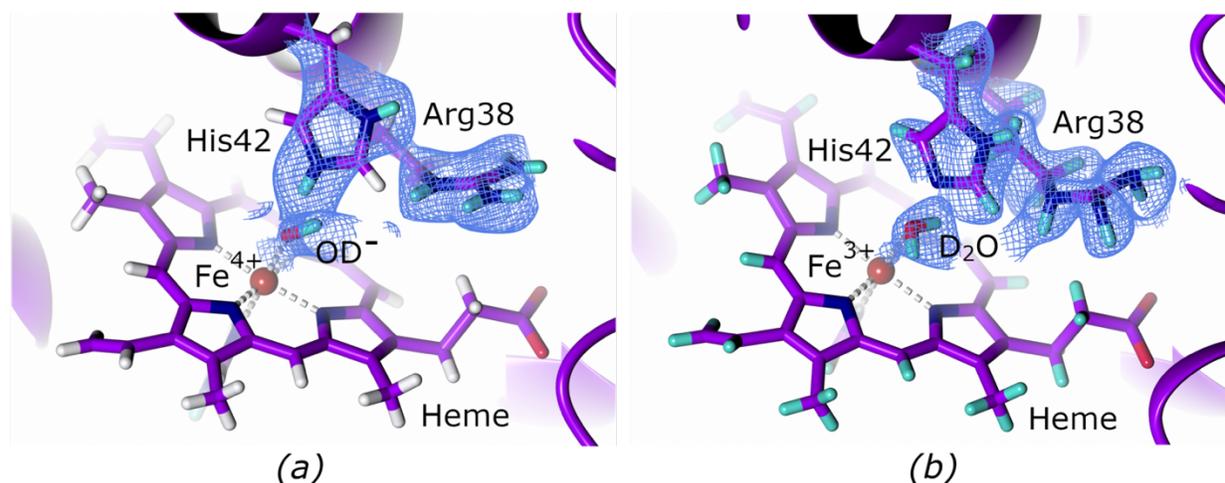


Figure 1.4 The active site of the heme-center ascorbate peroxidase (APX). **(A)** Compound II of APX with a positively charged His42 (PDB code 5JPR, H/D exchanged¹³⁴). $2F_o-F_c$ NSLD map ($\sigma = 1.50$) is displayed as blue mesh; H and D atoms displayed in white and turquoise respectively. **(B)** A neutral Arg38 residue in ascorbate-bound APX (PDB code 6XV4, perdeuterated¹³⁸). $2F_o-F_c$ NSLD map ($\sigma = 1.00$) is displayed as blue mesh; H and D atoms displayed in white and turquoise respectively.

A matter of contention is the 1.88 Å iron-oxygen bond length observed in the neutron study, which is considered to be much longer than the 1.76 Å bond length expected for a histidine coordinated Fe(IV) species and suggests a Fe(III) hydroxide species^{139,140}. The strengths of neutron diffraction to provide insight into the presence of light atoms and allow conclusions on protonation state can be juxtaposed to its weakness in assigning metal oxidation states making the use of complementary techniques such as spectroscopy crucial. A closer examination of the bond lengths in Compound II of CcP and APX has been performed by Kwon *et al.* using radiation damage-free XFEL diffraction to accurately establish the iron-oxygen bond length¹⁴¹. CcP displays a Fe-O distance of 1.76 Å while APX has a 1.87 Å bond length. The range of bond lengths observed led Kwon *et al.* to conclude that the bond length may “flex” due to protein dynamic motions in the heme active site. Indeed, a distal arginine residue in APX has been suggested to be dynamic and involved in proton delivery to the active site heme¹⁴², and the XFEL study indicates that the arginine has both heme-facing “in” and heme-distant “out” conformations. A 2.1 Å resolution neutron

protein diffraction structure of ferric APX bound to ascorbate by Kwon *et al.* indicates that this distal arginine is neutral, supporting its role as a dynamic proton donor (Figure 1.4 B) ¹³⁸. The dynamic hydrogen bonding rearrangement observed in APX, is not observed in CcP, where the distal arginine is present in only one conformation. This may explain the longer Fe-O bond distance in APX when compared to CcP, since the more dynamic heme hydrogen bonding environment may fine-tune the ferryl species. The implications that such fine-tuning may have for heme reactivity in different peroxidases remains to be established, however these findings highlight the effect that the active site environment can have on protein reactivity and selectivity.

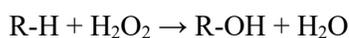
1.1.5.4 Lytic polysaccharide monooxygenase

The ability of enzymes to efficiently catalyze complex reactions under mild reaction conditions when contrasted to traditional chemical processes has made them invaluable for the production of value-added compounds ¹⁴³. In particular, the efficient degradation of cellulose by biocatalytic means for bioethanol is coming increasingly to the forefront as demands for sustainable energy production rise ¹⁴⁴. Fungi, in particular, employ numerous glycoside hydrolases to degrade lignocellulose as an energy source ¹⁴⁵. These glycoside hydrolases have been employed for commercial cellulose saccharification, however the complex structure of lignocellulose, particularly its biotic-attack resistant crystalline cellulose core limited solubilized sugar yields ^{146,147}. The discovery of a class of copper-dependent oxidoreductase termed lytic polysaccharide monooxygenases (LPMOs) (E.C. 1.14.99.56), however, resulted in increased cellulose yields by disruption of the crystalline cellulose structure by oxidative cleavage ^{148,149}. Fungal LPMOs are secreted proteins with a planar active site located at the enzyme surface for carbohydrate substrate binding ¹⁵⁰. The active site copper is coordinated by a characteristic “histidine brace” in the equatorial plane composed on the N-terminal histidine amino group and imidazole ring nitrogen and a further imidazole nitrogen atom from a second conserved histidine ¹⁵¹. LPMOs perform oxidative glycosidic bond cleavage by catalyzing hydroxylation of the C1 or C4 position of the glycosidic bond ¹⁵². In the O₂-dependent pathway, the reaction is initiated by one electron reduction of the resting state Cu(II) to Cu(I) after which oxygen binds and is spontaneously reduced to superoxide (O₂^{•-}). Rapid superoxide formation following one

electron reduction of an equatorially bound oxygen by Cu(I) has been shown by electron paramagnetic resonance (EPR) and stopped-flow absorption spectroscopy studies and was supported by DFT studies with coordination geometry informed by X-ray absorption near edge structure (XANES) and extended X-ray absorption fine structure (EXAFS) of the oxidized and reduced copper center and ¹⁵³. Following superoxide formation, the LPMO requires further precisely timed reduction and protonation steps to form the hydroxylated product, as illustrated in the overall reaction below, however the intermediates involved in the mechanism remains an area of active experimental and computational investigation ¹⁵⁴.



An essential step in elucidating the LPMO mechanism is characterization of the hydrogen atom abstracting species (HAA), a highly reactive species that abstracts a hydrogen atom from the glycosidic carbon to be hydroxylated ¹⁵⁵. Multiple candidates have been proposed using structural, spectroscopic, small molecule copper complexes and computational studies including superoxide ^{156–158}, hydroperoxy ¹⁵⁹, hydroxy ¹⁶⁰ and oxyl species ^{154,161–164}. Additionally, it has been proposed that hydrogen peroxide (H₂O₂) acts as the co-substrate in a reaction mechanism proceeds via an oxyl HAA intermediate ^{165,166}. This reaction mechanism requires the resting state Cu(II) to be reduced to Cu(I) in a “priming reduction” after which multiple hydroxylation reactions can be catalyzed.



In an initial study to investigate oxygen binding in LPMO, Bacik *et al.* collected a 1.1 Å X-ray diffraction dataset and a 2.1 Å neutron diffraction dataset on a chitin-binding LPMO from *Jonesia denitrificans* (JdLPMO10A) ^{167,168}. The X-ray structure was interpreted as containing an oxygen cofactor modelled as a peroxide species coordinated in the equatorial plane. The crystal asymmetric unit contained two molecules, Molecule A and B, related by non-crystallographic symmetry. The dioxygen species was

modelled with a side-on coordination to the Molecule A copper and an end-on coordination to the molecule B copper. In the separately solved neutron structure, only molecule B was modelled with an end-on coordinated dioxygen species, however limited resolution did not permit unambiguous determination of the chemical nature of the dioxygen species. The X-ray and neutron data deposited by Bacik *et al.* was later revisited by Caldararu *et al.* who performed a joint X-ray neutron quantum refinement of the originality deposited data ¹⁶⁹. The joint quantum refinement noted a discrepancy in the positioning of the dioxygen species, with the Molecule A dioxygen more resembling an end-on positioning following introduction of a QM potential. Subsequent vacuum QM and QM/MM calculations informed by the quantum refined X-ray data-only structure suggest that the dioxygen species may be a superoxo species, however the discrepancy of the binding modes of the dioxygen species in Molecule A and B, the weak nuclear density of the dioxygen species in Molecule A and the absence of a reducing agent for one electron reduction of the resting state Cu(II) to ensure dioxygen binding made these initial findings of dioxygen activation ambiguous. To further investigate dioxygen activation at the LPMO active site and identify the activated oxygen species involved in the initial steps of catalysis, O'Dell *et al.* collected both high resolution X-ray and room temperature neutron diffraction dataset of LPMO9D from *Neurospora crassa* (*NcLPMO9D*) ¹⁷⁰. In order to characterize the activated oxygen species, *NcLPMO9D* crystals were chemically reduced to Cu(I) and the resulting intermediate was freeze-trapped prior to X-ray diffraction data collection. Structural analysis of the 1.20 Å data revealed the first evidence of a dioxygen species (potentially superoxo or peroxy) coordinated to the copper with η_1 end-on geometry – an exciting development in the characterization of the early steps of LPMO dioxygen activation. Analysis of the second *NcLPMO9D* molecule in the crystal structure related by non-crystallographic symmetry did not contain a coordinated activated dioxygen species, however an oxygen could be modelled in a “pre-binding” conformation, presumably prior to coordinating at the equatorial position. To gain more detailed insight into the protonation state of second-shell residues that may play a role in catalysis, a 2.12 Å structure of the Cu(II) resting state form of *NcLPMO9D* was solved using room temperature neutron protein crystallography. Analysis of the NSLD maps collected at room temperature revealed that the residue His157 is singly N^{ε2}-protonated (Figure 1.5). His157 is a second shell

residue pointing toward the oxygen pre-binding site and may therefore play a role in oxygen activation by promoting oxygen activation. The NSLD maps did not indicate the presence of a pre-bound oxygen, however data were collected at pH 5.6 (pD 6.0) while LPMOs perform catalysis in a more acidic environment of pH~5 where cellulases show optimum activity^{147,171,172}. It is therefore posited that under such acidic conditions, His157 would be positively charged and doubly N^{ε2} and N^{δ1}-protonated which may promote oxygen binding. DFT calculations confirmed that while oxygen binding to the neutral His157 is thermoneutral, oxygen binding to the positively charged His157 is highly thermodynamically favorable.

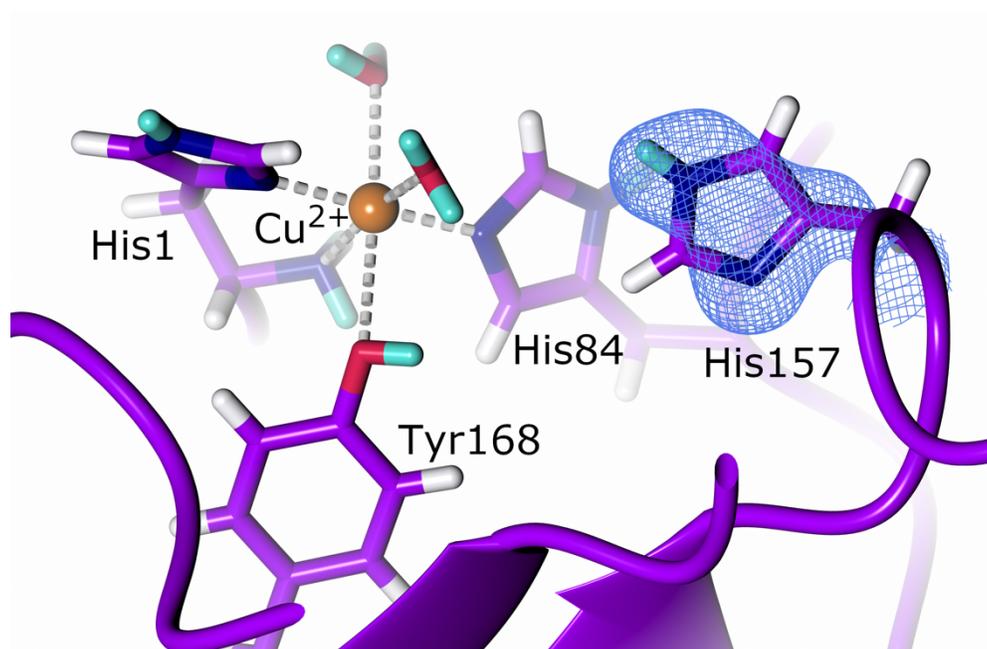


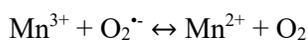
Figure 1.5 The active site of LPMO. The second-shell His157 is neutral in the copper(II) resting state. (PDB code 5TKI, H/D exchanged¹⁷⁰). $2F_o - F_c$ NSLD map ($\sigma = 1.50$) is displayed as blue mesh; H and D atoms displayed in white and turquoise respectively.

These findings show that His157 may play a role in ensuring efficient catalysis and illustrate the importance of second-shell residues. Indeed, DFT calculations of the LPMO mechanism assign this second-shell histidine a positive charge and have postulated that it may play a role as a proton donor of active site intermediates^{173,174}. The exact mechanism by which LPMOs achieve catalysis remains to be elucidated and the poorly soluble carbohydrate substrate presents challenges in crystallographic structural analysis, although a limited number have been determined¹⁷⁵⁻¹⁷⁷. Future spectroscopic, computational and structural

studies, including neutron protein diffraction studies investigating the *Nc*LPMO9D structure under functionally relevant acidic conditions and bound to a cryo-trapped intermediate ¹⁷⁸, may help shed light on LPMO catalysis.

1.1.5.5 Manganese superoxide dismutase

Reactive oxygen species (ROS) play a role in cellular signaling, however accumulation of these oxygen ions or radicals can result in cellular and organ damage ¹⁷⁹. Superoxide dismutases (SOD) are metalloenzymes that function to mitigate the damage from ROS by converting superoxide ($O_2^{\cdot-}$) into oxygen (O_2) or hydrogen peroxide (H_2O_2) ¹⁸⁰. The active site of SODs can contain an iron, manganese, copper or zinc cofactor dependent on the taxonomic domain ¹⁸¹. In eukaryotes, manganese superoxide dismutase (MnSOD) (E.C. 1.15.1.1) is situated in the mitochondrial matrix where it prevents the accumulation of superoxide formed due to mitochondrial electron leaks ¹⁸². MnSOD is functional as a homotetramer with the active site Mn ion coordinated to the residues His26, His74, Asp159, His163 ¹⁸³. The reaction mechanism of MnSOD proceeds by cycling through a Mn(III) and Mn(II) oxidation state as shown below ¹.



Mn(II) oxidizes superoxide to molecular oxygen and the formed Mn(III) reduces the superoxide species with concomitant protonation to form hydrogen peroxide ¹⁸⁴. The MnSOD mechanism requires concerted proton and electron transfer (CPET) to achieve a thermodynamically favorable reaction during catalysis, as has been observed for multiple enzymes involved in redox reactions ¹⁸⁵. In addition to the active site residues, mutagenesis, structural and computational studies have shown that second-shell residues including His30, Tyr34, Gln143, Glu162, and Tyr166 play an important role in the active site metal redox cycling and protonation ^{186–191}. The strength of neutron protein crystallography to provide detailed

information on hydrogen atom positions is particularly advantageous for the study of CPET in a metalloenzyme in which redox states must be maintained and knowledge of protonation states provide information on proton transfer.

To this end, Azadmanesh *et al.* collected room temperature neutron diffraction data on an oxidized Mn(III)SOD and reduced Mn(II)SOD to a resolution of 2.20 Å and 2.30 Å, respectively¹⁹². The heavy atom positions of the MnSOD backbone and side-chains were determined using X-ray datasets of the reduced and oxidized forms and these coordinates were used as the starting models for neutron data-only refinement to avoid including X-ray induced artefacts at the manganese center. Analysis of the NSLD maps provided unprecedented insight into the protonation states of second shell residues particularly His30, Tyr34, Gln143 and Tyr166 while also revealing the role of active site waters and emphasizing the role of short-strong hydrogen bonds (SSHBs) and low-barrier hydrogen bonds (LBHBs) in stabilization during catalysis. Oxidized Mn(III) center was found to be five-coordinated with a hydroxide (termed WAT1) ligand, confirming earlier DFT calculations (Figure 1.6 A)¹⁹³⁻¹⁹⁶. WAT1 was stabilized by hydrogen bonding to Gln143 via its D⁶²² deuteron. Significantly, following reduction to Mn(II), WAT1 was protonated by this Gln143 to form water and a deprotonated glutamine present as an amide ion (Figure 1.6 B). This finding, although unexpected given a glutamine pKa in the range of 16-18, can be explained by active site pKa modulation¹⁹⁷ as well as the Trp123 residue that was found to form a SSHB with Gln143, potentially stabilizing the amide form. Furthermore, complementary DFT calculations supported the role of Gln143 as a proton donor to WAT1 as opposed to the previously proposed Tyr34¹⁹¹. The role of Tyr34 as a proton donor to WAT1 was further put into question by the finding that in the Mn(III) form Tyr34 is deprotonated forming a SSHB with a neighboring WAT2 molecule, ruling out its role as a proton donor upon Mn(III) reduction. One of the active sites of the reduced Mn(II) form revealed a sixth ligand, characterized as a hydroxide from NSLD maps, in the position opposite the coordinating Asp159 supporting previous structural findings¹⁹⁸. This hydroxide association is additionally characterized by bond lengthening between Mn(II) and Asp159 and a reduction in the electropositivity of Mn(II). It is hypothesized that this

represents an intermediate prior to WAT1 protonation since decrease in the Mn(II) positive charge by hydroxide coordination may result in electronegative polarization of the WAT1 hydroxide, thereby promoting Gln143 deprotonation. Analysis of the NSLD omit maps further revealed unusual protonation states of the active site channel His30 which is associated with Tyr166.

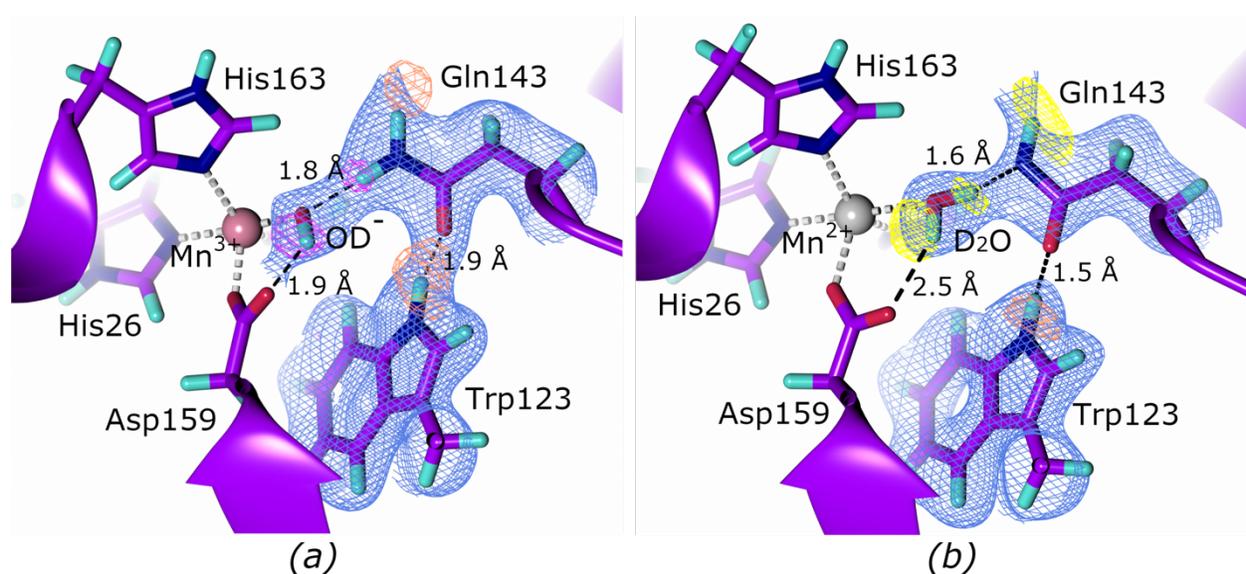


Figure 1.6 The active site of manganese superoxide dismutase. **(A)** The oxidized active site with a doubly protonated Gln143 (PDB code 7KKS, perdeuterated¹⁹²). $2F_o-F_c$ NSLD map ($\sigma = 1.00$) is displayed as blue mesh; H and D atoms displayed in white and turquoise respectively. F_o-F_c NSLD omit map is displayed as mesh with $\sigma = 3.00$ in orange and $\sigma = 3.50$ in magenta. **(B)** The reduced active site with a singly protonated Gln143 (PDB code 7KKW, perdeuterated¹⁹²). $2F_o-F_c$ NSLD map ($\sigma = 1.00$) is displayed as blue mesh; H and D atoms displayed in white and turquoise respectively. F_o-F_c NSLD omit map is displayed as mesh with $\sigma = 2.50$ in yellow and $\sigma = 3.00$ in orange.

Both His30 and Tyr166 are also intricately involved in the change of protonation states at the active site. In the oxidized Mn(II), Azadmanesh *et al.* observe a doubly deprotonated His30 imidazolate anion with the nearest D atom refined as being associated to the hydroxyl group of Tyr166. The authors interpret this as Tyr166 interchanging between an ionized and protonated form that interacts with His30 that alternates in a concerted fashion between a deprotonated imidazolate anion and N^{ε2}-protonated form. Although unusual, histidine residues in an imidazolate form have been observed and may play an important role in changes of protonation states during catalysis¹⁹⁹. Upon reduction to the Mn(II) form, His30 was found in a N^{δ1}-protonated form with WAT2 having provided the proton, while a LBHB forms between

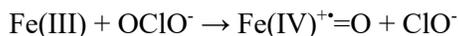
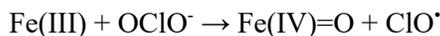
Tyr166 and the N^{e2} of His30. These changes in protonation state in the Tyr166-His30 pair can be related to the changes in protonation at the active site via the solvent molecule WAT2 which is involved in change in protonation states and is replenished from bulk solvent upon catalytic use. The insights gained from analysis of the neutron crystallography structures allowed Azadmanesh *et al.* to propose a reaction mechanism in which changes to metal redox states are coupled to two internal proton transfers as well as two external proton transfers from solvent molecules. The unexpected finding that Gln143 is the WAT1 proton donor is significant for potential future studies of MnSOD upon interaction with substrate or, indeed, with cryogenically trapped intermediates.

The neutron protein diffraction structures of the oxidized and reduced form of MnSOD provide a wealth of insight into CPET and the important role that second-shell residues have in modulation of active site protonation states during catalysis. The neutron structures revealed novel proton donors and unexpected protonation states which may assist in the unravelling of the reaction mechanism in further oxidoreductases in which electron and indeed proton transfers are crucial.

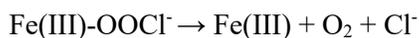
1.1.5.6 Chlorite dismutase

The generation of a covalent oxygen-oxygen bond is thus far a biologically rare reaction that has only been attributed to Photosystem II and a nitric oxide dismutase²⁰⁰⁻²⁰². A third enzyme, chlorite dismutase, has however garnered much interest for its ability to form an oxygen-oxygen bond during turnover of chlorite (OCIO⁻) to chloride (Cl⁻) and molecular dioxygen (O₂)²⁰³. Chlorite dismutases (Clds) are a family of heme *b*-containing oxidoreductases (EC 1.13.11.49) present in prokaryotes²⁰⁴. Clds can be divided into two clades which differ in their oligomeric organization and subunit structure. Clade 1, the “long” Clds, are present as homopentamers or homohexamers and contain a α -helix rich N-terminal domain and a heme *b*-containing ferredoxin-like domain^{205,206}. Clade 2, on the other hand, are termed “short” Clds, form homodimers and are shorter than their Clade 1 counterparts because they lack the N-terminal α -helix rich domain while maintaining a similar C-terminal domain^{207,208}. The reaction mechanism of Clds is postulated to follow either a homolytic or heterolytic route following chlorite binding²⁰⁹. During homolytic

cleavage²¹⁰, chlorine monoxide (ClO[•]) and Compound II are produced, while heterolytic cleavage²¹¹ produces hypochlorite (ClO⁻) and Compound I as shown below:



In both cases, following chlorite cleavage, there is a rebinding step that produces peroxyhypochlorite (OOCl⁻) which subsequently results in formation of chloride and dioxygen²⁰³:



The active-site architecture of ClDs includes a proximal histidine that ligates the heme iron, as well as conserved glutamate, lysine and two tryptophan residues²⁰⁹. In the resting state, the distal side of the heme iron is occupied by a water molecule. The only conserved charged residue is an arginine in an otherwise hydrophobic active site. This arginine has been postulated to play a catalytic role and has been found to have an “inward” conformation pointing toward the heme and an “outward” conformation facing toward the substrate entry channel^{206,212}. This flexible arginine residue is believed to be involved in substrate recognition, enzyme stability as well as playing a role in homolytic/heterolytic cleavage and the subsequent intermediate recombination reaction^{204,209,213}. The turnover of ClDs show a strong dependence on pH, a feature that was linked to the pH-dependent protonation state of the distal arginine residue^{214,215}. In order to investigate the role of arginine in ClD catalysis and determine the role of its protonation state in chlorite degradation, Schaffner *et al.* collected a room temperature 2.35 Å neutron diffraction dataset at pH 9.0 as well as a 2.0 Å X-ray diffraction dataset on the same crystal for joint refinement on the clade 2 dimeric chlorite dismutase from *Cyanothece* sp. PCC7425 (CCld)²¹⁶. In addition, they collected X-ray data at pH 6.5 and 8.5 and performed stopped-flow, UV-vis and Resonance Raman spectroscopy to determine the role of pH and characterize the reaction mechanism. Analysis of the NSLD density maps at pH 9.0 of

the proximal active site residues indicated a rigid hydrogen bonding network composed of the coordinating His114, Glu167 and Lys92 whose H-bond lengths were mostly invariant with a change in pH when compared to the X-ray diffraction structures at pH 6.5 and 8.5. On the distal face, the iron was coordinated to a hydroxide (the deprotonated form of water molecule W501) and a second water molecule (W502). The active site arginine (Arg127) was found to be in the “outward” conformation stabilized by a hydrogen bond with Gln74 (Figure 1.7). Furthermore, Arg127 remained fully protonated at pH 9.0, ruling out its role as a distal base that modulates chlorite degradation during catalysis. These findings agree with studies that found that the guanidinium group in arginine remains charged at high pH in internal protein sites^{217,218}.

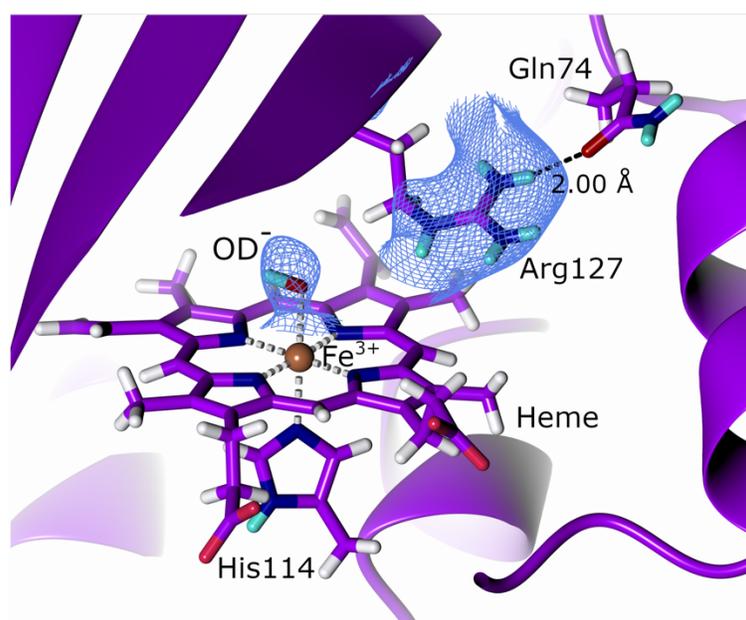


Figure 1.7 The active site of chlorite dismutase. The active-site Arg127 remains fully protonated in the outward conformation (PDB code 5NKU, H/D exchanged²¹⁶). $2F_o - F_c$ NSLD map ($\sigma = 1.40$) is displayed as blue mesh; H and D atoms displayed in white and turquoise respectively.

The structural and spectroscopic findings led the authors to conclude that CCl₂ most likely follows a homolytic cleavage pathway with formation of Compound II, however some Compound I formation observed at acidic conditions means that heterolytic cleavage cannot be unequivocally ruled out. In this mechanism, Arg127 plays a role in substrate recognition in its “outward” conformation and it is possible that it may adopt an “inward” conformation during catalysis by displacing W502 and stabilizing an

intermediate. The role of Arg127 in CCl₄ was further examined in mutagenesis studies which confirmed that its role in the catalytic efficiency of CCl₄ was small ²¹⁹. It was found that disrupting the Arg127 hydrogen bond to Gln74 by mutating it to a valine (Q74V) resulted in a more flexible Arg127 that could adopt an “inward” conformation in CCl₄, while an aspartate mutant (Q74E) locked the Arg127 in an “outward” conformation in a salt bridge. It was found that the conformational dynamics of Arg127 has little effect on CCl₄ catalysis since both mutants display similar catalytic efficiency which agrees with the neutron diffraction findings. It was found, however, that Arg127 plays a role in the thermal stability of CCl₄, with the more rigid Q74E showing increased thermal stability. Moreover, Arg127 acts as a gatekeeper to the active site and plays a role in heme coordination of intermediates during catalysis.

A strength of neutron protein diffraction as a probe is its absence of radiation damage and the coincident metal reduction that introduces artefacts in structures solved from X-ray diffraction ^{8,42}. In the neutron diffraction structure, Schaffner *et al.* observed a 2.33 Å distance between the iron and hydroxide oxygen atom, a distance the authors observed to be longer than expected ²¹⁶. Indeed, comparison of several structures indicated a correlation between ligand distances and the reducing power of the radiation used with distances measured following synchrotron radiation being the longest. The observations of the a longer iron-oxygen than anticipated could be attributed to the recent finding that chlorite dismutase shows unexpectedly high photosensitivity when exposed to UV-visible light ⁴³. This observation that even the visible light spectrum induces a degree of photo-inactivation in Cld provides perspective on the care that must be taken when studying photosensitive proteins, since the underlying chemistry may be more dependent on environmental and experimental factors than anticipated.

1.1.5.7 Amicyanin

Blue copper proteins are small type-I copper proteins that function as electron shuttles ²²⁰. Also termed cupredoxins, blue copper proteins function as intermediaries, accepting electrons from one molecule and donating them to another ²²¹. Amicyanin (E.C. 1.4.9.1) is a bacterial cupredoxin of 11.5 kDa containing a copper coordinated to two histidines, a cysteine, and a methionine (Figure 1.8 A) ²²². Inter- and intramolecular electron transfers play an important role in redox-reactions and cellular processes such as

respiration and photosynthesis, making an understanding of how proteins control of such electron transfer reactions crucial^{223,224}.

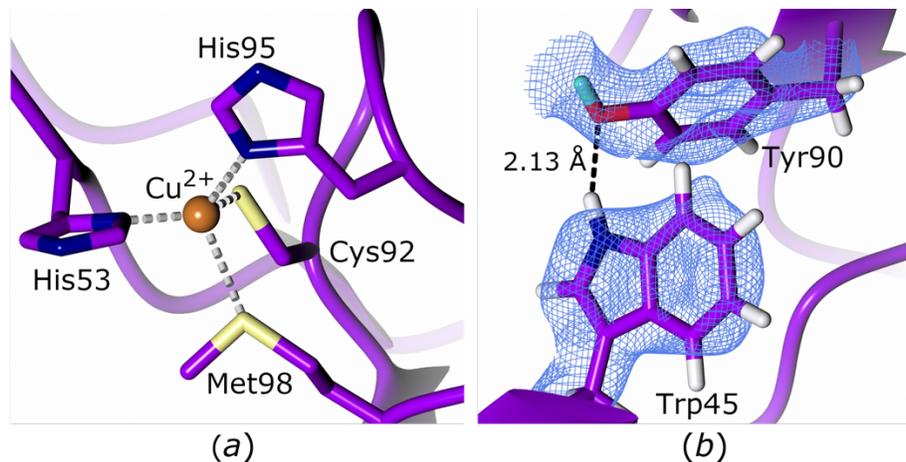


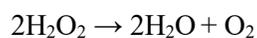
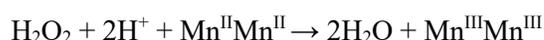
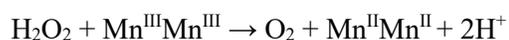
Figure 1.8 The active site of amicyanin and strong hydrogen bond with an unexchanged tryptophan(PDB code 3L45, H/D exchanged²²⁵). **(A)** Conserved active site residues coordinating the copper cofactor. **(B)** The hydrogen bond between the non-exchanged Trp45 and Tyr90. $2F_o-F_c$ NSLD map ($\sigma = 1.00$) is displayed as blue mesh; H and D atoms displayed in white and turquoise respectively.

In order to investigate the role of protein dynamics in electron transfer, Sukumar *et al.* performed a joint X-ray/neutron study on amicyanin from *Paracoccus denitrificans* that shuttles electrons from the tryptophan tryptophylquinone (TTQ) cofactor of methylamine dehydrogenase (MADH) to the heme of cytochrome *c*-551i^{224,225}. The copper of this amicyanin is coordinated by the imidazole nitrogens of His 53 and His95 and the sulfur atoms of Cys92 and Met98²²⁶. All three components of this ternary MADH:Amicyanin:cytochrome *c*-551i system have been structurally characterized, providing a structural framework for the neutron diffraction studies²²⁶⁻²²⁸. An X-ray dataset was collected to 1.5 Å resolution and used with a neutron diffraction dataset at 1.8 Å resolution for joint XN structure refinement to accurately determine the positions of hydrogen/deuterium (H/D) atoms and the extent of H/D exchange to investigate the dynamics and flexibility of amicyanin. Analysis of the electron density and NSLD maps revealed that 7 of 17 buried residues at an intramolecular depth of ≥ 3.8 Å displayed full H/D exchange with a valine residue at a depth of 5.5 Å also displaying full H/D exchange. This high level of exchange near the protein core indicates that the protein exhibits a high level of dynamic motion. In contrast, core structural elements

of the protein involved in the formation of β -strands did not exchange, confirming their role in protein structural stability. Additionally, 14 residues were found to have non-exchangeable hydrogen atoms, including Trp45 and Tyr90 which form a strong hydrogen bond that enhances thermal stability (Figure 1.8 B) ²²⁹. Residues in the vicinity of the copper as well as those involved in mediating electron transfer displayed significant levels of H/D exchange, further highlighting the dynamic nature of amicyanin while also supporting the role of these residues in long range electron transfer. These findings shed light on how protein dynamics can regulate electron transfer particularly the effect on electronic coupling. Electronic coupling, also termed H_{AB} , provides an indication of the likelihood that a reaction will take place when the activation energy is achieved ^{223,230}. Protein dynamics can alter the H_{AB} , by decreasing the distance of necessary through-space jumps during electron transfer and increasing the atomic packing density. These alterations would explain the surprisingly low calculated H_{AB} values for the amicyanin crystal structure when compared to the same system in solution ^{224,231}. The findings by Sukumar *et al.* illustrate the important role that protein dynamics has on reaction and transfer rate. In addition to examining H/D exchange levels, Sukumar *et al.* analyzed the presence and orientation of hydrogen bonds within the amicyanin structure. The observed hydrogen bonds in their oxidized amicyanin structure correspond well with those predicted by high-resolution X-ray diffraction structures, however they observed that the reduced form of amicyanin had five fewer hydrogen bonds confirming that the reduced state undergoes a conformational change ²³². Analysis of the number and type hydrogen bonds within 8Å of the copper indicated a high proportion of C-H \cdots X bonds which are weaker than conventional hydrogen bonds. The high number of these C-H \cdots X bonds may function collectively in stabilizing the of structure of amicyanin ²³³. The joint X-ray/neutron diffraction study presented by Sukumar *et al.* provides insight into the capability of neutron diffraction to study H/D exchange patterns to inform on protein dynamics and also reveal the importance of conventional and non-conventional H-bonds in the protein structure.

1.1.5.8 Manganese catalase

Hydrogen peroxide (H₂O₂) serves various functions in cellular metabolism, functioning as a signaling molecule and cellular pathway regulator, however this cytotoxic molecule can also be produced as a byproduct of aerobic metabolism, requiring rapid removal^{234–236}. Catalases are metalloproteins present in microbes, plants and animals and function to remove H₂O₂ before it results in potential oxidative cellular damage by degrading it into water and molecular oxygen²³⁷. There are two families of catalases with distinct cofactors, structures and chemistry: heme and non-heme catalases²³⁸. Heme catalases contain an iron porphyrin cofactor and degrade H₂O₂ *via* a Compound I intermediate²³⁹. Non-heme catalases, also known as manganese catalases (MnCat) (E.C. 1.11.1.6), have been found in bacteria and archaea²⁴⁰. Structural studies revealed that MnCat is a globular protein consisting of 30 kDa monomers forming a homohexameric structure^{241,242}. The active site contains two Mn ions connected by two solvent derived μ -oxo bridging molecules and a glutamic acid carboxyl group. The first coordination shell of this bimetallic core is completed by two glutamic acid and two histidine residues. The overall reaction of MnCat involves oxidative and reductive half reactions to get a net reaction in which two H₂O₂ molecules produce two molecules of water and molecular oxygen as shown below²⁴³:



In its resting state, MnCat is present as oxidized Mn^{III}Mn^{III}, which is reduced to Mn^{II}Mn^{II} during catalysis, however addition of oxidizing or reducing agents can also result in mixed valence states such as Mn^{II}Mn^{III} or Mn^{III}Mn^{IV}^{244–246}. During the MnCat reaction mechanism, the active site manganese molecules undergo changes in redox states and act as electron repositories in a reaction cycle that contains multiple intermediates with varied protonation states²⁴⁰. The identity of the bridging ligands during catalysis remains to be established, however it has been proposed that in the resting Mn^{III}Mn^{III} state, the active site contains

a μ -oxo and μ -OH⁻ bridge which converts to Mn^{II}Mn^{II} with a μ -OH⁻ and μ -H₂O bridge following H₂O₂ binding and oxidation ²⁴⁷. To structurally investigate the protonation states at the active site of the thermostable MnCat from *Thermus thermophilus*, Yamada *et al.* collected a 2.35 Å neutron diffraction dataset as well as a 1.37 Å X-ray diffraction dataset on the same crystal for joint refinement ²⁴⁸. Crystallization was performed under basic conditions to ensure Mn^{III}Mn^{III} oxidation state. Analysis of the NSLD maps of the active site confirmed the presence of the glutamate bridging ligand as well as a di- μ -oxodimanganese cluster (Figure 1.9 A). The μ -bridging oxygen atom closest to the μ -bridging Glu70 was found to be deprotonated in the μ -oxo form and formed a hydrogen bond to Thr39. However, as had also been observed in an earlier 1 Å resolution X-ray structure, the further active site bridging oxygen atom was disordered which complicated interpretation of the protonation state of the 2.35 Å resolution neutron structure ²⁴¹.

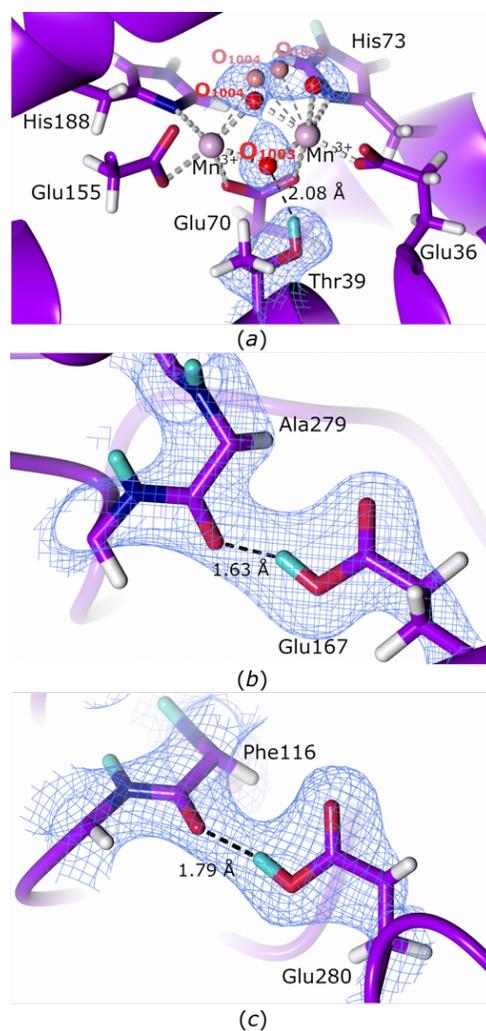


Figure 1.9 The active site and protonation states of manganese catalase glutamate residues with backbone carbonyl oxygen atoms (PDB code 6KK8, H/D exchanged²⁴⁸. $2F_o-F_c$ NSLD map ($\sigma = 1.00$) is displayed as blue mesh; H and D atoms displayed in white and turquoise respectively. **(A)** Manganese catalase active site with the bridging oxygen atoms coordinated to the two manganese cofactors. Bridging atoms O₁₀₀₄ and O₁₀₀₅ have alternate conformations shown in red and pale red. **(B)** Single protonation of Glu167 forming hydrogen bond with Ala279 carbonyl group oxygen. **(C)** Single protonation of Glu280 forming hydrogen bond with Phe116 carbonyl group oxygen.

The neutron diffraction structure also indicated protonation of Glu167 and Glu280 which are present on the hexamer boundaries, enabling them to form hydrogen bonds (Figure 1.9 B and 1.9 C). The protonated state of these glutamic acid residues under the basic crystallization conditions can be rationalized by the tight packing of the hexamer, effectively barricading these residues. While this neutron diffraction study did not allow complete interpretation of the bridging ligands at the active site, it represents an exciting first insight into the resting state of MnCat. The reaction mechanism of MnCat comprises several changes

of protonation and oxidation states which are ideal for study using neutron protein diffraction. Future studies following the binding of the H₂O₂ substrate or addition of reductants opens the possibility of further mechanistic findings.

1.1.6 Perspective

Visualization of hydrogen atoms during proton transfer, changes in protonation state and hydrogen bond network rearrangement is central to determining the reaction mechanism of oxidoreductases^{4,249}. A further important consideration in oxidoreductase mechanism is the delivery of electrons to drive catalysis. Electron delivery often occurs via a reductase partner protein through inter-protein electron transfer²⁵⁰. Association with the associated reductase often results in conformational changes and regulation of activity similar to the modulation by second-shell residues, as has been demonstrated for cytochrome P450cam (CYP450cam) in interaction with Putidaredoxin (Pdx)²⁵¹. Small angle neutron scattering (SANS) allows the association of the metalloprotein with its partner protein to be probed, providing the advantage of contrast matching through selective deuteration to highlight changes in only the component of interest while not inducing radiation damage⁷. Recent advances in these interactions has been made by SANS on the interactions of sulfite reductase (SiR)²⁵², cytochrome P450 reductase (CPR)²⁵³ and cellobiose dehydrogenase (CDH)^{5,254}. These SANS studies serve to further expand the insights gained into metalloproteins, thereby providing a more complete mechanistic picture. Neutron protein crystallography serves as a valuable complement to further structural techniques for elucidation of protein chemistry. As the flux of neutron sources and ability to cryo-trap intermediates in the protein reaction mechanism improves, our understanding of metalloprotein chemistry at an atomic level will greatly advance.

1.1.7 Conflict of interest

The authors declare no conflict of interest

1.1.8 Funding Information

GCS acknowledges support in part from the National Research Foundation (NRF), South Africa and from the Oak Ridge National Laboratory Graduate Opportunities (GO!) Program. FM acknowledges support from USDA NIFA Hatch 211001.

1.1.9 References

1. Holm, R. H., Kennepohl, P. & Solomon, E. I. Structural and functional aspects of metal sites in biology. *Chem. Rev.* **96**, 2239–2314 (1996).
2. Hemschemeier, A. & Happe, T. The plasticity of redox cofactors: from metalloenzymes to redox-active DNA. *Nat. Rev. Chem.* **2**, 231–243 (2018).
3. Natri, F. *et al.* Engineering Metalloprotein Functions in Designed and Native Scaffolds. *Trends Biochem. Sci.* **44**, 1022–1040 (2019).
4. Fontecilla-Camps, J. C. & Nicolet, Y. *Advances in Research on Metalloproteins*. **1122**, (Humana Press, 2014).
5. Bodenheimer, A. M., O'Dell, W. B., Stanley, C. B. & Meilleur, F. Structural studies of *Neurospora crassa* LPMO9D and redox partner CDHIIA using neutron crystallography and small-angle scattering. *Carbohydr. Res.* **448**, 200–204 (2017).
6. Meilleur, F., Coates, L., Cuneo, M. J., Kovalevsky, A. & Myles, D. A. A. The neutron macromolecular crystallography instruments at Oak Ridge national laboratory: Advances, challenges, and opportunities. *Crystals* **8**, 1–11 (2018).
7. Ashkar, R. *et al.* Neutron scattering in the biological sciences: progress and prospects. *Acta Crystallographica Section D: Structural Biology* **74**, (2018).
8. Meilleur, F., Kovalevsky, A. & Myles, D. A. A. IMAGINE: The neutron protein crystallography beamline at the high flux isotope reactor. in *Methods in enzymology* **634**, 69–85 (Elsevier Inc., 2020).
9. Putignano, V., Rosato, A., Banci, L. & Andreini, C. MetalPDB in 2018: A database of metal sites in biological macromolecular structures. *Nucleic Acids Res.* **46**, D459–D464 (2018).
10. Riordan, J. F. The role of metals in enzyme activity. *Ann. Clin. Lab. Sci.* **7**, 119–129 (1977).
11. Hosseinzadeh, P. & Lu, Y. Design and fine-tuning redox potentials of metalloproteins involved in electron transfer in bioenergetics. *Biochim. Biophys. Acta - Bioenerg.* **1857**, 557–581 (2016).
12. Maglio, O., Natri, F. & Lombardi, A. Structural and Functional Aspects of Metal Binding Sites in Natural and Designed Metalloproteins. *Ion. Interact. Nat. Synth. Macromol.* 361–450 (2012). doi:10.1002/9781118165850.ch11
13. Colquhoun, H. M., Stoddart, J. F. & Williams, D. J. Second-Sphere Coordination—a Novel Role for Molecular Receptors. *Angew. Chemie Int. Ed. English* **25**, 487–507 (1986).
14. Younus, H. Oxidoreductases: Overview and Practical Applications. in *Biocatalysis* (eds. Husain, Q. & Ullah, M. F.) 39–55 (Springer International Publishing, 2019). doi:10.1007/978-3-030-25023-2_3
15. Andreini, C., Bertini, I., Cavallaro, G., Holliday, G. L. & Thornton, J. M. Metal ions in biological catalysis: From enzyme databases to general principles. *J. Biol. Inorg. Chem.* **13**, 1205–1218 (2008).
16. Bowman, S. E. J., Bridwell-Rabb, J. & Drennan, C. L. Metalloprotein Crystallography: More than a Structure. *Acc. Chem. Res.* **49**, 695–702 (2016).
17. Engler, N., Ostermann, A., Niimura, N. & Parak, F. G. Hydrogen atoms in proteins: Positions and dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 10243–10248 (2003).
18. Hubbard, R. E. & Kamran Haider, M. Hydrogen Bonds in Proteins: Role and Strength. *Encycl. Life Sci.* (2010). doi:10.1002/9780470015902.a0003011.pub2
19. Bowie, J. U. Membrane protein folding: How important are hydrogen bonds? *Curr. Opin. Struct. Biol.* **21**, 42–49 (2011).
20. Woińska, M., Grabowsky, S., Dominiak, P. M., Woźniak, K. & Jayatilaka, D. Hydrogen atoms can be located accurately and precisely by x-ray crystallography. *Sci. Adv.* **2**, (2016).
21. Clabbers, M. T. B., Gruene, T., van Genderen, E. & Abrahams, J. P. Reducing dynamical electron scattering reveals hydrogen atoms. *Acta Crystallogr. Sect. A Found. Adv.* **75**, 82–93 (2019).
22. Förster, A. & Schulze-Briese, C. A shared vision for macromolecular crystallography over the next five years. *Struct. Dyn.* **6**, 064302 (2019).
23. Callaway, E. Revolutionary cryo-EM is taking over structural biology. *Nature* **578**, 201 (2020).

24. Neumann, P. & Tittmann, K. Marvels of enzyme catalysis at true atomic resolution: distortions, bond elongations, hidden flips, protonation states and atom identities. *Curr. Opin. Struct. Biol.* **29**, 122–133 (2014).
25. Takaba, K. *et al.* Subatomic resolution X-ray structures of green fluorescent protein. *IUCrJ* **6**, 387–400 (2019).
26. Cachau, R. E., Zhu, J. & Nicklaus, M. C. The upcoming subatomic resolution revolution. *Curr. Opin. Struct. Biol.* **58**, 53–58 (2019).
27. Nannenga, B. L. & Gonen, T. Protein structure determination by MicroED. *Curr. Opin. Struct. Biol.* **27**, 24–31 (2014).
28. Nannenga, B. L. & Gonen, T. MicroED opens a new era for biological structure determination. *Curr. Opin. Struct. Biol.* **40**, 128–135 (2016).
29. Luo, F. *et al.* Atomic structures of FUS LC domain segments reveal bases for reversible amyloid fibril formation. *Nat. Struct. Mol. Biol.* **25**, 341–346 (2018).
30. Nannenga, B. L. MicroED methodology and development. *Struct. Dyn.* **7**, (2020).
31. Palatinus, L. *et al.* Hydrogen positions in single nanocrystals revealed by electron diffraction. *Science*. **355**, 166–169 (2017).
32. Yip, K. M., Fischer, N., Paknia, E., Chari, A. & Stark, H. Atomic-resolution protein structure determination by cryo-EM. *Nature* **587**, 157–161 (2020).
33. Nakane, T. *et al.* Single-particle cryo-EM at atomic resolution. *Nature* **587**, 152–156 (2020).
34. Garman, E. F. & Weik, M. Radiation Damage in Macromolecular Crystallography. in *Protein Crystallography: Methods and Protocols* (eds. Wlodawer, A., Dauter, Z. & Jaskolski, M.) 467–489 (Springer New York, 2017). doi:10.1007/978-1-4939-7000-1_20
35. Hattne, J. *et al.* Analysis of Global and Site-Specific Radiation Damage in Cryo-EM. *Structure* **26**, 759-766.e4 (2018).
36. Clabbers, M. T. B. & Xu, H. Macromolecular crystallography using microcrystal electron diffraction. *Acta Crystallogr. Sect. D Struct. Biol.* **77**, 313–324 (2021).
37. Macedo, S. *et al.* Can soaked-in scavengers protect metalloprotein active sites from reduction during data collection? *J. Synchrotron Radiat.* **16**, 191–204 (2009).
38. Frankaer, C. G., Mossin, S., Ståhl, K. & Harris, P. Towards accurate structural characterization of metal centres in protein crystals: the structures of Ni and Cu T6 bovine insulin derivatives. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **70**, 110–122 (2014).
39. Corbett, M. C. *et al.* Photoreduction of the active site of the metalloprotein putidaredoxin by synchrotron radiation. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **63**, 951–960 (2007).
40. Wherland, S. & Pecht, I. Radiation chemists look at damage in redox proteins induced by X-rays. *Proteins Struct. Funct. Bioinforma.* **86**, 817–826 (2018).
41. De La Mora, E. *et al.* Structural changes caused by radiation-induced reduction and radiolysis: The effect of X-ray absorbed dose in a fungal multicopper oxidase. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 564–577 (2012).
42. Pfanzagl, V. *et al.* X-ray-induced photoreduction of heme metal centers rapidly induces active-site perturbations in a protein-independent manner. *J. Biol. Chem.* **295**, 13488–13501 (2020).
43. Mahor, D., Püschmann, J., Adema, D. R., Strampraad, M. J. F. & Hagedoorn, P. L. Unexpected photosensitivity of the well-characterized heme enzyme chlorite dismutase. *J. Biol. Inorg. Chem.* **25**, 1129–1138 (2020).
44. Horrell, S. *et al.* Serial crystallography captures enzyme catalysis in copper nitrite reductase at atomic resolution from one crystal. *IUCrJ* **3**, 271–281 (2016).
45. Yano, J. *et al.* X-ray damage to the Mn4Ca complex in single crystals of photosystem II: A case study for metalloprotein crystallography. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 12047–12052 (2005).
46. George, G. N. *et al.* X-ray-induced photo-chemistry and X-ray absorption spectroscopy of biological samples. *J. Synchrotron Radiat.* **19**, 875–886 (2012).
47. Nass, K. *et al.* Indications of radiation damage in ferredoxin microcrystals using high-intensity X-FEL beams. *J. Synchrotron Radiat.* **22**, 225–238 (2015).

48. Schlichting, I. Serial femtosecond crystallography: The first five years. *IUCrJ* **2**, 246–255 (2015).
49. Spence, J. C. H. XFELs for structure and dynamics in biology. *IUCrJ* **4**, 322–339 (2017).
50. Mastrorarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
51. de la Cruz, M. J., Martynowycz, M. W., Hattne, J. & Gonen, T. MicroED data collection with SerialEM. *Ultramicroscopy* **201**, 77–80 (2019).
52. Sears, V. F. Neutron scattering lengths and cross sections. *Neutron News* **3**, 26–37 (1992).
53. Blakeley, M. P., Langan, P., Niimura, N. & Podjarny, A. Neutron crystallography: opportunities, challenges, and limitations. *Curr. Opin. Struct. Biol.* **18**, 593–600 (2008).
54. Oksanen, E., Chen, J. C. H. & Fisher, S. Z. Neutron crystallography for the study of hydrogen bonds in macromolecules. *Molecules* **22**, 1–26 (2017).
55. Schröder, G. C., O'Dell, W. B., Myles, D. A. A., Kovalevsky, A. & Meilleur, F. IMAGINE: Neutrons reveal enzyme chemistry. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 778–786 (2018).
56. O'Dell, W. B., Bodenheimer, A. M. & Meilleur, F. Neutron protein crystallography: A complementary tool for locating hydrogens in proteins. *Arch. Biochem. Biophys.* **602**, 48–60 (2016).
57. Meilleur, F., Weiss, K. L. & Myles, D. A. A. Deuterium Labeling for Neutron Structure-Function-Dynamics Analysis. in *Methods in molecular biology (Clifton, N.J.)* **544**, 281–292 (2009).
58. Niimura, N. & Bau, R. Neutron protein crystallography: Beyond the folding structure of biological macromolecules. *Acta Crystallogr. Sect. A Found. Crystallogr.* **64**, 12–22 (2008).
59. Bennett, B. C., Gardberg, A. S., Blair, M. D. & Dealwis, C. G. On the determinants of amide backbone exchange in proteins: A neutron crystallographic comparative study. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **64**, 764–783 (2008).
60. Hazemann, I. *et al.* High-resolution neutron protein crystallography with radically small crystal volumes: Application of perdeuteration to human aldose reductase. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **61**, 1413–1417 (2005).
61. Adams, P. D., Mustyakimov, M., Afonine, P. V. & Langan, P. Generalized X-ray and neutron crystallographic analysis: More accurate and complete structures for biological macromolecules. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **65**, 567–573 (2009).
62. Ng, J. D. *et al.* Large-volume protein crystal growth for neutron macromolecular crystallography. *Acta Crystallogr. Sect. F Structural Biol. Commun.* **71**, 358–370 (2015).
63. Blakeley, M. P., Cianci, M., Helliwell, J. R. & Rizkallah, P. J. Synchrotron and neutron techniques in biological crystallography. *Chem. Soc. Rev.* **33**, 548–557 (2004).
64. Schröder, G. C. & Meilleur, F. Neutron Crystallography Data Collection and Processing for Modelling Hydrogen Atoms in Protein Structures. *J. Vis. Exp.* 1–38 (2020). doi:10.3791/61903
65. Siegbahn, P. E. M. & Himo, F. Recent developments of the quantum chemical cluster approach for modeling enzyme reactions. *J. Biol. Inorg. Chem.* **14**, 643–651 (2009).
66. Siegbahn, P. E. M. A quantum chemical approach for the mechanisms of redox-active metalloenzymes. *RSC Adv.* **11**, 3495–3508 (2021).
67. Van Der Kamp, M. W. & Mulholland, A. J. Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology. *Biochemistry* **52**, 2708–2728 (2013).
68. Zumft, W. G. Cell biology and molecular basis of denitrification. *Microbiol. Mol. Biol. Rev.* **61**, 533–616 (1997).
69. Yang, J. *et al.* A critical review of aerobic denitrification: Insights into the intracellular electron transfer. *Sci. Total Environ.* **731**, 139080 (2020).
70. Eady, R. R. & Hasnain, S. S. Denitrification. in *Comprehensive Coordination Chemistry II* 759–786 (Elsevier, 2003). doi:10.1016/B0-08-043748-6/08141-X
71. Horrell, S., Kekilli, D., Strange, R. W. & Hough, M. A. Recent structural insights into the function of copper nitrite reductases. *Metallomics* **9**, 1470–1482 (2017).
72. Merkle, A. C. & Lehnert, N. Binding and activation of nitrite and nitric oxide by copper nitrite reductase and corresponding model complexes. *Dalt. Trans.* **41**, 3355–3368 (2012).
73. Godden, J. *et al.* The 2.3 angstrom X-ray structure of nitrite reductase from *Achromobacter*

- cycloclastes. *Science*. **253**, 438–442 (1991).
74. Kataoka, K., Furusawa, H., Takagi, K., Yamaguchi, K. & Suzuki, S. Functional analysis of conserved aspartate and histidine residues located around the type 2 copper site of copper-containing nitrite reductase. *J. Biochem.* **127**, 345–350 (2000).
 75. Boulanger, M. J., Kukimoto, M., Nishiyama, M., Horinouchi, S. & Murphy, M. E. P. Catalytic roles for two water bridged residues (Asp-98 and His-255) in the active site of copper-containing nitrite reductase. *J. Biol. Chem.* **275**, 23957–23964 (2000).
 76. Murphy, M. E. P., Turley, S. & Adman, E. T. Structure of nitrite bound to copper-containing nitrite reductase from *Alcaligenes faecalis*: Mechanistic implications. *J. Biol. Chem.* **272**, 28455–28460 (1997).
 77. Lefterink, N. G. H. *et al.* Proton-coupled electron transfer in the catalytic cycle of *Alcaligenes xylosoxidans* copper-dependent nitrite reductase. *Biochemistry* **50**, 4121–4131 (2011).
 78. Tocheva, E. I., Rosell, F. I., Mauk, A. G. & Murphy, M. E. P. Side-On Copper-Nitrosyl Coordination by Nitrite Reductase. *Science*. **304**, 867–870 (2004).
 79. Antonyuk, S. V., Strange, R. W., Sawers, G., Eady, R. R. & Hasnain, S. S. Atomic resolution structures of resting-state, substrate- and product-complexed Cu-nitrite reductase provide insight into catalytic mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 12041–12046 (2005).
 80. Nojiri, M. *et al.* Structural basis of inter-protein electron transfer for nitrite reduction in denitrification. *Nature* **462**, 117–120 (2009).
 81. Tocheva, E. I., Rosell, F. I., Mauk, A. G. & Murphy, M. E. P. Stable copper-nitrosyl formation by nitrite reductase in either oxidation state. *Biochemistry* **46**, 12366–12374 (2007).
 82. Hough, M. A., Antonyuk, S. V., Strange, R. W., Eady, R. R. & Hasnain, S. S. Crystallography with Online Optical and X-ray Absorption Spectroscopies Demonstrates an Ordered Mechanism in Copper Nitrite Reductase. *J. Mol. Biol.* **378**, 353–361 (2008).
 83. Fukuda, Y. *et al.* Redox-coupled structural changes in nitrite reductase revealed by serial femtosecond and microfocus crystallography. *J. Biochem.* **159**, 527–538 (2016).
 84. Fukuda, Y. *et al.* Redox-coupled proton transfer mechanism in nitrite reductase revealed by femtosecond crystallography. *Proc. Natl. Acad. Sci.* **113**, 2928–2933 (2016).
 85. Halsted, T. P. *et al.* An unprecedented dioxygen species revealed by serial femtosecond rotation crystallography in copper nitrite reductase. *IUCrJ* **5**, 22–31 (2018).
 86. Halsted, T. P. *et al.* Catalytically important damage-free structures of a copper nitrite reductase obtained by femtosecond X-ray laser and room-temperature neutron crystallography. *IUCrJ* **6**, 761–772 (2019).
 87. Ghosh, S., Dey, A., Sun, Y., Scholes, C. P. & Solomon, E. I. Spectroscopic and computational studies of nitrite reductase: Proton induced electron transfer and backbonding contributions to reactivity. *J. Am. Chem. Soc.* **131**, 277–288 (2009).
 88. Brenner, S. *et al.* Demonstration of proton-coupled electron transfer in the copper-containing nitrite reductases. *J. Biol. Chem.* **284**, 25973–25983 (2009).
 89. Fukuda, Y., Hirano, Y., Kusaka, K., Inoue, T. & Tamada, T. High-resolution neutron crystallography visualizes an OH-bound resting state of a copper-containing nitrite reductase. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 4071–4077 (2020).
 90. Lintuluoto, M. & Lintuluoto, J. M. DFT Study on Nitrite Reduction Mechanism in Copper-Containing Nitrite Reductase. *Biochemistry* **55**, 210–223 (2016).
 91. Qin, X., Deng, L., Hu, C., Li, L. & Chen, X. Copper-Containing Nitrite Reductase Employing Proton-Coupled Spin-Exchanged Electron-Transfer and Multiproton Synchronized Transfer to Reduce Nitrite. *Chem. - A Eur. J.* **23**, 14900–14910 (2017).
 92. Merkle, A. C. & Lehnert, N. The side-on copper(I) nitrosyl geometry in copper nitrite reductase is due to steric interactions with isoleucine-257. *Inorg. Chem.* **48**, 11504–11506 (2009).
 93. Horrell, S. *et al.* Enzyme catalysis captured using multiple structures from one crystal at varying temperatures. *IUCrJ* **5**, 283–292 (2018).
 94. Hadt, R. G., Gorelsky, S. I. & Solomon, E. I. Anisotropic covalency contributions to superexchange

- pathways in type one copper active sites. *J. Am. Chem. Soc.* **136**, 15034–15045 (2014).
95. Rose, S. L. *et al.* An unprecedented insight into the catalytic mechanism of copper nitrite reductase from atomic-resolution and damage-free structures. *Sci. Adv.* **7**, (2021).
 96. Dooley, D. M. *et al.* Structure and Reactivity of Copper-Containing Amine Oxidases. in *Bioinorganic Chemistry of Copper* (eds. Karlin, K. D. & Tyeklár, Z.) 459–470 (Springer Netherlands, 1993). doi:10.1007/978-94-011-6875-5_37
 97. Wilmot, C. M. *et al.* Catalytic mechanism of the quinoenzyme amine oxidase from *Escherichia coli*: Exploring the reductive half-reaction. *Biochemistry* **36**, 1608–1620 (1997).
 98. Floris, G. & Mondovi, B. *Copper Amine Oxidases: Structures, Catalytic Mechanisms and Role in Pathophysiology*. (Taylor & Francis Group, 2009).
 99. Parsons, M. *et al.* Crystal structure of a quinoenzyme: copper amine oxidase of *Escherichia coli* at 2 Å resolution. *Structure* **3**, 1171–1184 (1995).
 100. Bardsley, W. G., Crabbe, M. J. C. & Scott, I. V. The amine oxidases of human placenta and pregnancy plasma. *Biochem. J.* **139**, 169–181 (1974).
 101. Brazeau, B. J., Johnson, B. J. & Wilmot, C. M. Copper-containing amine oxidases. Biogenesis and catalysis; a structural perspective. *Arch. Biochem. Biophys.* **428**, 22–31 (2004).
 102. Medda, R. *et al.* The reaction mechanism of copper amine oxidase: detection of intermediates by the use of substrates and inhibitors. *Biochemistry* **34**, 16375–16381 (1995).
 103. Klema, V. J. & Wilmot, C. M. The role of protein crystallography in defining the mechanisms of biogenesis and catalysis in copper amine oxidase. *Int. J. Mol. Sci.* **13**, 5375–5405 (2012).
 104. Rinaldi, A., Giartosio, A., Floris, G., Medda, R. & Agrò, A. F. Lentil seedlings amine oxidase: Preparation and properties of the copper-free enzyme. *Biochem. Biophys. Res. Commun.* **120**, 242–249 (1984).
 105. Kishishita, S. *et al.* Role of copper ion in bacterial copper amine oxidase: Spectroscopic and crystallographic studies of metal-substituted enzymes. *J. Am. Chem. Soc.* **125**, 1041–1055 (2003).
 106. Chiu, Y. C. *et al.* Kinetic and structural studies on the catalytic role of the aspartic acid residue conserved in copper amine oxidase. *Biochemistry* **45**, 4105–4120 (2006).
 107. Murakawa, T. *et al.* Probing the catalytic mechanism of copper amine oxidase from *Arthrobacter globiformis* with halide ions. *J. Biol. Chem.* **290**, 23094–23109 (2015).
 108. Murakawa, T. *et al.* Neutron crystallography of copper amine oxidase reveals keto/enolate interconversion of the quinone cofactor and unusual proton sharing. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 10818–10824 (2020).
 109. Murray, J. M. *et al.* The active site base controls cofactor reactivity in *Escherichia coli* amine oxidase: X-ray crystallographic studies with mutational variants. *Biochemistry* **38**, 8217–8227 (1999).
 110. Banco, M. T. *et al.* Neutron structures of the *Helicobacter pylori* 5?-methylthioadenosine nucleosidase highlight proton sharing and protonation states. *Proc. Natl. Acad. Sci.* **113**, 13756–13761 (2016).
 111. Shoji, M. *et al.* Unique protonation states of aspartate and topaquione in the active site of copper amine oxidase. *RSC Adv.* **10**, 38631–38639 (2020).
 112. Wilce, M. C. J. *et al.* Crystal Structures of the Copper-Containing Amine Oxidase from *Arthrobacter globiformis* in the Holo and Apo Forms: Implications for the Biogenesis of Topaquione. *Biochemistry* **36**, 16116–16133 (1997).
 113. Hasegawa, K., Ono, T. A. & Noguchi, T. Vibrational Spectra and Ab Initio DFT Calculations of 4-Methylimidazole and Its Different Protonation Forms: Infrared and Raman Markers of the Protonation State of a Histidine Side Chain. *J. Phys. Chem. B* **104**, 4253–4265 (2000).
 114. Poulos, T. L. Thirty years of heme peroxidase structural biology. *Arch. Biochem. Biophys.* **500**, 3–12 (2010).
 115. Keilin, D. & Hartree, E. F. Purification of horse-radish peroxidase and comparison of its properties with those of catalase and methaemoglobin. *Biochem. J.* **49**, 88–106 (1951).
 116. George, P. Chemical nature of the secondary hydrogen peroxide compound formed by cytochrome-

- c peroxidase and horseradish peroxidase. *Nature* **169**, 612–3 (1952).
117. Dawson, J. H. Probing structure-function relations in heme-containing oxygenases and peroxidases. *Science* **240**, 433–9 (1988).
 118. Groves, J. T. & Boaz, N. C. Fishing for peroxidase protons. *Science*. **345**, 142–143 (2014).
 119. Groves, J. T. Enzymatic C-H bond activation: Using push to get pull. *Nat. Chem.* **6**, 89–91 (2014).
 120. Sligar, S. G. Glimpsing the Critical Intermediate in Cytochrome P450 Oxidations. *Science*. **330**, 924–925 (2010).
 121. Karlin, K. D. Model offers intermediate insight. *Nature* **463**, 168–169 (2010).
 122. Behan, R. K. & Green, M. T. On the status of ferryl protonation. *J. Inorg. Biochem.* **100**, 448–459 (2006).
 123. Terner, J. *et al.* Resonance Raman spectroscopy of oxoiron(IV) porphyrin π -cation radical and oxoiron(IV) hemes in peroxidase intermediates. *J. Inorg. Biochem.* **100**, 480–501 (2006).
 124. Hersleth, H. P., Ryde, U., Rydberg, P., Görbitz, C. H. & Andersson, K. K. Structures of the high-valent metal-ion haem-oxygen intermediates in peroxidases, oxygenases and catalases. *J. Inorg. Biochem.* **100**, 460–476 (2006).
 125. Meharena, Y. T., Doukov, T., Li, H., Soltis, S. M. & Poulos, T. L. Crystallographic and single-crystal spectral analysis of the peroxidase ferryl intermediate. *Biochemistry* **49**, 2984–2986 (2010).
 126. Moody, P. C. E. & Raven, E. L. The Nature and Reactivity of Ferryl Heme in Compounds i and II. *Acc. Chem. Res.* **51**, 427–435 (2018).
 127. Poulos, T. L. *et al.* The crystal structure of cytochrome c peroxidase. *J. Biol. Chem.* **255**, 575–580 (1980).
 128. Jones, D. K., Dalton, D. A., Rosell, F. I. & Raven, E. L. Class I heme peroxidases: Characterization of soybean ascorbate peroxidase. *Arch. Biochem. Biophys.* **360**, 173–178 (1998).
 129. Volkov, A. N., Nicholls, P. & Worrall, J. A. R. The complex of cytochrome c and cytochrome c peroxidase: The end of the road? *Biochim. Biophys. Acta - Bioenerg.* **1807**, 1482–1503 (2011).
 130. Sivaraja, M., Goodin, D. B., Smith, M. & Hoffman, B. M. Identification by ENDOR of Trp191 as the free-radical site in cytochrome c peroxidase compound ES. *Science* **245**, 738–40 (1989).
 131. Casadei, C. M. *et al.* Neutron cryo-crystallography captures the protonation state of ferryl heme in a peroxidase. *Science*. **345**, 193–197 (2014).
 132. Vidossich, P. *et al.* On the role of water in peroxidase catalysis: A theoretical investigation of HRP compound i formation. *J. Phys. Chem. B* **114**, 5161–5169 (2010).
 133. Shaik, S., Kumar, D., de Visser, S. P., Altun, A. & Thiel, W. Theoretical Perspective on the Structure and Mechanism of Cytochrome P450 Enzymes †. *Chem. Rev.* **105**, 2279–2328 (2005).
 134. Kwon, H. *et al.* Direct visualization of a Fe(IV)-OH intermediate in a heme enzyme. *Nat. Commun.* **7**, 1–6 (2016).
 135. Asada, K. Ascorbate peroxidase – a hydrogen peroxide-scavenging enzyme in plants. *Physiol. Plant.* **85**, 235–241 (1992).
 136. Yosca, T. H. *et al.* Iron(IV)hydroxide pK(A) and the role of thiolate ligation in C-H bond activation by cytochrome P450. *Science* **342**, 825–9 (2013).
 137. Ledray, A. P., Krest, C. M., Yosca, T. H., Mitra, K. & Green, M. T. Ascorbate Peroxidase Compound II Is an Iron(IV) Oxo Species. *J. Am. Chem. Soc.* **142**, 20419–20425 (2020).
 138. Kwon, H. *et al.* Visualizing the protons in a metalloenzyme electron proton transfer pathway. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 6484–6490 (2020).
 139. Stone, K. L., Behan, R. K. & Green, M. T. Resonance Raman spectroscopy of chloroperoxidase compound II provides direct evidence for the existence of an iron(IV)-hydroxide. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 12307–12310 (2006).
 140. Green, M. T. Application of Badger’s rule to heme and non-heme iron-oxygen bonds: An examination of ferryl protonation states. *J. Am. Chem. Soc.* **128**, 1902–1906 (2006).
 141. Kwon, H. *et al.* XFEL Crystal Structures of Peroxidase Compound II. *Angew. Chemie Int. Ed.* (2021). doi:10.1002/anie.202103010
 142. Efimov, I. *et al.* Proton delivery to ferryl heme in a heme peroxidase: Enzymatic use of the grotthuss

- mechanism. *J. Am. Chem. Soc.* **133**, 15376–15383 (2011).
143. Wilttschi, B. *et al.* Enzymes revolutionize the bioproduction of value-added compounds: From enzyme discovery to special applications. *Biotechnol. Adv.* **40**, 107520 (2020).
 144. Liao, J. C., Mi, L., Pontrelli, S. & Luo, S. Fuelling the future: Microbial engineering for the production of sustainable biofuels. *Nat. Rev. Microbiol.* **14**, 288–304 (2016).
 145. Lynd, L. R., Weimer, P. J., Zyl, W. H. Van & Isak, S. Microbial Cellulose Utilization : Fundamentals and Biotechnology Microbial Cellulose Utilization : Fundamentals and Biotechnology Downloaded from <http://mmbr.asm.org/> on February 6 , 2013 by INDIAN INSTITUTE OF TECHNOLOGY MADRAS. *Microbiol. Mol. Biol. Rev.* **66**, 506–577 (2002).
 146. Harris, P. V., Xu, F., Kreel, N. E., Kang, C. & Fukuyama, S. New enzyme insights drive advances in commercial ethanol production. *Curr. Opin. Chem. Biol.* **19**, 162–170 (2014).
 147. Cragg, S. M. *et al.* Lignocellulose degradation mechanisms across the Tree of Life. *Curr. Opin. Chem. Biol.* **29**, 108–119 (2015).
 148. Himmel, M. E. *et al.* Biomass recalcitrance: Engineering plants and enzymes for biofuels production. *Science*. **315**, 804–807 (2007).
 149. Bissaro, B., Várnai, A., Røhr, Å. K. & Eijsink, V. G. H. Oxidoreductases and Reactive Oxygen Species in Conversion of Lignocellulosic Biomass. *Microbiol. Mol. Biol. Rev.* **82**, 1–51 (2018).
 150. Tandrup, T., Frandsen, K. E. H., Johansen, K. S., Berrin, J.-G. & Lo Leggio, L. Recent insights into lytic polysaccharide monooxygenases (LPMOs). *Biochem. Soc. Trans.* **46**, 1431–1447 (2018).
 151. Quinlan, R. J. *et al.* 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 (2011).
 152. Frommhagen, M., Westphal, A. H., van Berkel, W. J. H. & Kabel, M. A. Distinct Substrate Specificities and Electron-Donating Systems of Fungal Lytic Polysaccharide Monooxygenases. *Front. Microbiol.* **9**, 1–22 (2018).
 153. Kjaergaard, C. H. *et al.* Spectroscopic and computational insight into the activation of O₂ by the mononuclear Cu center in polysaccharide monooxygenases. *Proc. Natl. Acad. Sci.* **111**, 8797–8802 (2014).
 154. Walton, P. H. & Davies, G. J. On the catalytic mechanisms of lytic polysaccharide monooxygenases. *Curr. Opin. Chem. Biol.* **31**, 195–207 (2016).
 155. Meier, K. K. *et al.* Oxygen Activation by Cu LPMOs in Recalcitrant Carbohydrate Polysaccharide Conversion to Monomer Sugars. *Chem. Rev.* **118**, 2593–2635 (2018).
 156. Phillips, C. M., Beeson, W. T., Cate, J. H. & Marletta, M. A. Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem. Biol.* **6**, 1399–1406 (2011).
 157. Beeson, W. T., Phillips, C. M., Cate, J. H. D. & Marletta, M. A. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J. Am. Chem. Soc.* **134**, 890–892 (2012).
 158. Li, X., Beeson, W. T., Phillips, C. M., Marletta, M. A. & Cate, J. H. D. Structural Basis for Substrate Targeting and Catalysis by Fungal Polysaccharide Monooxygenases. *Structure* **20**, 1051–1061 (2012).
 159. Neisen, B. D., Gagnon, N. L., Dhar, D., Spaeth, A. D. & Tolman, W. B. Formally Copper(III)-Alkylperoxo Complexes as Models of Possible Intermediates in Monooxygenase Enzymes. *J. Am. Chem. Soc.* **139**, 10220–10223 (2017).
 160. Dhar, D. & Tolman, W. B. Hydrogen atom abstraction from hydrocarbons by a copper(III)-hydroxide complex. *J. Am. Chem. Soc.* **137**, 1322–1329 (2015).
 161. Kim, S., Ståhlberg, J., Sandgren, M., Paton, R. S. & Beckham, G. T. Quantum mechanical calculations suggest that lytic polysaccharide monooxygenases use a copper-oxyl, oxygen-rebound mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 149–154 (2014).
 162. Lee, J. Y. & Karlin, K. D. Elaboration of copper-oxygen mediated CH activation chemistry in consideration of future fuel and feedstock generation. *Curr. Opin. Chem. Biol.* **25**, 184–193 (2015).
 163. Wang, B. *et al.* QM/MM Studies into the H₂O₂-Dependent Activity of Lytic Polysaccharide Monooxygenases: Evidence for the Formation of a Caged Hydroxyl Radical Intermediate. *ACS*

- Catal.* **8**, 1346–1351 (2018).
164. Wang, B., Wang, Z., Davies, G. J., Walton, P. H. & Rovira, C. Activation of O₂ and H₂O₂ by Lytic Polysaccharide Monooxygenases. *ACS Catal.* **10**, 12760–12769 (2020).
 165. Bissaro, B. *et al.* Oxidative cleavage of polysaccharides by monocopper enzymes depends on H₂O₂. *Nat. Chem. Biol.* **13**, 1123–1128 (2017).
 166. Hangasky, J. A., Iavarone, A. T. & Marletta, M. A. Reactivity of O₂ versus H₂O₂ with polysaccharide monooxygenases. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 4915–4920 (2018).
 167. Bacik, J. P. *et al.* Neutron and high-resolution room-temperature X-ray data collection from crystallized lytic polysaccharide monooxygenase. *Acta Crystallogr. Sect. Struct. Biol. Commun.* **71**, 1448–1452 (2015).
 168. Bacik, J. P. *et al.* Neutron and Atomic Resolution X-ray Structures of a Lytic Polysaccharide Monooxygenase Reveal Copper-Mediated Dioxygen Binding and Evidence for N-Terminal Deprotonation. *Biochemistry* **56**, 2529–2532 (2017).
 169. Caldararu, O., Oksanen, E., Ryde, U. & Hedegård, E. D. Mechanism of hydrogen peroxide formation by lytic polysaccharide monooxygenase. *Chem. Sci.* **10**, 576–586 (2019).
 170. O'Dell, W. B., Agarwal, P. K. & Meilleur, F. Oxygen Activation at the Active Site of a Fungal Lytic Polysaccharide Monooxygenase. *Angew. Chemie - Int. Ed.* **56**, 767–770 (2017).
 171. Schüle, M. Enzymatic properties of cellulases from *Humicola insolens*. *J. Biotechnol.* **57**, 71–81 (1997).
 172. Boer, H. & Koivula, A. The relationship between thermal stability and pH optimum studied with wild-type and mutant *Trichoderma reesei* cellobiohydrolase Cel7A. *Eur. J. Biochem.* **270**, 841–848 (2003).
 173. Span, E. A., Suess, D. L. M., Deller, M. C., Britt, R. D. & Marletta, M. A. The Role of the Secondary Coordination Sphere in a Fungal Polysaccharide Monooxygenase. *ACS Chem. Biol.* **12**, 1095–1103 (2017).
 174. Hedegård, E. D. & Ryde, U. Molecular mechanism of lytic polysaccharide monooxygenases. *Chem. Sci.* **9**, 3866–3880 (2018).
 175. Frandsen, K. E. H. *et al.* The molecular basis of polysaccharide cleavage by lytic polysaccharide monooxygenases. *Nat. Chem. Biol.* **12**, 298–303 (2016).
 176. Simmons, T. J. *et al.* Structural and electronic determinants of lytic polysaccharide monooxygenase reactivity on polysaccharide substrates. *Nat. Commun.* **8**, (2017).
 177. Tandrup, T. *et al.* Oligosaccharide Binding and Thermostability of Two Related AA9 Lytic Polysaccharide Monooxygenases. *Biochemistry* **59**, 3347–3358 (2020).
 178. Schröder, G. C., O'Dell, W. B., Swartz, P. D. & Meilleur, F. Preliminary results of neutron and X-ray diffraction data collection on a lytic polysaccharide monooxygenase under reduced and acidic conditions. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **77**, 128–133 (2021).
 179. Auten, R. L. & Davis, J. M. Oxygen toxicity and reactive oxygen species: The devil is in the details. *Pediatr. Res.* **66**, 121–127 (2009).
 180. Fridavich, I. Superoxide radical and superoxide dismutases. *Annu. Rev. Biochem.* **64**, 97–112 (1995).
 181. Abreu, I. A. & Cabelli, D. E. Superoxide dismutases—a review of the metal-associated mechanistic variations. *Biochim. Biophys. Acta - Proteins Proteomics* **1804**, 263–274 (2010).
 182. Jastroch, M., Divakaruni, A. S., Mookerjee, S., Treberg, J. R. & Brand, M. D. Mitochondrial proton and electron leaks. *Essays Biochem.* **47**, 53–67 (2010).
 183. Borgstahl, G. E. O. *et al.* The structure of human mitochondrial manganese superoxide dismutase reveals a novel tetrameric interface of two 4-helix bundles. *Cell* **71**, 107–118 (1992).
 184. Miller, A. F., Padmakumar, K., Sorkin, D. L., Karapetian, A. & Vance, C. K. Proton-coupled electron transfer in Fe-superoxide dismutase and Mn-superoxide dismutase. *J. Inorg. Biochem.* **93**, 71–83 (2003).
 185. Chang, C. J., Chang, M. C. Y., Damrauer, N. H. & Nocera, D. G. Proton-coupled electron transfer: A unifying mechanism for biological charge transport, amino acid radical initiation and propagation,

- and bond making/breaking reactions of water and oxygen. *Biochim. Biophys. Acta - Bioenerg.* **1655**, 13–28 (2004).
186. Guan, Y. *et al.* Crystal structure of Y34F mutant human mitochondrial manganese superoxide dismutase and the functional role of tyrosine 34. *Biochemistry* **37**, 4722–4730 (1998).
 187. Srnec, M., Aquilante, F., Ryde, U. & Rulišek, L. Reaction mechanism of manganese superoxide dismutase studied by combined quantum and molecular mechanical calculations and multiconfigurational methods. *J. Phys. Chem. B* **113**, 6074–6086 (2009).
 188. Heimdal, J., Kaukonen, M., Srnec, M., Rulišek, L. & Ryde, U. Reduction potentials and acidity constants of Mn superoxide dismutase calculated by QM/MM free-energy methods. *ChemPhysChem* **12**, 3337–3347 (2011).
 189. Hsieh, Y. *et al.* Probing the active site of human manganese superoxide dismutase: The role of glutamine 143. *Biochemistry* **37**, 4731–4739 (1998).
 190. Hearn, A. S. *et al.* Catalytic and structural effects of amino acid substitution at histidine 30 in human manganese superoxide dismutase: Insertion of valine C γ into the substrate access channel. *Biochemistry* **42**, 2781–2789 (2003).
 191. Perry, J. J. P. *et al.* Contribution of human manganese superoxide dismutase tyrosine 34 to structure and catalysis. *Biochemistry* **48**, 3417–3424 (2009).
 192. Azadmanesh, J., Lutz, W. E., Coates, L., Weiss, K. L. & Borgstahl, G. E. O. Direct detection of coupled proton and electron transfers in human manganese superoxide dismutase. *Nat. Commun.* **12**, 2079 (2021).
 193. Li, J., Fisher, C. L., Konecny, R., Bashford, D. & Noodleman, L. Density functional and electrostatic calculations of manganese superoxide dismutase active site complexes in protein environments. *Inorg. Chem.* **38**, 929–939 (1999).
 194. Han, W. G., Lovell, T. & Noodleman, L. Coupled redox potentials in manganese and iron superoxide dismutases from reaction kinetics and density functional/electrostatics calculations. *Inorg. Chem.* **41**, 205–218 (2002).
 195. Rulišek, L. & Ryde, U. Structure of reduced and oxidized manganese superoxide dismutase: A combined computational and experimental approach. *J. Phys. Chem. B* **110**, 11511–11518 (2006).
 196. Kaukonen, M., Söderhjelm, P., Heimdal, J. & Ryde, U. Proton transfer at metal sites in proteins studied by quantum mechanical free-energy perturbations. *J. Chem. Theory Comput.* **4**, 985–1001 (2008).
 197. Eriksson, M. A., Härd, T. & Nilsson, L. On the pH dependence of amide proton exchange rates in proteins. *Biophys. J.* **69**, 329–339 (1995).
 198. Borgstahl, G. E. O., Pokross, M., Chehab, R., Sekher, A. & Snell, E. H. Cryo-trapping the six-coordinate, distorted-octahedral active site of manganese superoxide dismutase. *J. Mol. Biol.* **296**, 951–959 (2000).
 199. Lyubimov, A. Y., Lario, P. I., Moustafa, I. & Vrielink, A. Atomic resolution crystallography reveals how changes in pH shape the protein microenvironment. *Nat. Chem. Biol.* **2**, 259–264 (2006).
 200. Rutherford, A. W. Photosystem II, the water-splitting enzyme. *Trends Biochem. Sci.* **14**, 227–232 (1989).
 201. Ettwig, K. F. *et al.* Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. *Nature* **464**, 543–548 (2010).
 202. Zhu, B. *et al.* Nitric oxide dismutase (NOD) genes as a functional marker for the diversity and phylogeny of methane-driven oxygenic denitrifiers. *Front. Microbiol.* **10**, 1–9 (2019).
 203. Lee, A. Q., Streit, B. R., Zdilla, M. J., Abu-Omar, M. M. & DuBois, J. L. Mechanism of and exquisite selectivity for O-O bond formation by the heme-dependent chlorite dismutase. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 15654–15659 (2008).
 204. Hofbauer, S., Schaffner, I., Furtmüller, P. G. & Obinger, C. Chlorite dismutases - a heme enzyme family for use in bioremediation and generation of molecular oxygen. *Biotechnol. J.* **9**, 461–473 (2014).
 205. Kumar, S. *et al.* Oxidative stress induced by chlorine dioxide as an insecticidal factor to the Indian

- meal moth, *Plodia interpunctella*. *Pestic. Biochem. Physiol.* **124**, 48–59 (2015).
206. Kostan, J. *et al.* Structural and functional characterisation of the chlorite dismutase from the nitrite-oxidizing bacterium ‘*Candidatus Nitrospira defluvii*’: Identification of a catalytically important amino acid residue. *J. Struct. Biol.* **172**, 331–342 (2010).
207. Mlynek, G. *et al.* Unexpected diversity of chlorite dismutases: A catalytically efficient dimeric enzyme from *Nitrobacter winogradskyi*. *J. Bacteriol.* **193**, 2408–2417 (2011).
208. Celis, A. I. *et al.* A dimeric chlorite dismutase exhibits O₂-generating activity and acts as a chlorite antioxidant in *Klebsiella pneumoniae* MGH 78578. *Biochemistry* **54**, 434–446 (2015).
209. Schaffner, I. *et al.* Mechanism of chlorite degradation to chloride and dioxygen by the enzyme chlorite dismutase. *Arch. Biochem. Biophys.* **574**, 18–26 (2015).
210. Keith, J. M., Abu-Omar, M. M. & Hall, M. B. Computational investigation of the concerted dismutation of chlorite ion by water-soluble iron porphyrins. *Inorg. Chem.* **50**, 7928–7930 (2011).
211. Mayfield, J. A., Blanc, B., Rodgers, K. R., Lukat-Rodgers, G. S. & DuBois, J. L. Peroxidase-Type Reactions Suggest a Heterolytic/Nucleophilic O–O Joining Mechanism in the Heme-Dependent Chlorite Dismutase. *Biochemistry* **52**, 6982–6994 (2013).
212. Goblirsch, B. R., Streit, B. R., DuBois, J. L. & Wilmot, C. M. Structural features promoting dioxygen production by *Dechloromonas aromatica* chlorite dismutase. *J. Biol. Inorg. Chem.* **15**, 879–888 (2010).
213. Hofbauer, S. *et al.* Transiently produced hypochlorite is responsible for the irreversible inhibition of chlorite dismutase. *Biochemistry* **53**, 3145–3157 (2014).
214. Streit, B. R., Blanc, B., Lukat-Rodgers, G. S., Rodgers, K. R. & Dubois, J. L. How active-site protonation state influences the reactivity and ligation of the heme in chlorite dismutase. *J. Am. Chem. Soc.* **132**, 5711–5724 (2010).
215. Blanc, B. *et al.* Understanding how the distal environment directs reactivity in chlorite dismutase: Spectroscopy and reactivity of Arg183 Mutants. *Biochemistry* **51**, 1895–1910 (2012).
216. Schaffner, I. *et al.* Molecular Mechanism of Enzymatic Chlorite Detoxification: Insights from Structural and Kinetic Studies. *ACS Catal.* **7**, 7962–7976 (2017).
217. Harms, M. J., Schlessman, J. L., Sue, G. R. & Bertrand García-Moreno, E. Arginine residues at internal positions in a protein are always charged. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 18954–18959 (2011).
218. Fitch, C. A., Platzer, G., Okon, M., Garcia-Moreno, B. E. & McIntosh, L. P. Arginine: Its pK_a value revisited. *Protein Sci.* **24**, 752–761 (2015).
219. Schmidt, D. *et al.* Arresting the Catalytic Arginine in Chlorite Dismutases: Impact on Heme Coordination, Thermal Stability, and Catalysis. *Biochemistry* (2021). doi:10.1021/acs.biochem.0c00910
220. De Rienzo, F., Gabdoulline, R. R., Menziani, M. C. & Wade, R. C. Blue copper proteins: A comparative analysis of their molecular interaction properties. *Protein Sci.* **9**, 1439–1454 (2000).
221. Sykes, A. G. Active-Site Properties Of The Blue Copper Proteins. *Adv. Inorg. Chem.* **36**, 377–408 (1991).
222. Choi, M. & Davidson, V. L. Cupredoxins - A study of how proteins may evolve to use metals for bioenergetic processes. *Metallomics* **3**, 140–151 (2011).
223. Williamson, H. R., Dow, B. A. & Davidson, V. L. Mechanisms for control of biological electron transfer reactions. *Bioorg. Chem.* **57**, 213–221 (2014).
224. Davidson, V. L. & Jones, L. H. Electron transfer from copper to heme within the methylamine dehydrogenase-amicyanin-cytochrome c-551i complex. *Biochemistry* **35**, 8120–8125 (1996).
225. Sukumar, N., Mathews, F. S., Langan, P. & Davidson, V. L. A joint x-ray and neutron study on amicyanin reveals the role of protein dynamics in electron transfer. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 6817–6822 (2010).
226. Durley, R., Chen, L., Scott Mathews, F., Lim, L. W. & Davidson, V. L. Crystal structure analysis of amicyanin and apoamicyanin from *paracoccus denitrificans* at 2.0 Å and 1.8 Å resolution. *Protein Sci.* **2**, 739–752 (1993).

227. Chen, L., Durley, R., Mathews, F. & Davidson, V. Structure of an electron transfer complex: methylamine dehydrogenase, amicyanin, and cytochrome c551i. *Science*. **264**, 86–90 (1994).
228. Chen, L. *et al.* Crystal Structure of an Electron-Transfer Complex between Methylamine Dehydrogenase and Amicyanin. *Biochemistry* **31**, 4959–4964 (1992).
229. Dow, B. A. *et al.* The sole tryptophan of amicyanin enhances its thermal stability but does not influence the electronic properties of the type 1 copper site. *Arch. Biochem. Biophys.* **550–551**, 20–27 (2014).
230. Cave, R. J. & Newton, M. D. Calculation of electronic coupling matrix elements for ground and excited state electron transfer reactions: Comparison of the generalized Mulliken-Hush and block diagonalization methods. *J. Chem. Phys.* **106**, 9213–9226 (1997).
231. Ferrari, D. *et al.* Catalysis and electron transfer in protein crystals: The binary and ternary complexes of methylamine dehydrogenase with electron acceptors. *Biochim. Biophys. Acta - Proteins Proteomics* **1647**, 337–342 (2003).
232. Zhu, Z. *et al.* Molecular basis for interprotein complex-dependent effects on the redox properties of amicyanin. *Biochemistry* **37**, 17128–17136 (1998).
233. Iwaoka, M. *Chalcogen bonds in protein architecture. Noncovalent Forces* (2015). doi:10.1007/978-3-319-14163-3_9
234. Lennicke, C., Rahn, J., Lichtenfels, R., Wessjohann, L. A. & Seliger, B. Hydrogen peroxide - Production, fate and role in redox signaling of tumor cells. *Cell Commun. Signal.* **13**, 1–19 (2015).
235. Gough, D. R. & Cotter, T. G. Hydrogen peroxide: A Jekyll and Hyde signalling molecule. *Cell Death Dis.* **2**, 1–8 (2011).
236. Fang, F. C. Antimicrobial reactive oxygen and nitrogen species: Concepts and controversies. *Nat. Rev. Microbiol.* **2**, 820–832 (2004).
237. Glorieux, C. & Calderon, P. B. Catalase, a remarkable enzyme: Targeting the oldest antioxidant enzyme to find a new cancer treatment approach. *Biol. Chem.* **398**, 1095–1108 (2017).
238. Chelikani, P., Fita, I. & Loewen, P. C. Diversity of structures and properties among catalases. *Cell. Mol. Life Sci.* **61**, 192–208 (2004).
239. Zámocký, M. & Koller, F. Understanding the structure and function of catalases: Clues from molecular evolution and in vitro mutagenesis. *Prog. Biophys. Mol. Biol.* **72**, 19–66 (1999).
240. Crichton, R. Manganese – Oxygen Generation and Detoxification. in *Biological Inorganic Chemistry* 459–473 (Elsevier, 2019). doi:10.1016/B978-0-12-811741-5.00016-3
241. Antonyuk, S. V. *et al.* Structures of macromolecular compounds: Three-dimensional structure of the enzyme dimanganese catalase from *Thermus thermophilus* at 1 Å resolution. *Kristallografiya* **45**, 111–122 (2000).
242. Barynin, V. V. *et al.* Crystal structure of manganese catalase from *Lactobacillus plantarum*. *Structure* **9**, 725–738 (2001).
243. Whittaker, J. W. Non-heme manganese catalase - The ‘other’ catalase. *Arch. Biochem. Biophys.* **525**, 111–120 (2012).
244. Kono, Y. & Fridovich, I. Inhibition and reactivation of Mn-catalase. Implications for valence changes at the active site manganese. *J. Biol. Chem.* **258**, 13646–13648 (1983).
245. Waldo, G. S., Fronko, R. M. & Penner-Hahn, J. E. Inactivation and Reactivation of Manganese Catalase: Oxidation-State Assignments Using X-ray Absorption Spectroscopy. *Biochemistry* **30**, 10486–10490 (1991).
246. Waldo, G. S. & Penner-Hahn, J. E. Mechanism of Manganese Catalase Peroxide Disproportionation: Determination of Manganese Oxidation States during Turnover. *Biochemistry* **34**, 1507–1512 (1995).
247. Boelrijk, A. E. M. & Dismukes, G. C. Mechanism of hydrogen peroxide dismutation by a dimanganese catalase mimic: Dominant role of an intramolecular base on substrate binding affinity and rate acceleration. *Inorg. Chem.* **39**, 3020–3028 (2000).
248. Yamada, T., Yano, N., Hosoya, T. & Kusaka, K. Single-crystal time-of-flight neutron Laue methods: application to manganese catalase from *Thermus thermophilus* HB27. *J. Appl.*

- Crystallogr.* **52**, 972–983 (2019).
249. Stephanos, J. J. & Addison, A. W. *Chemistry of Metalloproteins: Problems and Solutions in Bioinorganic Chemistry*. (John Wiley & Sons, Incorporated, 2014).
250. Antonyuk, S. V., Han, C., Eady, R. R. & Hasnain, S. S. Structures of protein-protein complexes involved in electron transfer. *Nature* **496**, 123–126 (2013).
251. Liou, S. H., Mahomed, M., Lee, Y. T. & Goodin, D. B. Effector Roles of Putidaredoxin on Cytochrome P450cam Conformational States. *J. Am. Chem. Soc.* **138**, 10163–10172 (2016).
252. Tavolieri, A. M. *et al.* NADPH-dependent sulfite reductase flavoprotein adopts an extended conformation unique to this diflavin reductase. *J. Struct. Biol.* **205**, 170–179 (2019).
253. Freeman, S. L. *et al.* Solution structure of the cytochrome P450 reductase–cytochrome c complex determined by neutron scattering. *J. Biol. Chem.* **293**, 5210–5219 (2018).
254. Bodenheimer, A. M. *et al.* Structural investigation of cellobiose dehydrogenase IIA: Insights from small angle scattering into intra- and intermolecular electron transfer mechanisms. *Biochim. Biophys. Acta - Gen. Subj.* **1862**, 1031–1039 (2018).

1.2 IMAGINE: neutrons reveal enzyme chemistry

The following work was reprinted with permission from: Schröder, G. C., O'Dell, W. B., Myles, D. A. A., Kovalevsky, A. & Meilleur, F. IMAGINE: Neutrons reveal enzyme chemistry. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 778–786 (2018).

1.2.1 Synopsis

Capabilities of the IMAGINE neutron protein diffractometer at the Oak Ridge National Laboratory High Flux Isotope Reactor and highlights of the first 5 years of the scientific program are reviewed.

1.2.2 Abstract

Neutron diffraction is exquisitely sensitive to the position of hydrogen atoms in protein crystal structures. IMAGINE is a high intensity, quasi-Laue neutron crystallography beam line developed at the High Flux Isotope Reactor (HFIR) at Oak Ridge National Laboratory (ORNL). This state-of-the-art facility for neutron-diffraction has enabled detailed structural analysis of macromolecules. IMAGINE is especially suited to resolve individual hydrogen atoms in protein structures, enabling neutron protein structures to be determined at or near atomic resolutions from crystals with volumes of less than 1 mm³ and unit cell edge of less than 150 Å. Beam line features include elliptical focusing mirrors that deliver neutrons into a 2.0 × 3.2 mm² focal spot at the sample position, and variable short and long wavelength cutoff optics that provide automated exchange between multiple wavelength configurations. This review gives an overview of the IMAGINE beam line at the HFIR, presents examples of the scientific questions being addressed at this beamline, and highlights important findings in enzyme chemistry that have been made using the neutron diffraction capabilities offered by IMAGINE.

1.2.3 Keywords

Neutron crystallography, enzyme mechanism, deuteron, proton.

1.2.4 Neutron protein crystallography

Hydrogen atoms play a central role in most enzyme catalyzed reactions. A detailed knowledge of the hydrogen atom positions can therefore help provide critical understanding of the chemistry involved. Structural techniques such as Nuclear Magnetic Resonance (NMR) and X-ray crystallography allow

enzyme structures to be readily determined at the atomic level, but provide limited direct information on the localization of hydrogen atoms in the structures. NMR provides detailed information on enzyme structure and dynamics, and can potentially reveal the protonation state of titratable residues, but is limited in practice to proteins with molecular weights around 40kD. In contrast, X-ray analysis can provide atomic resolution structures with effectively no upper limit on the size of protein to be studied, and has enabled more than 120,000 protein crystal structures to be determined (<http://www.rcsb.org/pdb/>]. Many of these structures lack valuable information on the position and orientation of hydrogen atoms because hydrogen atoms, with only one electron, scatter X-rays only weakly (Table 1.2). It is therefore difficult - and in most cases impossible - to visualize hydrogen atoms in X-ray crystal structures due to hydrogen atoms mobility, polarizability and radiation damage of chemical groups. Information on hydrogen atom position is therefore only accessible under the most ideal circumstances using X-ray diffraction techniques.

Table 1.2 Neutron scattering lengths and X-ray scattering factors for atoms in biological macromolecules. X-ray scattering factors decrease with increasing scattering angles while neutron scattering lengths show no angular dependence.

	H	D	C	N	O
Neutron coherent scattering length (10^{-12} cm)	-0.374	0.667	0.665	0.936	0.581
Neutron incoherent cross section (Barns)	80.27	2.05	0.00	0.49	0.00
X-ray Scattering Factors (10^{-12} cm) $\sin\theta/\lambda=0$	0.28	0.28	1.69	1.97	2.25
X-ray Scattering Factors (10^{-12} cm) $\sin\theta/\lambda=0.5\text{\AA}^{-1}$	0.02	0.02	0.48	0.53	0.62

Hydrogen atoms can be more readily located by neutron diffraction analysis. This is because the coherent scattering length of hydrogen and deuterium are similar in magnitude to that of carbon, nitrogen and oxygen (Table 1.2) ¹. The nuclear diffraction signal from both hydrogen and deuterium are therefore strong enough that they are clearly visible as peaks in nuclear scattering length density (SLD) maps at typical resolutions of 1.8 to 2.2Å. However, although of comparable magnitude to heavier atoms, the scattering from hydrogen is of the opposite sign and gives rise to negative peaks in Fourier maps, which

can complicate analysis. In contrast, the neutron scattering length of deuterium is of the same sign and approximate magnitude as that of the heavier atoms and thus gives a strong positive peak in the SLD maps. Neutron diffraction is therefore the method of choice for visualizing the positions of hydrogen or deuterium atoms.

A disadvantage of neutron diffraction is that the signal-to-noise ratio from hydrogen-containing samples is weak. This is due to the large incoherent neutron scattering cross section of hydrogen (80.27 Barns, Table 1.2), which contributes to a significant background, making the detection of the signal difficult. In contrast, the incoherent scattering cross section of deuterium (2.05 Barns) is much smaller, reducing the background and thereby increasing the signal-to-noise ratio of the data. For this reason, hydrogenated protein solutions or fully-grown crystals used for neutron diffraction studies are subjected to H₂O/D₂O solvent exchange, resulting in the substitution of all titratable and solvent accessible protons with deuterium atoms. This H/D exchange reduces the background noise from crystals by a factor of ~3. When possible, replacing all protein hydrogen atoms with deuterium to obtain a perdeuterated form through protein expression in deuterated media has a much larger impact, improving the signal-to-noise ratio of crystallographic data by a factor of ~10^{2,3}.

Interested readers are referred to comprehensive reviews for historical, technical and scientific perspectives on neutron protein crystallography for further reading^{2,4-12}.

1.2.5 Overview of the IMAGINE beamline

To compensate the weak scattering inherent to biological samples and the low flux of available neutron beams, almost all reactor-based protein diffractometers, and all accelerator-based diffractometers, operate in broad band-pass Laue or quasi-Laue geometries, maximizing the flux delivered to the sample. The IMAGINE diffractometer at HFIR is designed for the collection of high-resolution (~1.5 Å) quasi-Laue data from single crystals (>0.3 mm³) with moderately large unit cell sizes (up to ~150 Å). The instrument is located on the end-station of cold guide CG-4, which receives neutrons from a 19×12 mm²

section of the CG-4 guide and delivers a flux of 2×10^9 n/s/cm² over the useful 2–10 Å wavelength range (Figure 1.10). The neutron wavelengths and bandpass delivered to the sample are varied by automated exchange of 3 neutron flat mirrors and 3 pairs of neutron filters that provide short wavelength ($\lambda_{min} = 2.0$, 2.8 and 3.3 Å) and long wavelength ($\lambda_{max} = 3.0$, 4.0 and 4.5 Å) cut-offs, respectively. It is also possible to leave the filter position open for full Laue mode analysis, which enables crystals to be screened in minutes. A pair of elliptically shaped mirrors then collects and focuses the resulting beam vertically and horizontally down to 2.0×3.2 mm² at the sample position with full-width vertical and horizontal divergence of 0.5° and 0.6°, respectively. The 2.8–4.0 Å bandpass quasi-Laue configuration delivers $\sim 3 \times 10^7$ n/s/cm² to sample¹³. The IMAGINE neutron spectrum at the sample position is presented in Figure 1.11.

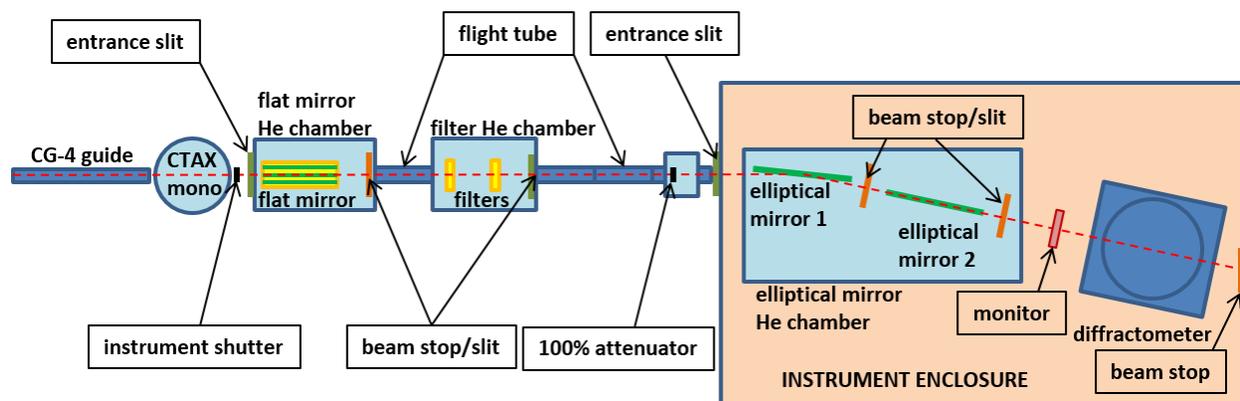


Figure 1.10 IMAGINE beamline layout. Neutron optics for IMAGINE begin immediately downstream of the monochromator for the CTAX instrument.

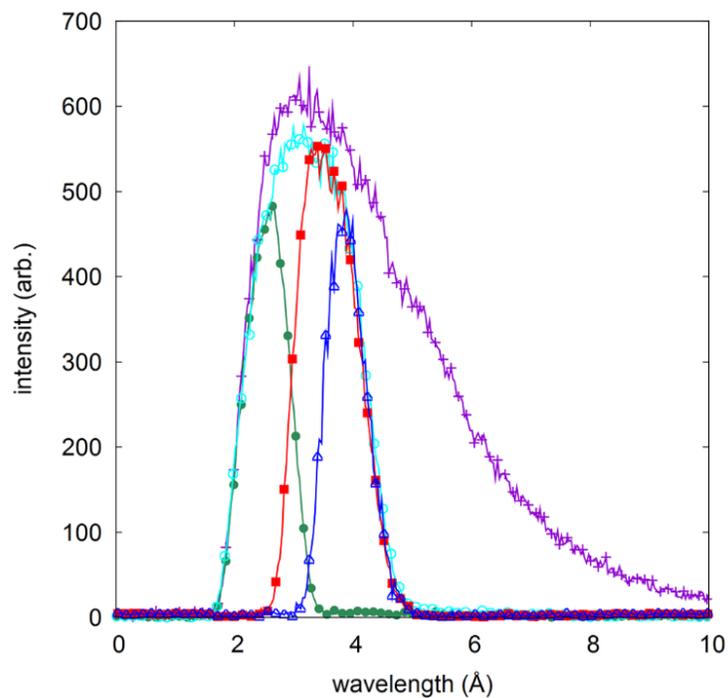


Figure 1.11 IMAGINE neutron spectra at the sample position. (+) 2.0–10.0 Å Laue configuration, (●) 2.0–3.0 Å bandpass, (○) 2.0–4.5 Å bandpass, (■) 2.8–4.5 Å bandpass and (△) 3.3–4.5 Å bandpass.

The diffractometer uses a cylindrical neutron image plate (NIP) detector geometry that is available commercially from ARINAX, France (Figure 1.12). The detector system has a broad dynamic range (>10,000), high spatial resolution (250 μm), and high detective quantum efficiency (DQE) of 47% for the standard 2.8–4.5 Å quasi-Laue configuration.

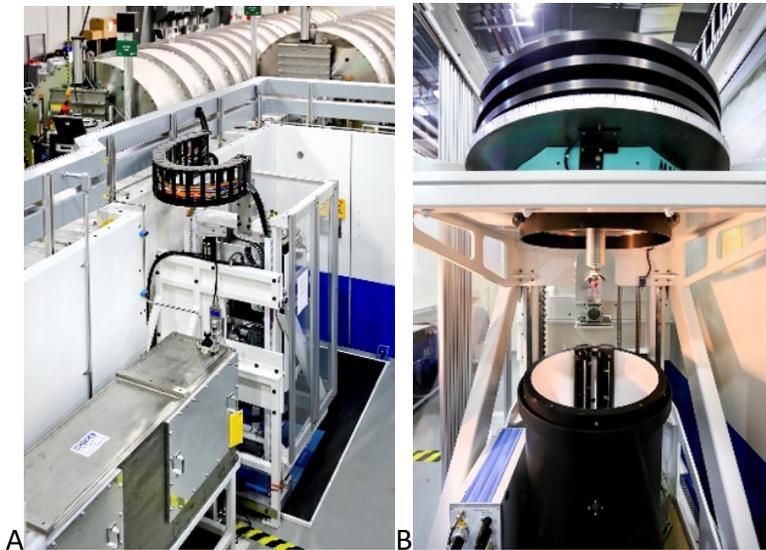


Figure 1.12 IMAGINE installed on CG-4D. A. IMAGINE is installed in a 4”-thick steel wall enclosure next to the Bio-SANS instrument. The elliptical mirror chamber sits directly upstream of the instrument. B. The image plate detector is mounted inside a cylindrical drum with the radius of 200 mm. The erasing and reading system is placed in the backscattering position. Crystals are mounted on a sample stick with automated ϕ -axis rotation for room temperature data collection. (Photo credit: Jason Richards, ORNL)

1.2.6 Scientific highlights

A comprehensive list of neutron and joint x-ray/neutron structures of protein deposited in the Protein Data Bank from the IMAGINE beamline is presented in Table 1.3. A total of 13 neutron structures have been deposited since the beamline first came online in 2013. In this section, we highlight 6 impactful case studies from data obtained at IMAGINE.

Table 1.3 A comprehensive list of neutron and joint X-ray/neutron structures of proteins deposited in the Protein Database from the IMAGINE beamline. “Uniform” deuterium labeling indicates that proteins were expressed from organisms grown in the culture media prepared from deuterated carbon source and D₂O. The label “85% uniform” deuterium labeling is used for HIV-1 protease which was expressed from *E. coli* grown in culture media prepared with hydrogenated carbon source and D₂O. “None” indicates that the proteins were hydrogenated. Proteins were expressed in *E. coli* except for lytic polysaccharide monooxygenase which was expressed in *P. pastoris*. Blank entries are given for X-ray d_{\min} when refinement was performed against neutron data only.

Protein	PDB	Deuterium Labeling	Crystallization Method	H/D Exchange Method	Crystal Volume (mm ³)	Measurement Time (days)	Neutron d_{\min} (Å)	X-ray d_{\min} (Å)	Ref.
Rubredoxin	4K9F	uniform	macroseeded sitting drop	crystallization	0.7	3	1.75	--	(Meilleur <i>et al.</i> , 2013)
Dihydrofolate reductase	4PDJ	none	macroseeded sitting drop	vapor 14 days	3.6	17	2	1.6	(Wan <i>et al.</i> , 2014a; Wan <i>et al.</i> , 2014b)
Cholesterol oxidase	5KWF	none	macroseeded sitting drop	soak 24 hours vapor 24 months	0.41	20	2.21	1.5	(Golden <i>et al.</i> , 2017)
Lytic polysaccharide monooxygenase	5TKI	none	macroseeded sitting drop	vapor 4 weeks	0.35	22	2.12	1.5	(O'Dell <i>et al.</i> , 2017a)
HIV-1Protease	5E5K	85% uniform	sitting drop	vapor 4 weeks	0.3	20	2.3	1.75	(Gerlits <i>et al.</i> , 2016)
Carbonic anhydrase II	5C8I	none	sitting drop	vapor 4 weeks	0.7	22	2.2	1.56	(Aggarwal <i>et al.</i> , 2016)

Table 1.3 (continued).

T4 Lysozyme (80K)	5VNQ	uniform	microseeded sitting drop	crystallization	0.7	11	2.2	--	(Li <i>et al.</i> , 2017)
T4 lysozyme (mutant)	5XPE	uniform	macroseeded hanging drop	crystallization	0.9	8	2.1	1.65	(Hiromoto <i>et al.</i> , 2017)
Concanavalin A	5WEY	none	sitting drop	vapor 1 week	1	15	2.5	1.8	(Gerlits <i>et al.</i> , 2017)
H-RAS	4RSG	none	sitting drop	vapor 8 weeks soak 1 week	0.7	23	1.91	--	(Knihtila <i>et al.</i> , 2015)
5'-methylthioadenosine nucleosidase (MTAN)	5JPC	none	sitting drop	vapor 1 week	1.5	7	2.5	2.1	(Banco <i>et al.</i> , 2016)
	5K1Z	none	sitting drop	vapor 1 week	1.5	9	2.6	2.25	
	5CCD	none	sitting drop	vapor 1 week	1.5	8	2.6	2.2	

1.2.6.1 HIV-1 protease

Human Immunodeficiency Virus type 1 (HIV-1) protease is an enzyme responsible for the maturation of viral particles into infectious HIV-1 virions. Without effective HIV-1 protease activity, the virions remain non-infectious, making HIV-1 protease a key target for successful antiretroviral (ART) drugs.

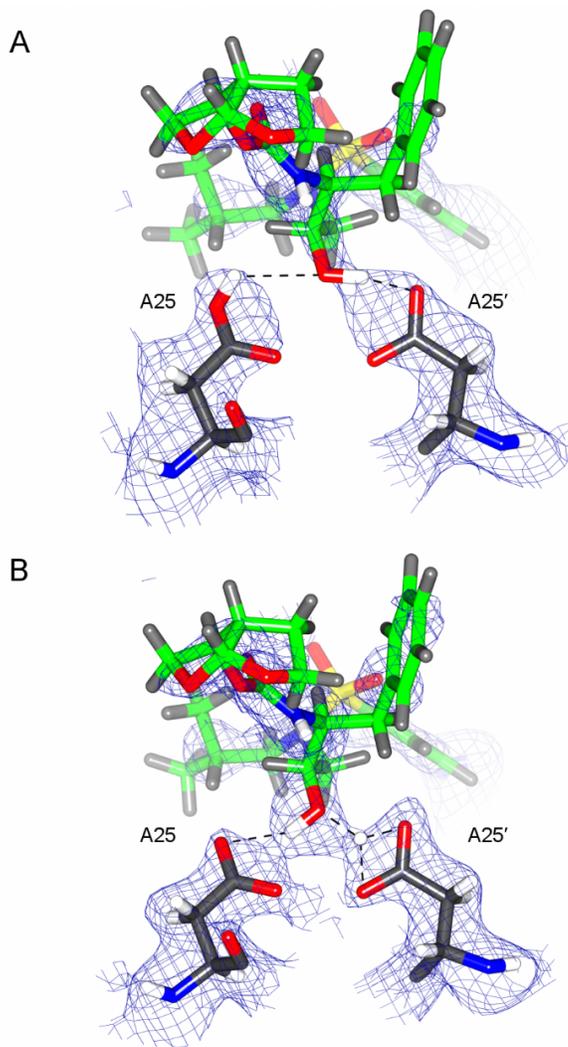


Figure 1.13 HIV-1 protease in complex with clinical drug darunavir (PDB 5E5J, 5E5K). Upon homodimerization, the catalytic site contains two closely positioned aspartic acid residues. **(A)** At pH 4.3, darunavir accepts a H-bond from A25 and donates a H-bond to A25'. **(B)** At pH 6.0, the two hydrogens undergo a transfer reaction in the catalytic site, captured for the first time by neutron crystallography. Carbon atoms belonging to darunavir are colored in green. ^1H atoms are shown in light grey while atoms that have undergone exchange to ^2H are shown in white. The blue mesh represents neutron SLD $2F_o - F_c$ maps contoured at $\sigma = 1.0$.

Neutron crystallography was used to probe the structure of HIV-1 protease in complex with the clinical drug darunavir¹⁸. The study shed light on the pH sensitivity of darunavir binding to the enzyme's active site. By determining structures at different pH values, the positions of hydrogen atoms before and after a pH-induced two-proton transfer between the drug and the enzyme were directly observed (Figure 1.13). The low-pH proton configuration in the catalytic site, critical for the catalytic action of this enzyme, was shown to be triggered by electrostatic effects arising from protonation state changes of surface residues far from the active site. These insights can help in the design of new, more effective ART drugs and could only be obtained by neutron crystallography since other structural methods (X-ray crystallography/cryo electron microscopy) do not permit visualization of the labile protons. Additionally, this study demonstrates that neutrons provide detailed information about proton transfer reactions in proteins.

1.2.6.2 *H. pylori* 5-methylthioadenosine nucleosidase (HpMTAN)

Gastrointestinal infection by the bacterium *Helicobacter pylori* is strongly associated with the development of gastric cancer. *H. pylori* 5'-methylthioadenosine nucleosidase (HpMTAN) is an interesting drug target because of its vital role in the production of menaquinone, vitamin K₂, using the alternative menaquinone biosynthetic pathway. HpMTAN offers a unique target for treating *H. pylori* infections without affecting the survival of the human microbiome. Neutron crystallography was performed to determine the hydrogen atom positions in the HpMTAN active site and thereby gain insight into the catalytic mechanism and transition state stabilization²⁴. The HpMTAN is responsible for hydrolyzing 6-amino-6-deoxy-futalosine in the second step of this alternative menaquinone biosynthetic pathway. Substrate binding of the adenine moiety is mediated almost exclusively by hydrogen bonds, and the proposed catalytic mechanism requires multiple proton transfer events. Of particular interest is the protonation state of residue D198, which possesses a pK_a above 8 and functions as a general acid to initiate the enzymatic reaction. In this study four neutron structures of HpMTAN were obtained (three of which were measured on IMAGINE (PDB 5CCD, 5JPC and 5K1Z)) and the results support a mechanism of D198 pK_a elevation through the unexpected sharing of a proton with atom N7 of the adenine moiety with

unconventional hydrogen-bond geometry. Additionally, the neutron structures also highlight active site features that promote the stabilization of the transition state and slight variations in these interactions that result in very different binding affinities between the Immucillian A (ImmA) (3R,4S)-4-(4-Chlorophenylthiomethyl)-1-[(9-deaza-adenin-9-yl)methyl]-3-hydroxypyrrolidine (DADMe-ImmA) analogs (Figure 1.14). Understanding how this enzyme functions, will provide insight into fine-tuning a drug to recognize only a specific target. This high specificity would eliminate some of the problematic off-target side effects and interactions that arise when broad spectrum antibiotics are used.

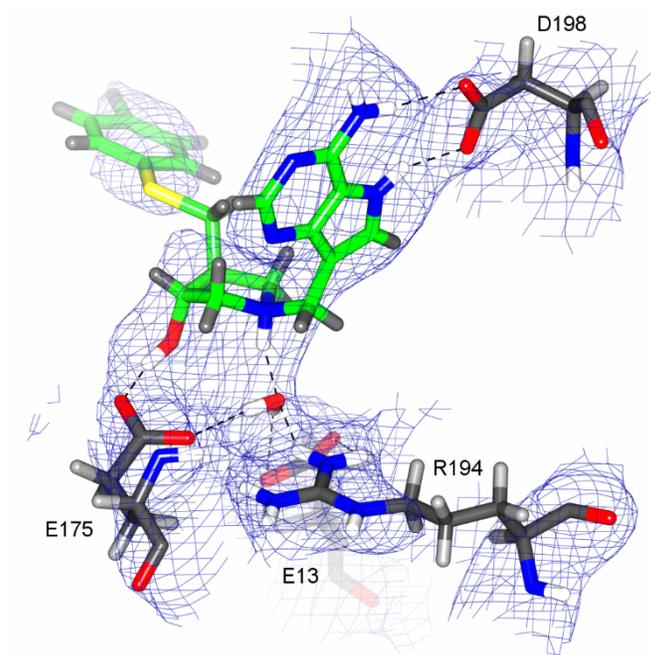


Figure 1.14 *Hp*MTAN in complex with the late transition state mimic *p*-ClPh-Thio-DADMe-ImmA (PDB 5K1Z). Substrate–D198 hydrogen bonds that mediate binding are unambiguously described by the neutron SLD. In addition, hydrogen bonding interactions with the tetrahedral transition state mimic are also apparent. Ligand carbon atoms are shown in green.¹H atoms are shown in light grey while atoms that have undergone exchange to ²H are shown in white. The blue mesh represents neutron SLD $2F_o - F_c$ map contoured at $\sigma = 1.0$.

1.2.6.3 *Neurospora crassa* lytic polysaccharide monooxygenase 9D (*Nc*LPMO9D)

Lytic polysaccharide monooxygenases (LPMOs) are copper-dependent metalloenzymes that activate molecular oxygen for the net insertion of one oxygen atom into the carbon-hydrogen bond of a glycosidic carbon in polymeric carbohydrates. LPMOs, thus, can disrupt the crystalline structure of

cellulose in lignocellulosic plant biomass, and can be used in industry to enhance the efficiency of cellulose degradation by other hydrolytic enzymes for biofuels production. Therefore, detailed knowledge of the LPMO enzymatic mechanism, which requires molecular oxygen, two electrons, and two protons, is important but remains poorly understood. An examination of protonation states enabled by neutron crystallography of the LPMO enzymatic resting state and density functional theory calculations identified a role for conserved histidine (H157) in promoting oxygen activation (Figure 1.15)^{17,25}. The result provides a new structural description of oxygen activation by LPMOs. Fully characterizing the LPMO reaction mechanism may inform strategies for increasing the efficiency of enzymatic polysaccharide hydrolysis for applications including bioethanol production.

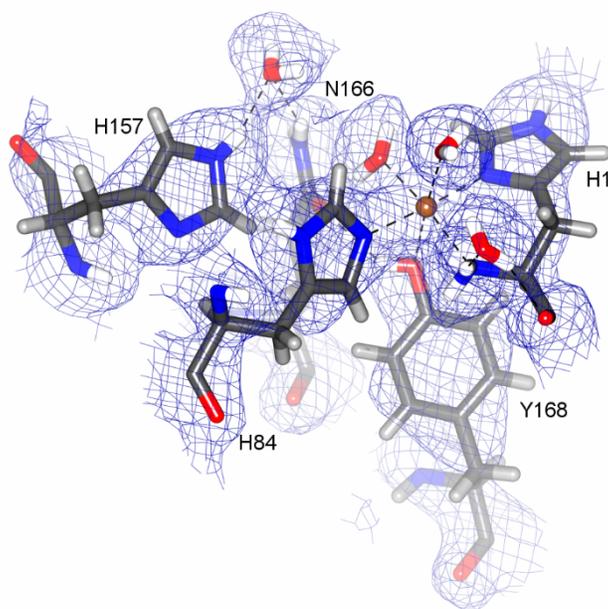


Figure 1.15 Active site of *NcLPMO9D* in the enzymatic resting state (PDB 5TKI). The conformation and tautomeric state of singly-protonated H157 revealed by neutron protein crystallography was confirmed to stabilize molecular oxygen (not shown) near the LPMO active site by quantum chemical calculations. The H1 backbone amide N and carbonyl O form H-bonds with the “pocket” water molecule which has been hypothesized to participate in substrate interactions. ¹H atoms are shown in light grey while atoms that have undergone exchange to ²H are shown in white. The copper atom is shown in brown. The blue mesh represents neutron SLD $2F_o - F_c$ map contoured at $\sigma = 1.0$.

1.2.6.4 Cholesterol oxidase

The protein microenvironment surrounding the flavin cofactor in flavoenzymes is key to the efficiency and diversity of reactions catalyzed by this class of enzymes. Cholesterol oxidase is found in some virulent bacteria that cause lung infections in immunocompromised individuals, such as people with acquired immune deficiency syndrome (AIDS). It catalyzes the oxidation of cholesterol to cholest-5-en-3-one via a hydride transfer to N5 of the flavin moiety of the bound flavin adenine dinucleotide (FAD) cofactor. In this study, neutron protein crystallography and high-level density functional theory were combined to map the position of the enzyme's hydrogen atoms in the oxidized state (prior to cholesterol oxidation) ¹⁶. The neutron crystallographic data revealed an unusually elongated main-chain nitrogen to hydrogen bond distance positioning the hydrogen atom towards the flavin N5 reactive center (Figure 1.16). Investigation of the structural features which could cause such an unusual conformation revealed a positively charged lysine side chain, conserved in other flavin-mediated oxidoreductases, in a second shell away from the cofactor acting to polarize the peptide bond through interaction with the carbonyl oxygen atom. Such an interaction between the protein and the cofactor may serve to position the receiving orbital of the flavin N5 for optimal alignment with the substrate (cholesterol) hydride. This work provides a model for the hydride transfer mechanism catalyzed by glucose-methanol-choline oxidoreductases.

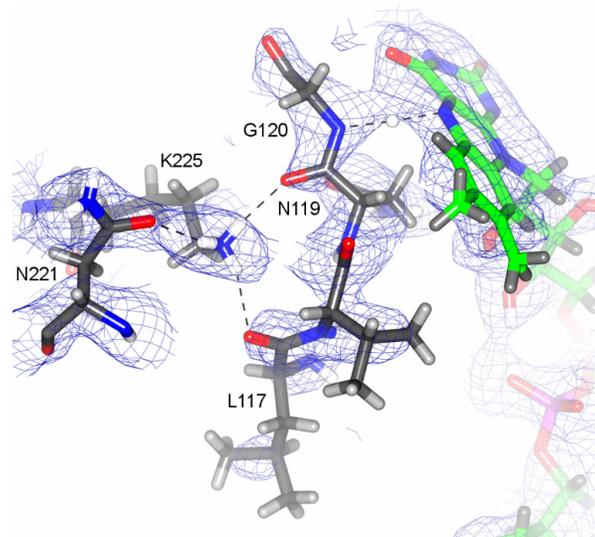


Figure 1.16 The cholesterol oxidase active site and its vicinity (PDB 5KWF). The positive charge of K225 facilitates the elongation of the G120 N–D bond by polarizing the peptide bond between N119 and G120. The G120 amide proton (^2H) is positioned equidistantly (1.66 Å) between G120N and FAD N5. FAD carbon atoms are shown in green. ^1H atoms are shown in light grey while atoms that have undergone exchange to ^2H are shown in white. The blue mesh represents neutron SLD $2F_o - F_c$ map contoured at $\sigma = 1.0$.

1.2.6.5 Dihydrofolate reductase

Ultra-high-resolution X-ray and neutron crystallography were paired to reveal the catalytic mechanism of dihydrofolate reductase (DHFR), an archetypal enzyme for studying proton and hydride transfer reactions in biological systems ¹⁴. DHFR catalyzes the NADPH-dependent reduction of dihydrofolate (DHF) to tetrahydrofolate (THF). In a complex that closely resembles the catalytically active state, DHFR stabilizes a particular substrate conformer and likely elevates the pK_a of the substrate atom that is protonated directly via a water molecule. An important step in the mechanism involves a proton donation to the N5 atom of DHF. Neutron and ultrahigh-resolution X-ray crystallographic studies of the pseudo-Michaelis ternary complex of *Escherichia coli* DHFR with folate and NADP^+ revealed that the N3 atom of the folate moiety is protonated and that D27 is negatively charged (Figure 1.17). The structure supports the existence of the keto tautomer owing to protonation of the N3 atom, suggesting that tautomerization is unnecessary for catalysis. In a 1.05 Å resolution X-ray structure of the ternary complex, conformational disorder of the M20 side chain is coupled to electron density for a partially-occupied water

within hydrogen-bonding distance of the N5 atom of folate. These findings suggest direct protonation of substrate by solvent. Thus, a catalytic mechanism for DHFR could be put forward, which involves stabilization of the keto tautomer of the substrate, elevation of the pK_a value of the N5 atom of DHF by D27, and protonation of N5 by water that gains access to the active site through fluctuation of the M20.

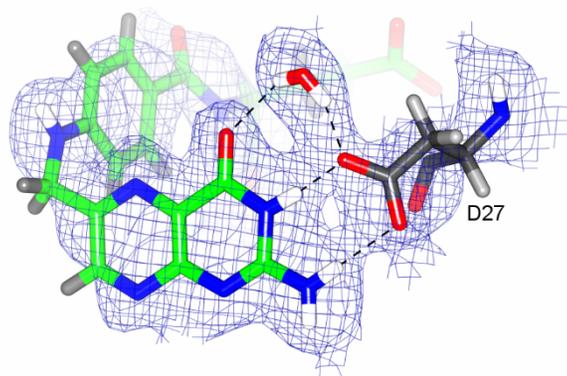


Figure 1.17 DHFR D27–folate hydrogen bonding(4PDJ). A deuterated water molecule is hydrogen bonded with both D27 and the N3-protonated (^2H) folate. Folate carbon atoms are shown in green. ^1H atoms are shown in light grey while atoms that have undergone exchange to ^2H are shown in white. The blue mesh represents neutron SLD $2F_o-F_c$ map contoured at $\sigma = 1.0$.

1.2.6.6 H-RAS GTPase

The small GTPase H-RAS is a prototype for nucleotide-binding proteins that function by cycling between bound guanosine triphosphate (GTP) and guanosine diphosphate (GDP). The neutron structure of H-RAS exemplifies the uniquely rich informational content provided by a single neutron diffraction experiment ²³. Previous efforts to elucidate the GTP hydrolysis mechanism of H-RAS have relied on the assumption that the γ -phosphate group of GTP bound at the active site of the enzyme would be deprotonated as is observed in solution ²⁶. This study, however, suggests that the γ -phosphate group of the GTP ligand remains protonated at physiological pH when bound at the active site of H-RAS in the ground state (Figure 1.18). The neutron structure revealed the modulation of the pK_a of the gamma-phosphate group by the H-RAS active site microenvironment and opens the possibility that the pre-catalytic state of the GTP hydrolysis reaction catalyzed by H-RAS is different than previously thought.

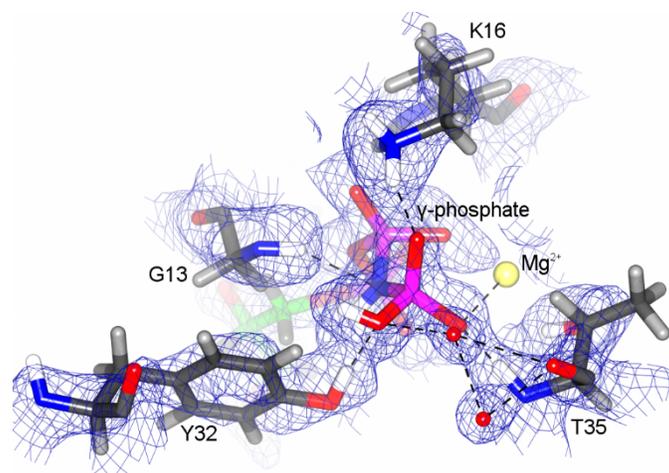


Figure 1.18 H-RAS in complex with GTP analogue GppNHp (PDB 4RSG). Unexpectedly, clear neutron SLD could be observed for a protonated (^2H) γ -phosphate in the $2F_o-F_c$ map. The deuterium atoms of the nucleophilic water (hydrogen bonded to the γ -phosphate and T35) are not visible in the density map, suggesting rotational freedom. Ligand carbon atoms are shown in green. ^1H atoms are shown in light grey while atoms that have undergone exchange to ^2H are shown in white. The blue mesh represents neutron SLD $2F_o-F_c$ map contoured at $\sigma = 1.0$.

1.2.7 Neutron protein crystallography at cryogenic temperature

Cold (2-10 Å) neutrons commonly used in the crystallographic experiments do not cause any observable radiation damage in protein crystals and thus most experiments are performed at room temperature. However, cryo-cooling of protein crystals in X-ray crystallography has been shown (in some cases) to provide improvements in data quality by reducing B-factors^{27,28}. Additionally, cryo-cooling functionality opens the important possibility for freeze-trapping studies of enzymatic reaction intermediates. Until recently, cryo-cooling has been challenging for large protein crystals since the solvent must be rapidly (flash)-cooled to a vitreous glass state to avoid ice formation that disrupts the crystal lattice. Protocols for cryo-neutron protein crystallography were first developed to cool and maintain large protein crystals (1-5mm³) at cryogenic temperatures (<15 K) using a closed cycle refrigerator²⁹. Data have been collected at cryogenic temperatures for lysozyme, concanavalin A and rubredoxin^{27,29}. In each case, crystals were mounted in fiber loops and flash-cooled by plunging directly into liquid N₂ (77K). The crystals were then transferred under liquid N₂ and mounted and cooled to 15 K on the cold head of a displax cryostat. A closed-cycle refrigerator (CCR) has been recently commissioned as a standard sample environment for

IMAGINE, expanding the capabilities of the instrument for neutron diffraction analysis at cryogenic temperatures. The CCR provides a constant sample temperature in the range of 4 to 450 K (Figure 19).



Figure 1.19 An automated lift moves the CCR up and down. Crystals are mounted directly onto the CCR cold finger *via* custom magnetic bases and then sealed with a heat shield (not shown) and a dome-shaped vacuum enclosure (visible at the bottom center of the photograph). A motorized vertical lift facilitates positioning of the mounted sample into the neutron beam. (Photo credit: Genevieve Martin, ORNL)

To demonstrate the suitability of the cryo-cooling set-up at IMAGINE to collect cryogenic data from protein crystals, proof-of-principle experiments have been completed in which a data set was collected on T4L lysozyme at 80K²⁰ (Figure 1.20). This set-up expands the capabilities of IMAGINE, allowing new scientific questions to be addressed by using the freeze-trapping technique to 'quench' kinetic processes and capture reaction intermediaries. For example, a data set was recently collected at 10K on *NcLPMO9D*. The crystal was frozen to capture the activation of molecular dioxygen at the active site of the enzyme³⁰.

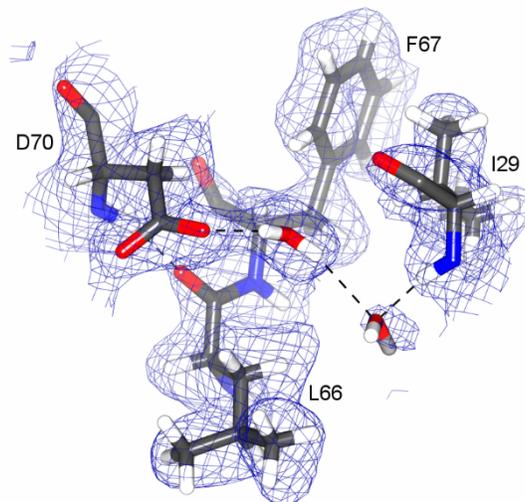


Figure 1.20 T4 Lysozyme “cavity 1” as observed at 80 K (PDB 5VNQ). Hydrogen atoms that were substituted with ^2H by covalent labeling and exchange are shown in white. The blue mesh represents neutron SLD $2F_o - F_c$ map contoured at $\sigma = 1.0$.

1.2.8 Instrument development

Collecting data with high completeness is complicated by the cylindrical detector geometry as well as a significant blind region arising from the curvature of the Ewald sphere at long wavelengths. Working with a single axis goniometer, crystals are typically reoriented to increase completeness by collecting data in a second crystal orientation. A miniaturized kappa goniometer head used at synchrotron beamlines has been modified and built for incorporation on IMAGINE³¹. The Arinax MiniKappa MK3 provides an additional kappa and phi axis (Table 1.4). This will enable data to be collected at multiple crystal settings within a single data collection run, which is expected to increase the data collection efficiency of IMAGINE.

Table 1.4 Mini kappa specifications.

	Kappa	Phi	Chi
Rotational precision (°)	±0.05	±0.09	N/A
Travel range (°)	0 to 255	360	0 to 48
Resolution(mdeg)	0.35	1.45	N/A

The CrystalPlan software has been adapted for IMAGINE to optimize data collection strategies using the mini-Kappa goniometer ³². CrystalPlan uses the UB matrix determined from a short Laue geometry alignment image to predict an optimal set of crystal orientations that will maximize the completeness using a minimal number of crystal settings. These settings are then loaded and run in the instrument GUI data collection script. This will allow IMAGINE users to maximize use of available beam time per sample per experiment, and increase the data collection efficiency.

1.2.9 Conclusions

IMAGINE is a world class neutron diffractometer that specializes in neutron protein crystallography. The instrument offers access to crystals with unit cells of up to 150 Å, complementing the capabilities of the ORNL Spallation Neutron Source MaNDi time-of-flight instrument which is optimized for larger unit cells ³³. Recent upgrades have added to the capabilities of IMAGINE by providing access to data collection at variable temperatures (4K-450K), which will broaden the science that can be done on this beam line. Additionally, low temperature macromolecular neutron diffraction data can now be obtained from cryo-cooled protein crystals, making experiments with freeze-trapped intermediates in proteins possible. The unique combination of scientific and technical expertise as well as facilities for biodeuteration, crystal growth and macromolecular crystallography at ORNL provides the user community with supporting capabilities for sample preparation that help drive the productivity and development of the structural biology program at IMAGINE, allowing new scientific questions to be addressed.

1.2.10 Acknowledgements

The authors acknowledge engineering, sample environment development and project management support for the IMAGINE instrument and IMAGINE CCR from Amy Jones, John Wenzel and Lakeisha Walker, Neutron Sciences Directorate, Oak Ridge National Laboratory. The authors acknowledge Lee Robertson and colleagues, Neutron Sciences Directorate, Oak Ridge National Laboratory for design and implementation of the neutron optics system for the IMAGINE beamline. The authors thanks Vickie Lynch for adapting CrystalPlan. Research conducted at ORNL's High Flux Isotope Reactor was sponsored by the Scientific User Facilities Division, Office of Basic Energy Sciences, US Department of Energy. This research is supported by UT Battelle, LLC under Contract DE-AC05-00OR22725 for the U.S. Department of Energy, Office of Science. IMAGINE acquisition and installation was partially supported by the National Science Foundation under Grant No. 0922719. GCS is supported in part by the National Research Foundation (NRF), South Africa.

1.2.11 References

1. Sears, V. F. Neutron scattering lengths and cross sections. *Neutron News* **3**, 26-37 (1992).
2. O'Dell, W. B., Bodenheimer, A. M. & Meilleur, F. Neutron protein crystallography: A complementary tool for locating hydrogens in proteins. *Arch Biochem Biophys* **602**, 48-60, doi:10.1016/j.abb.2015.11.033 (2016).
3. Shu, F., Ramakrishnan, V. & Schoenborn, B. P. Enhanced visibility of hydrogen atoms by neutron crystallography on fully deuterated myoglobin. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 3872-3877, doi:10.1073/pnas.060024697 (2000).
4. Niimura, N. & Podjarny, A. D. in *IUCr Monographs on Crystallography* Vol. 25 (Oxford University Press, Oxford, 2011).
5. Blakeley, M. P. Neutron macromolecular crystallography. *Crystallogr Rev* **15**, 157-218, doi:10.1080/08893110902965003 (2009).
6. Blakeley, M. P., Hasnain, S. S. & Antonyuk, S. V. Sub-atomic resolution X-ray crystallography and neutron crystallography: promise, challenges and potential. *IUCrJ* **2**, 464-474, doi:10.1107/S2052252515011239 (2015).
7. Blakeley, M. P., Langan, P., Niimura, N. & Podjarny, A. Neutron crystallography: opportunities, challenges, and limitations. *Current opinion in structural biology* **18**, 593-600, doi:10.1016/j.sbi.2008.06.009 (2008).
8. Meilleur, F., Myles, D. A. & Blakeley, M. P. Neutron Laue macromolecular crystallography. *European biophysics journal : EBJ* **35**, 611-620, doi:10.1007/s00249-006-0074-6 (2006).
9. Myles, D. A. Neutron protein crystallography: current status and a brighter future. *Current opinion in structural biology* **16**, 630-637, doi:10.1016/j.sbi.2006.08.010 (2006).
10. Langan, P. & Chen, J. C. Seeing the chemistry in biology with neutron crystallography. *Phys Chem Chem Phys* **15**, 13705-13712, doi:10.1039/c3cp51760h (2013).
11. Chen, J. C. & Unkefer, C. J. Fifteen years of the Protein Crystallography Station: the coming of age of macromolecular neutron crystallography. *IUCrJ* **4**, 72-86, doi:10.1107/S205225251601664X (2017).
12. Oksanen, E., Chen, J. C. & Fisher, S. Z. Neutron Crystallography for the Study of Hydrogen Bonds in Macromolecules. *Molecules* **22**, E596, doi:10.3390/molecules22040596 (2017).
13. Meilleur, F. *et al.* The IMAGINE instrument: first neutron protein structure and new capabilities for neutron macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **69**, 2157-2160, doi:10.1107/S0907444913019604 (2013).
14. Wan, Q. *et al.* Toward resolving the catalytic mechanism of dihydrofolate reductase using neutron and ultrahigh-resolution X-ray crystallography. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 18225-18230, doi:10.1073/pnas.1415856111 (2014).
15. Wan, Q. *et al.* Preliminary joint X-ray and neutron protein crystallographic studies of ecDHFR complexed with folate and NADP⁺. *Acta crystallographica. Section F, Structural biology communications* **70**, 814-818, doi:10.1107/S2053230X1400942X (2014).
16. Golden, E. *et al.* An extended N-H bond, driven by a conserved second-order interaction, orients the flavin N5 orbital in cholesterol oxidase. *Scientific reports* **7**, 40517, doi:10.1038/srep40517 (2017).
17. O'Dell, W. B., Agarwal, P. K. & Meilleur, F. Oxygen Activation at the Active Site of a Fungal Lytic Polysaccharide Monooxygenase. *Angewandte Chemie* **56**, 767-770, doi:10.1002/anie.201610502 (2017).
18. Gerlits, O. *et al.* Long-Range Electrostatics-Induced Two-Proton Transfer Captured by Neutron Crystallography in an Enzyme Catalytic Site. *Angewandte Chemie* **55**, 4924-4927, doi:10.1002/anie.201509989 (2016).
19. Aggarwal, M. *et al.* Neutron structure of human carbonic anhydrase II in complex with methazolamide: mapping the solvent and hydrogen-bonding patterns of an effective clinical drug. *IUCrJ* **3**, 319-325, doi:10.1107/S2052252516010514 (2016).

20. Li, L. *et al.* Neutron crystallographic studies of T4 lysozyme at cryogenic temperature. *Protein Sci* **26**, 2098-2104, doi:10.1002/pro.3231 (2017).
21. Hiromoto, T. *et al.* Neutron structure of the T26H mutant of T4 phage lysozyme provides insight into the catalytic activity of the mutant enzyme and how it differs from that of wild type. *Protein Sci* **26**, 1953-1963, doi:10.1002/pro.3230 (2017).
22. Gerlits, O. O., Coates, L., Woods, R. J. & Kovalevsky, A. Mannobiose Binding Induces Changes in Hydrogen Bonding and Protonation States of Acidic Residues in Concanavalin A As Revealed by Neutron Crystallography. *Biochemistry* **56**, 4747-4750, doi:10.1021/acs.biochem.7b00654 (2017).
23. Knihtila, R., Holzapfel, G., Weiss, K., Meilleur, F. & Mattos, C. Neutron Crystal Structure of RAS GTPase Puts in Question the Protonation State of the GTP gamma-Phosphate. *J Biol Chem* **290**, 31025-31036, doi:10.1074/jbc.M115.679860 (2015).
24. Banco, M. T. *et al.* Neutron structures of the *Helicobacter pylori* 5'-methylthioadenosine nucleosidase highlight proton sharing and protonation states. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 13756-13761, doi:10.1073/pnas.1609718113 (2016).
25. O'Dell, W. B., Swartz, P. D., Weiss, K. L. & Meilleur, F. Crystallization of a fungal lytic polysaccharide monooxygenase expressed from glycoengineered *Pichia pastoris* for X-ray and neutron diffraction. *Acta crystallographica. Section F, Structural biology communications* **73**, 70-78, doi:10.1107/S2053230X16020318 (2017).
26. Cheng, H., Sukal, S., Callender, R. & Leyh, T. S. gamma-phosphate protonation and pH-dependent unfolding of the Ras.GTP.Mg²⁺ complex: a vibrational spectroscopy study. *J Biol Chem* **276**, 9931-9935, doi:10.1074/jbc.M009295200 (2001).
27. Blakeley, M. P., Kalb, A. J., Helliwell, J. R. & Myles, D. A. The 15-K neutron structure of saccharide-free concanavalin A. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 16405-16410, doi:10.1073/pnas.0405109101 (2004).
28. Casadei, C. M. *et al.* Heme enzymes. Neutron cryo-crystallography captures the protonation state of ferryl heme in a peroxidase. *Science* **345**, 193-197, doi:10.1126/science.1254398 (2014).
29. Myles, D. A. A., Dauvergne, F., Blakeley, M. P. & Meilleur, F. Neutron protein crystallography at ultra-low (< 15 K) temperatures. *J Appl Crystallogr* **45**, 686-692, doi:10.1107/S0021889812019784 (2012).
30. O'Dell, W. B. & *Structural Investigations of the Fungal Lytic Polysaccharide Monooxygenase Reaction Mechanism using X-ray and Neutron Protein Crystallography*. Doctor of Philosophy thesis, North Carolina State University, (2017).
31. *Mini-Kappa Goniometer Head*, <<https://www.arinax.com/mini-kappa-goniometer-head/>> (
32. Zikovskiy, J., Peterson, P. F., Wang, X. P., Frost, M. & Hoffmann, C. CrystalPlan: an experiment-planning tool for crystallography. *J Appl Crystallogr* **44**, 418-423, doi:10.1107/s0021889811007102 (2011).
33. Coates, L. *et al.* The Macromolecular Neutron Diffractometer MaNDi at the Spallation Neutron Source. *J Appl Crystallogr* **48**, 1302-1306, doi:10.1107/S1600576715011243 (2015).

1.3 A Guide to neutron crystallography data collection and processing for modelling hydrogen atoms in protein structures

The following work was adapted from: Schröder, G. C. & Meilleur, F. Neutron Crystallography Data Collection and Processing for Modelling Hydrogen Atoms in Protein Structures. *J. Vis. Exp.* 1–38 (2020). doi:10.3791/61903. For full Protocol details, please refer to the aforementioned publication.

1.3.1 Keywords

Neutron protein crystallography, neutrons, proteins, hydrogen atoms, protonation, enzymes, metalloproteins, structural biology, polysaccharide monooxygenases, reaction mechanism, X-ray crystallography.

1.3.2 Summary

Neutron protein crystallography is a structural technique that permits the localization of hydrogen atoms, thereby providing important mechanistic details of protein function. We present here the workflow for mounting a protein crystal, neutron diffraction data collection, structure refinement and analysis of the neutron scattering length density maps.

1.3.3 Abstract

Neutron crystallography is a structural technique that allows determination of hydrogen atom positions within biological macromolecules, yielding mechanistically important information about protonation and hydration states while not inducing radiation damage. X-ray diffraction, in contrast, provides only limited information on the position of light atoms and X-ray beam rapidly induces radiation damage of photosensitive redox cofactors and metal centers. Presented here is the workflow employed for the IMAGINE and MaNDi beamlines at Oak Ridge National Laboratory (ORNL) to obtain a neutron diffraction structure once a protein crystal of suitable size ($> 0.1 \text{ mm}^3$) has been grown. We demonstrate mounting of hydrogenated protein crystals in quartz capillaries for neutron diffraction data collection. Also presented is the vapor exchange process of the mounted crystals with D_2O -containing buffer to ensure replacement of hydrogen atoms at exchangeable sites with deuterium. The incorporation of deuterium

reduces the background arising from the incoherent scattering of hydrogen atoms and prevents density cancellation caused by their negative coherent scattering length. Sample alignment and room temperature data collection strategies are illustrated using quasi-Laue data collection at IMAGINE at the High Flux Isotope Reactor (HFIR). Furthermore, crystal mounting and rapid freezing in liquid nitrogen for cryo-data collection to trap labile reaction intermediates is demonstrated at the MaNDi time-of-flight instrument at the Spallation Neutron Source (SNS). Preparation of the model coordinate and diffraction data files and visualization of the neutron scattering length density (SLD) maps will also be addressed. Structure refinement against neutron data-only or against joint X-ray/neutron data to obtain an all-atom structure of the protein of interest will finally be discussed. The process of determining a neutron structure will be demonstrated using crystals of the lytic polysaccharide monooxygenase *Neurospora crassa* LPMO9D, a copper-containing metalloprotein involved in the degradation of recalcitrant polysaccharides *via* oxidative cleavage of the glycosidic bond.

1.3.4 Introduction

Neutron macromolecular crystallography is a technique that provides a unique window into the structure and underlying chemistry of proteins. Conceptually similar to X-ray diffraction, neutron diffraction provides atomistic details of macromolecular structure, however, the interaction of neutrons with nuclei enables localization of light atoms, often difficult to detect with X-ray diffraction¹. During X-ray diffraction, X-rays scatter from the electron cloud, making light atoms such as hydrogen (H) poorly visible in electron density maps that do not have near sub-Ångström resolution². In contrast, the scattering intensity of neutrons depends on complex interactions with the nucleus, with isotopes of the same element displaying different scattering lengths. Therefore, light atoms and their isotopes, such as hydrogen (¹H) and deuterium (²H or D), have comparable visibility to the backbone carbon, nitrogen and oxygen atoms in neutron scattering length density (SLD) maps. Furthermore, since the magnitude of neutron scattering is independent of the number of electrons, scattering from light elements is not obscured by heavy elements when they are in close vicinity to each other, as is observed in X-ray scattering. The enhanced visibility of H and its isotope D when employing neutron diffraction provides valuable information about the

protonation state of catalytically important residues, cofactors and ligands and aids the orientation of water molecules, revealing important information about catalytic mechanisms and protein chemistry³. Neutron diffraction also offers the advantage of being a non-destructive technique, particularly suited to biological samples sensitive to ionization such as proteins with metal centers or photosensitive redox cofactors². The primary focus of this article is to provide an overview of the workflow to obtain a high-quality neutron protein crystal structure. We refer the interested reader to Podjarny *et al.*⁴, Blakeley⁵, Blakeley *et al.*⁶ and O'Dell *et al.*³ for an excellent overview of neutron protein diffraction and Ashkar *et al.*⁷ for further applications of neutron scattering.

Neutrons are primarily generated during nuclear reactions employing either of two processes: nuclear fission at reactor sources or spallation at accelerator-based sources⁸. Reactor sources provide a continuous neutron beam by employing nuclear fission of the ²³⁵U isotope while spallation neutron sources produce a pulsed neutron beam by bombarding a target, for example a liquid metal such as mercury, with protons⁹. Oak Ridge National Laboratory (ORNL) in Oak Ridge, Tennessee, hosts both a steady-state neutron source at the High Flux Isotope Reactor (HFIR) and a 60 Hz pulsed source at the Spallation Neutron Source (SNS). The IMAGINE beamline, located at the HFIR, is a neutron diffractometer optimized for biological macromolecules (Figure 1.21)¹⁰. IMAGINE employs a neutron image plate detector to measure quasi-Laue data using a narrow bandpass in the range of 2.8 – 4.5 Å from single crystals with unit cell edges <150 Å. The Macromolecular Neutron Diffractometer (MaNDi), located at the SNS, is a time-of-flight (TOF) Laue neutron diffractometer equipped with a spherical detector array frame (DAF) (Figure 1.22)¹¹. MaNDi measures data from single crystals with unit cell edges in the range of 10 – 300 Å by employing a tunable 2 Å-wavelength bandwidth between 2.0 – 6.0 Å¹².

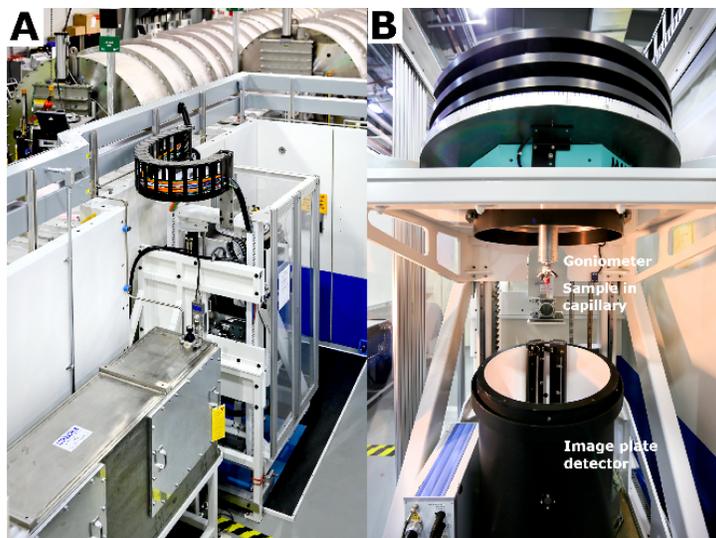


Figure 1.21 The IMAGINE Instrument at the High Flux Isotope Reactor. **(A)** The IMAGINE instrument in the cold neutron guide hall. **(B)** Sample in mounted in a quartz capillary attached with putty to the goniometer. The sample and detector table closes to position the crystal and the cylindrical image plate in the neutron beam. Modified with the permission of the International Union of Crystallography¹³. Images provided with permission of Genevieve Martin, Oak Ridge National Laboratory.

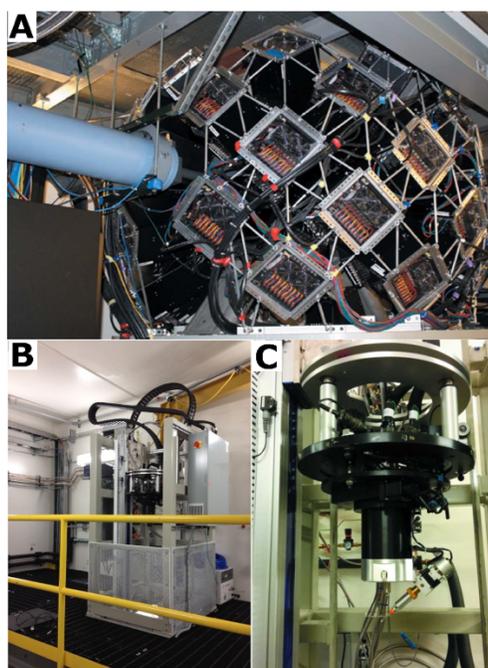


Figure 1.22 The MaNDi Instrument at the Spallation Neutron Source. **(A)** The MaNDi Anger camera detector array. Reproduced with permission the International Union of Crystallography¹¹. **(B)** MaNDi moveable sample stage. **(C)** Sample mounted in quartz capillary mounted on the goniometer at MaNDi for room temperature data collection. Images provided with permission of Genevieve Martin, Oak Ridge National Laboratory.

The process of generating neutrons is highly energy intensive, resulting in relatively weak neutron beam fluxes when contrasted to X-ray beam fluxes at synchrotron sources¹⁴. To ensure sufficient signal-to-noise ratios during data collection, it is necessary to grow crystals of suitable size and quality¹⁵. Typically, crystals with volumes $> 0.1 \text{ mm}^3$ are needed to collect data with adequate statistics¹⁶. In addition to lower fluxes, inherent properties of the interaction between neutrons and the sample nuclei must be taken into consideration¹⁷. The scattering length of neutrons differs for isotopes of the same element, a property which can be advantageously exploited in small angle neutron scattering (SANS) to mask or highlight regions of a sample – a process known as contrast matching¹⁸. In diffraction experiments, the negative coherent neutron scattering length of H (-3.741 fm for ^1H) can lead to cancellation of neutron scattering density map features since the coherent neutron scattering lengths of other biologically relevant atoms, including carbon (6.6511 fm for ^{12}C), nitrogen (9.37 fm for ^{14}N), oxygen (5.803 fm for ^{16}O), phosphorus (5.13 fm for ^{31}P) and sulfur (2.804 fm for ^{32}S), are positive (Table 1.5)^{12,14}. Furthermore, the large incoherent scattering length of H (25.274 fm), increases the background during data collection, hampering the quality of the dataset and compromising data resolution⁷. To circumvent these limitations introduced by H it is necessary, for neutron diffraction, to exchange H for its isotope deuterium, $^2\text{H}(\text{D})$, which has a positive coherent neutron scattering length (6.671 fm) and significantly lower incoherent scattering length (4.04 fm)²⁰. This can be achieved by perdeuteration, a process in which the protein is expressed by organisms grown in fully deuterated media ensuring complete incorporation of D at H sites²¹. It is also possible to partially deuterate protein by replacing H with D solely at the exchangeable sites (titratable groups) while the non-exchangeable carbon-bound sites remain hydrogenated²². This can be achieved by growth of hydrogenated protein crystals in deuterated mother liquor²³. Most commonly however, H/D exchange of hydrogenated proteins is performed by vapor exchange following growth of suitably large crystals in H_2O -based buffer²⁴. In such cases, crystals are mounted in a quartz capillary and vapor-equilibrated with D_2O -based mother liquor.

Table 1.5 Neutron scattering lengths and incoherent scattering values. Adapted from Sears, 1992¹⁷.

Isotope	Coherent scattering length (fm)	Incoherent scattering length (fm)
1H	-3.741	25.274
2H	6.671	4.04
12C	6.6511	0
14N	9.37	2
16O	5.803	0
23Na	3.63	3.59
24Mg	5.66	0
31P	5.13	0.2
32S	2.804	0
35Cl	11.65	6.1
39K	3.74	1.4
40Ca	4.8	0
55Mn	-3.73	1.79
56Fe	9.94	0
63Cu	6.43	0.22
64Zn	5.22	0

The limited neutron fluxes at neutron sources result in longer data collection times, ranging from days to several weeks²⁵. At ORNL, both IMAGINE and MaNDi employ a narrow wavelength bandpass in the 2–6 Å range to optimize data collection²⁶. Data can be collected at room temperature or at cryo-temperature. Cryo-data collection can potentially improve data quality and opens up the possibility for freeze-trapping catalytic intermediates. Following neutron diffraction data collection, an X-ray dataset is typically collected on the same crystal at the same temperature or on a crystal grown under identical conditions²⁷. Data collection at the same temperature allows structure refinement to be performed against both X-ray and neutron data, preventing any potential temperature-induced artefacts such as changes in the visibility and position of waters or the occupancies of residues with alternate conformations²⁸. Joint X-ray neutron data refinement increases the data-to-parameter ratio and provides the advantage of allowing the protein backbone coordinates to be refined against the X-ray data, while the neutron diffraction data is used to refine the position of the H/D atoms²⁹. This is particularly useful using partially deuterated samples,

where density cancellation due to H atoms at non-exchangeable sites on the protein is present. Although the number of X-ray structures far exceeds the number of neutron structures deposited in the Protein Data Bank (PDB), software packages initially designed for the refinement of X-ray data have been expanded to encompass neutron data as well^{3,30,31}. Following data collection, models can be refined using refinement packages such as *phenix.refine*, CNSsolve (nCNS) or SHELXL^{29,32-34}. During the refinement process, neutron scattering density maps can be visualized for manual fitting using COOT³⁵. Following structure solution, the coordinates and the neutron and/or X-ray data diffraction data files can be submitted to the PDB, who will validate and deposit the model, making it available for public access^{19,30,31}.

Structural analysis of proteins is a multifaceted approach in which numerous techniques are used to probe their function and mechanism³⁶. Neutron protein crystallography provides valuable chemical insights to expand on and complement findings from additional studies such as X-ray diffraction, spectroscopy, nuclear magnetic resonance (NMR) or micro crystal electron diffraction (microED)³⁷. Neutron protein diffraction is uniquely positioned to provide insights into enzymatic mechanisms, since H atoms are central to their chemistry. The absence of radiation damage induced by neutrons make them a probe exceptionally suited to the study of metalloproteins³⁸. We present here a representative example of the process of neutron protein diffraction from sample preparation to data collection, refinement and analysis (Figure 1.23). Crystals of sufficient size for neutron diffraction experiments have been grown of the metalloprotein *Neurospora crassa* LPMO9D (*NcLPMO9D*). *NcLPMO9D* is a copper-containing metalloprotein involved in the degradation of recalcitrant cellulose by oxygen atom insertion at the glycosidic bond^{39,40}. The *NcLPMO9D* active site contains a mononuclear copper center within a characteristic “histidine-brace” composed of the N-terminal histidine and a second conserved histidine (Figure 1.24)⁴¹. The N-terminal of fungal LPMOs is methylated but the post-translational modification does not occur during recombinant expression in yeast. In the *NcLPMO9D* resting state, the copper center is present in a Cu²⁺ oxidation state and is activated by single electron reduction to Cu¹⁺, allowing molecular oxygen to bind and be activated by rapidly being reduced to a superoxide species^{42,43}. The overall

NcLPMO9D reaction requires further addition of one electron and two protons to form the hydroxylated polysaccharide product⁴⁴. The identity of the activated oxygen species responsible for hydrogen atom abstraction (HAA) from the polysaccharide substrate has not been identified and intensive structural and computational studies are currently ongoing^{45,46}. Given the redox chemistry at the *NcLPMO9D* active site, mitigation of radiation damage is particularly pertinent. We illustrate here room temperature and cryo-temperature data collection on *NcLPMO9D* crystals to determine the *NcLPMO9D* structure in the resting state and in the activated reduced form, respectively⁴⁷. Emphasis will be given to protein crystal mounting, beamline instrument setup for data collection, the preparation of the data and coordinate files and the refinement steps necessary to model an all-atom neutron structure.

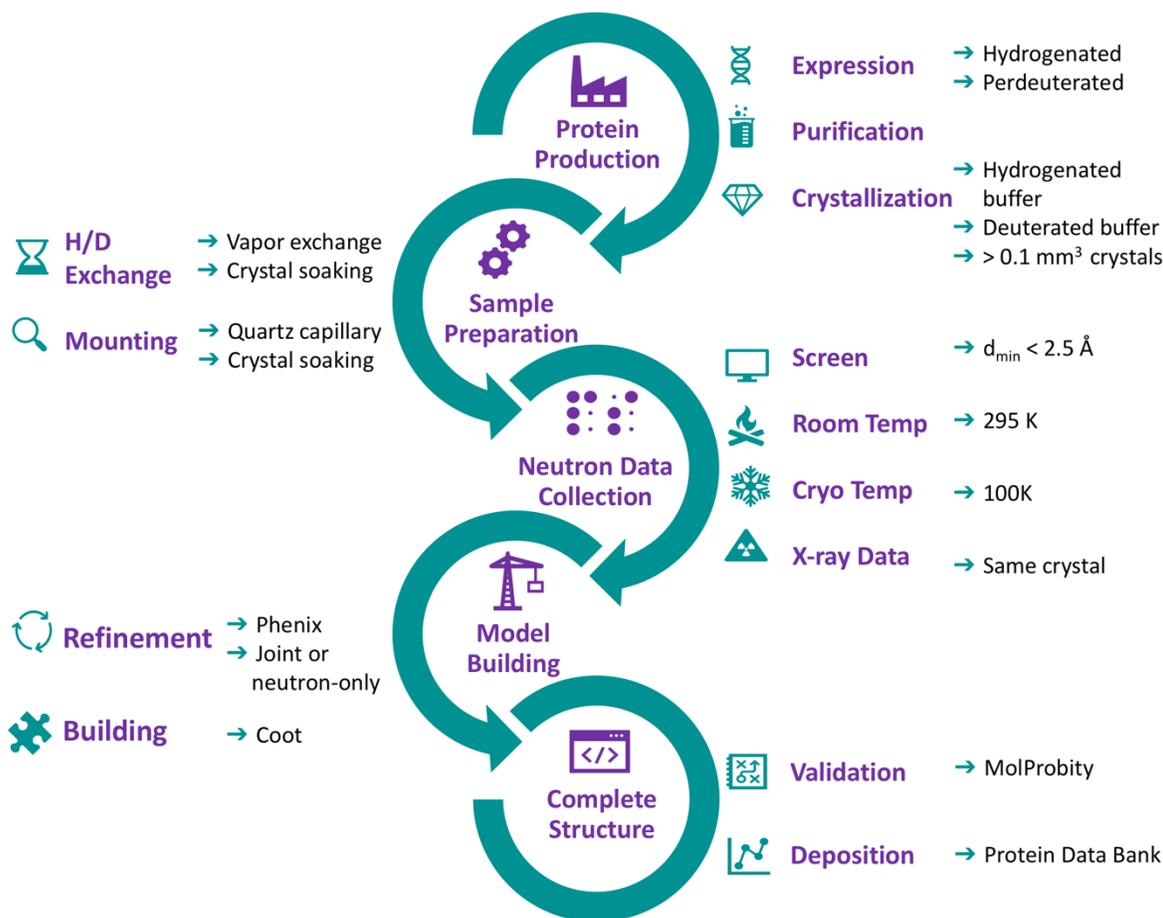


Figure 1.23 Flow chart of neutron protein crystallography workflow. **Protein Production:** In order to obtain a neutron structure, protein is first expressed. Bacterial expression in H₂O- or D₂O-based media is typically used to produce a high yield of hydrogenated or perdeuterated recombinant protein, respectively. The protein is purified in H₂O-based buffer and then crystallized in either H₂O- or D₂O-based crystallization buffer to grow crystals to a minimum size of 0.1 mm³. **Sample Preparation:** Prior to neutron diffraction data collection, H₂O-grown crystals undergo H/D exchange to exchange the protein titratable H atoms with D. H/D exchange can be done by direct soaking of the crystals in deuterated crystallization buffer, equilibration of the crystallization drop with a D₂O-based reservoir, or by mounting the crystals in quartz capillaries for vapor exchange with deuterated crystallization buffer. **Neutron Data Collection:** Following H/D exchange, potential crystals are screened to determine the diffraction quality. Crystals with a minimum resolution of 2.5 Å are considered suitable for a full dataset to be collected. Crystals are mounted in quartz capillaries for data collection at room temperature or flash frozen in a cryo-loop for data collection at cryogenic temperature. An X-ray dataset is collected on the same (or an identical) crystal at the same temperature. **Model Building:** Refinement is performed using *phenix.refine* against both neutron and X-ray data or against the neutron data only. Manual model building of the protein structure is performed in Coot using the neutron SLD maps. **Complete Structure:** Following completion of the protein structure, the coordinate model is validated and deposited in the Protein Data Bank.

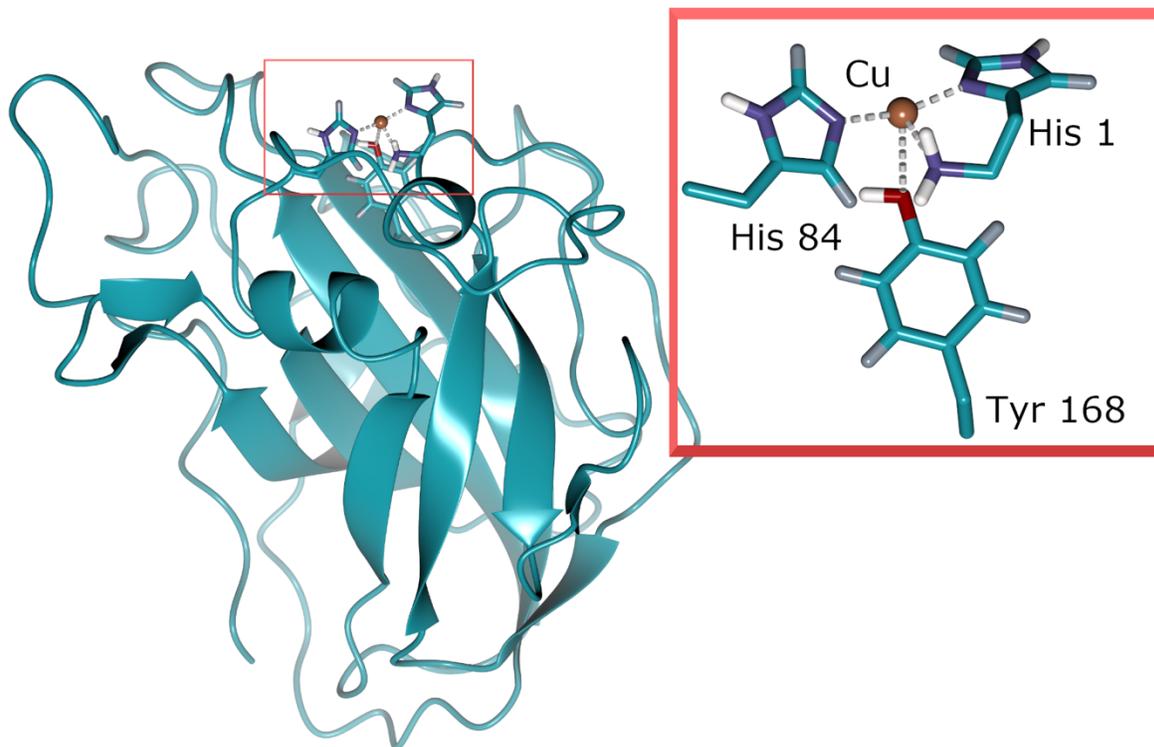


Figure 1.24 Structure of the lytic polysaccharide monooxygenase *NcLPMO9D*. The *NcLPMO9D* copper-active site is located on a flat polysaccharide binding surface. The copper is coordinated by two histidine residues in a classical “histidine brace” as well as an axial tyrosine residue.

1.3.5 Representative results

Neutron diffraction data on crystals of a lytic polysaccharide monooxygenase from *Neurospora crassa* (*NcLPMO9D*) were collected on IMAGINE at the HFIR at room temperature and at MaNDi at the SNS under cryo-conditions following the protocol described above. Crystals of the hydrogenated protein grown in H₂O-based buffer with volume greater than 0.1 mm³ were used (Figure 1.25) Crystals were mounted in quartz capillaries and vapor exchange with the D₂O-based buffer was performed for three weeks prior to data collection.

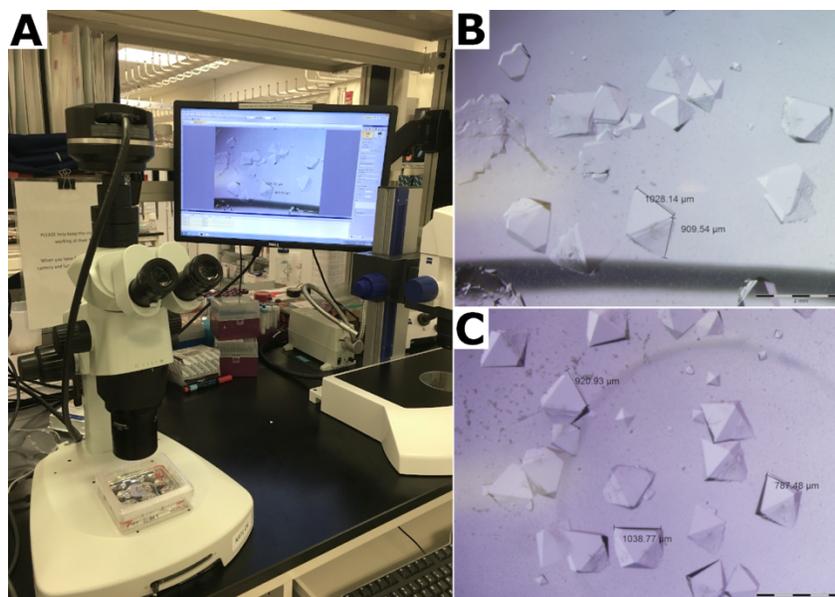


Figure 1.25 Crystal with sufficient volume in sitting drop crystallization tray. **(A)** Large crystals are grown in sitting drops set up in 9-well siliconized glass plates. **(B and C)** Crystals are measured to identify those with volume $> 0.1 \text{ mm}^3$.

Room temperature data collection was performed on the IMAGINE beamline (Figure 1.21). A four hour white-beam test led to high resolution diffraction suggesting that the crystal was of suitable size and quality for a full dataset to be collected. In addition to providing preliminary information on the diffraction quality of the crystal, the initial broad bandpass exposure can be used to index the diffraction pattern and determine the crystal orientation matrix. Given the $P2_1$ space group of the crystal, a data collection strategy of 18 frames with a collection time of 20 hours per frame was implemented. As with X-ray diffraction data collection, higher symmetry space groups require fewer frames (i.e., less angular coverage) to collect a complete dataset. The data were collected in quasi-Laue mode using a wavelength range of 2.8 – 4.0 Å. Following data collection, the data were indexed, integrated scaled and merged to give a neutron SLD file in MTZ format at a resolution of 2.14 Å. Data were evaluated to be of sufficient quality following similar guidelines for X-ray data analysis, although a completeness of 80 % and a $CC_{1/2}$ of at least 0.3 were considered acceptable since neutron protein diffraction is a flux-limited technique.

Following room-temperature neutron diffraction data collection, the same crystal was used to collect a room temperature X-ray diffraction dataset at 1.90 Å resolution. The X-ray data were used to determine the positions of the “heavier” atoms including C,N,O and S. The structure refined against the X-ray data alone was then used as the starting model to perform a joint refinement against the X-ray and neutron data. Phenix ReadySet was used to add H atoms at non-exchangeable sites, H and D atoms at exchangeable sites and D atoms to water molecules of the starting X-ray model. Following this model preparation, iterative refinements were performed against both datasets. Interactive model building was performed in Coot by visually inspecting the density maps to orientate side-chains and water molecules accordingly. The neutron data were primarily used to determine protonation states and water molecule orientations. Comparison of the electron density map of residues such as serine and tryptophan and the corresponding neutron SLD map illustrate the information that can be gained on protonation states at H/D exchangeable sites from neutron protein diffraction (Figure 1.26). A map overlay of electron and neutron SLD maps for water molecules also indicate that while hydrogen bond interactions can be inferred from X-ray data, neutrons provide clear information regarding the orientation of these hydrogen bonds (Figure 1.27). Neutron SLD F_o-F_c omit maps were generated to determine protonation states and H/D orientation of side-chains. Illustrated are the neutron SLD maps obtained for tyrosine and threonine residues, in which the neutron F_o-F_c maps clearly indicate positive peaks signifying the presence of H/D (Figure 1.28). The collected neutron diffraction data also provided valuable information about multiple protonation states, such as the $-ND_3^+$ group of lysine (Figure 1.29). Refinement statistics (R_{work} and R_{free}) were closely monitored during model optimization to prevent over-fitting. Final statistics gave an X-ray R_{work} of 12.77 % and an R_{free} of 18.21%, and a neutron R_{work} of 14.48 % and an R_{free} of 21.41% with 389 water molecules present.

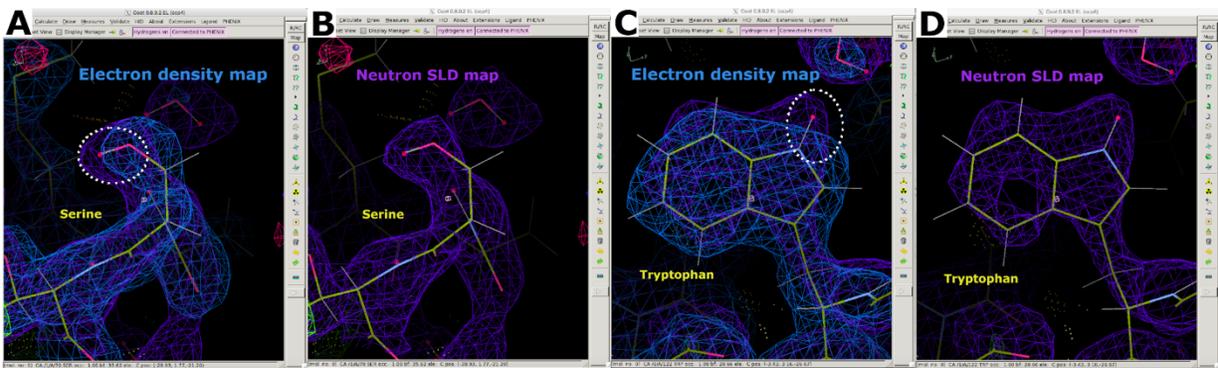


Figure 1.26 Additional information from neutron SLD maps. **(A)** $2F_o-F_c$ electron density map (blue) displays the positions of the “heavier” atoms in serine. **(B)** $2F_o-F_c$ neutron SLD map (purple) clearly displays the position of the “lighter” D atom in serine. **(C)** $2F_o-F_c$ electron density map (blue) displays the positions of the “heavier” atoms in tryptophan. **(D)** $2F_o-F_c$ neutron SLD map (purple) clearly displays the position of the “lighter” D atom in tryptophan.

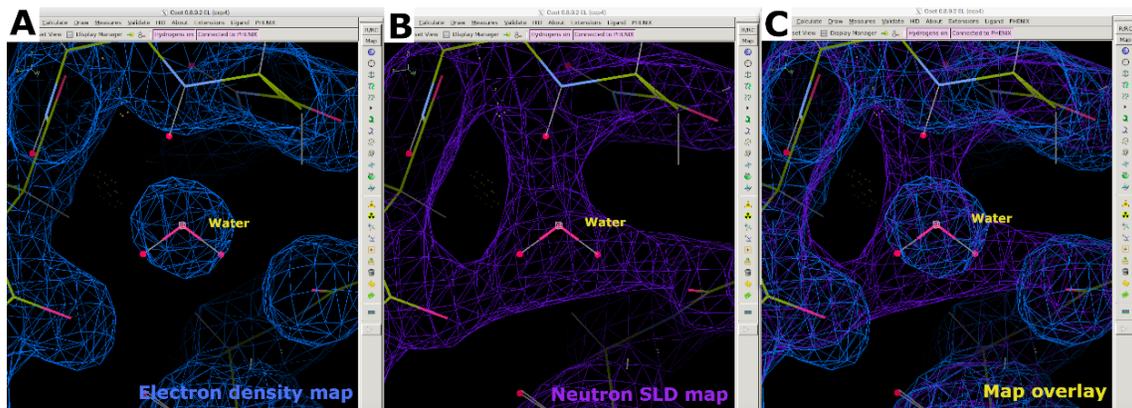


Figure 1.27 Water molecule positioning. **(A)** The spherical shape of an $2F_o-F_c$ electron density map (blue) feature for water. **(B)** The $2F_o-F_c$ neutron SLD map (purple) provides information about the water orientation and hydrogen bond interaction. **(C)** Map overlay of electron and neutron SLD maps of water. The $2F_o-F_c$ neutron SLD map is displayed in purple and $2F_o-F_c$ electron density map is displayed in blue.

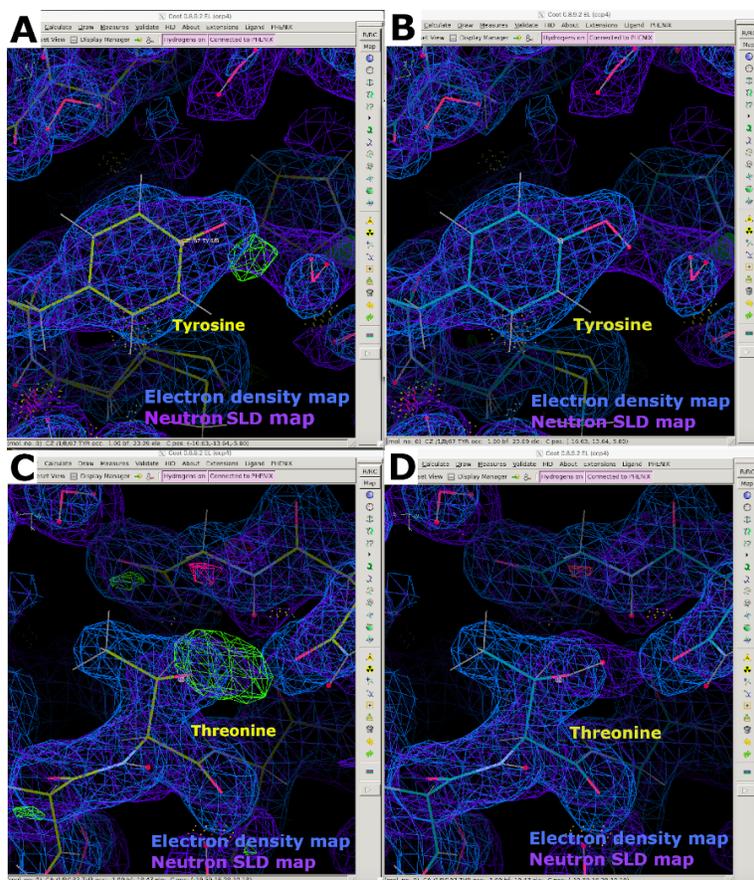


Figure 1.28 Neutron SLD $F_o - F_c$ omit maps. **(A)** The $F_o - F_c$ neutron SLD map (green) provides clear information on the H/D orientation of tyrosine residues. The $2F_o - F_c$ neutron SLD map is displayed in purple and $2F_o - F_c$ electron density map is displayed in blue. **(B)** Tyrosine residue with correct H/D orientation. **(C)** $F_o - F_c$ neutron SLD map (green) provides clear information on the H/D orientation of threonine residues. **(D)** Threonine residue with correct H/D orientation

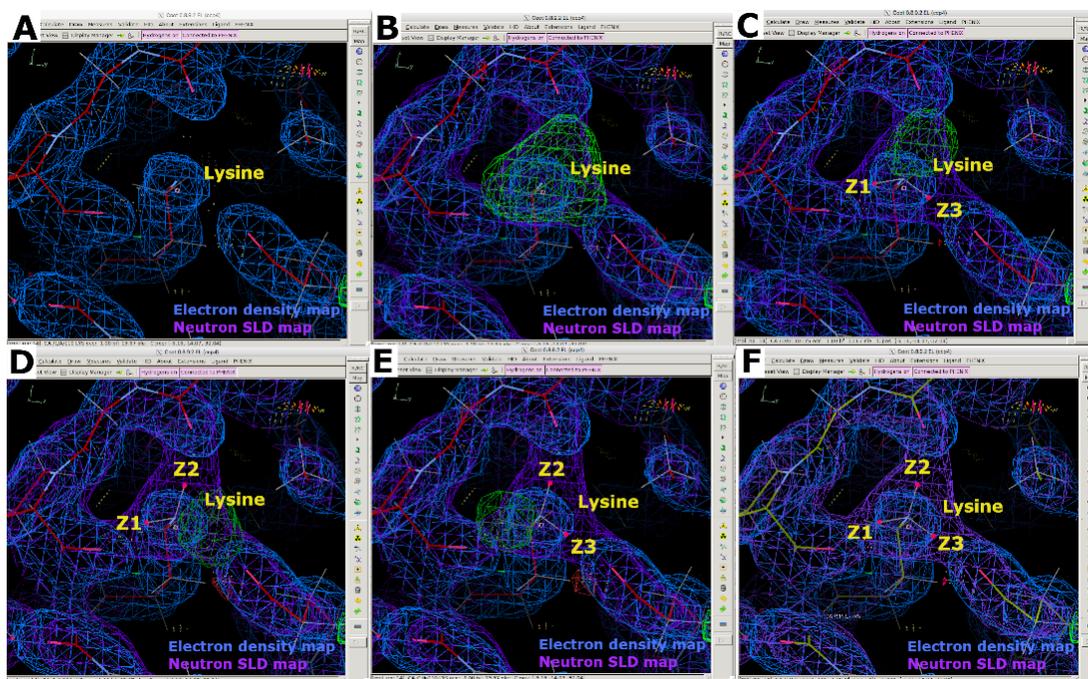


Figure 1.29 Multiple protonation states displayed with neutron SLD maps. **(A)** The $2F_o-F_c$ electron density map (blue) only provides the position of the N atom of lysine ϵ -ammonium group. **(B-E)** The F_o-F_c neutron SLD omit map (green) clearly demonstrates the positively charged $-NH_3$ group. The $2F_o-F_c$ neutron SLD map is displayed in purple and $2F_o-F_c$ electron density map is displayed in blue. **(F)** Overlay of electron density and neutron SLD maps.

Cryo-temperature data were collected on *NcLPMO9D* following an ascorbate soak to reduce the copper active site from Cu^{II} to Cu^I on the MaNDi beamline (Figure 1.22)⁴⁶. Data were collected using TOF Laue mode following a neutron diffraction test using a 4 hour exposure to verify the quality of diffraction. Given the space group of the crystal, a data collection strategy of 18 frames with a collection dose of 80 Coulombs per frame was devised. The data were collected in TOF-Laue mode at a wavelength range of 2.0 – 4.0 Å. Following data collection, the data were indexed, integrated, scaled and merged to give a neutron SLD file in MTZ format at a resolution of 2.40 Å^{48,49}.

Following data collection, the 2.40 Å cryo-temperature *NcLPMO9D* neutron diffraction dataset was used for neutron-only data refinement. The neutron data were phased by molecular replacement using PDB 5TKH as the starting model. Phenix ReadySet was used to add H atoms at non-exchangeable sites and H/D atoms with partial occupancies at exchangeable sites. Water molecules were removed from the starting

model with PDB Tools. Model preparation was followed by refinement with *phenix.refine* using the neutron scattering table. Interactive model building was performed in Coot, with water molecules being added using the positive peaks of the F_o-F_c map and positioned according to potential hydrogen bond interactions (Figure 1.30 A and Figure 1.30 B). When analyzing neutron SLD maps, water molecules are clearly visible if they are highly ordered, however their density may be spherical or ellipsoidal if they are not well-ordered (Figure 1.30 C-E).

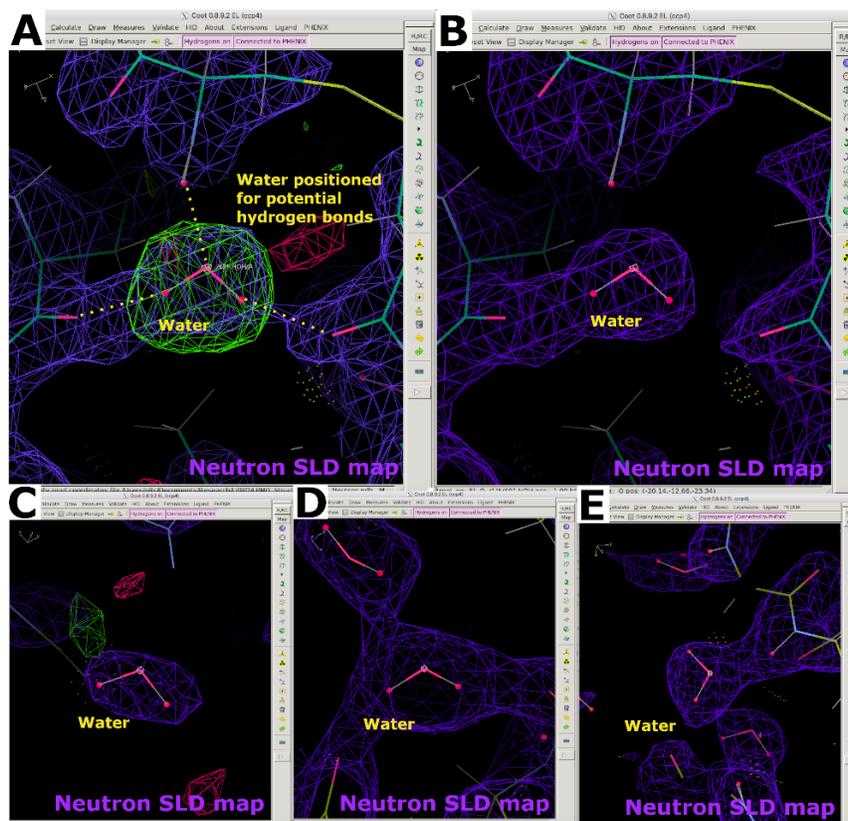


Figure 1.30 Appearance of water molecules in neutron SLD maps. (A) Water molecules are positioned according to F_o-F_c neutron SLD maps (green) and potential hydrogen bonds. The $2F_o-F_c$ neutron SLD map is displayed in purple. (B) Correctly positioned water molecule. (C-E) The various shapes of neutron SLD maps for water molecules depending on B-factors and hydrogen bond interactions.

Neutron SLD maps were used to provide valuable information on the orientation of residues such as asparagine, in which differentiating between the carbonyl and amino groups can be challenging when using X-ray diffraction data alone (Figure 1.31 A and Figure 1.31 B). Peaks in F_o-F_c neutron SLD omit maps were also very informative in determining the protonation states of histidine residues at the N_δ - or N_ϵ -

position (Figure 1.31 C and 1.31 D). The protonation state of residues with multiple H/D exchangeable sites can also be determined using neutron SLD maps. This was clearly illustrated with an F_o-F_c neutron SLD omit map of arginine, which is known to have a positive charge (Figure 1.31 E and 1.31 F). As previously, over-fitting was prevented by monitoring R_{work} and R_{free} . Final statistics gave an R_{work} of 22.58% and an R_{free} of 30.84%. Given that neutron protein diffraction is a flux limited technique in which the negative scattering length and large incoherent scattering factor of H must be taken into account, it can be expected that a neutron data-only refinement would have poorer statistics than a joint X-ray/neutron-data refinement with fewer visible water molecules.

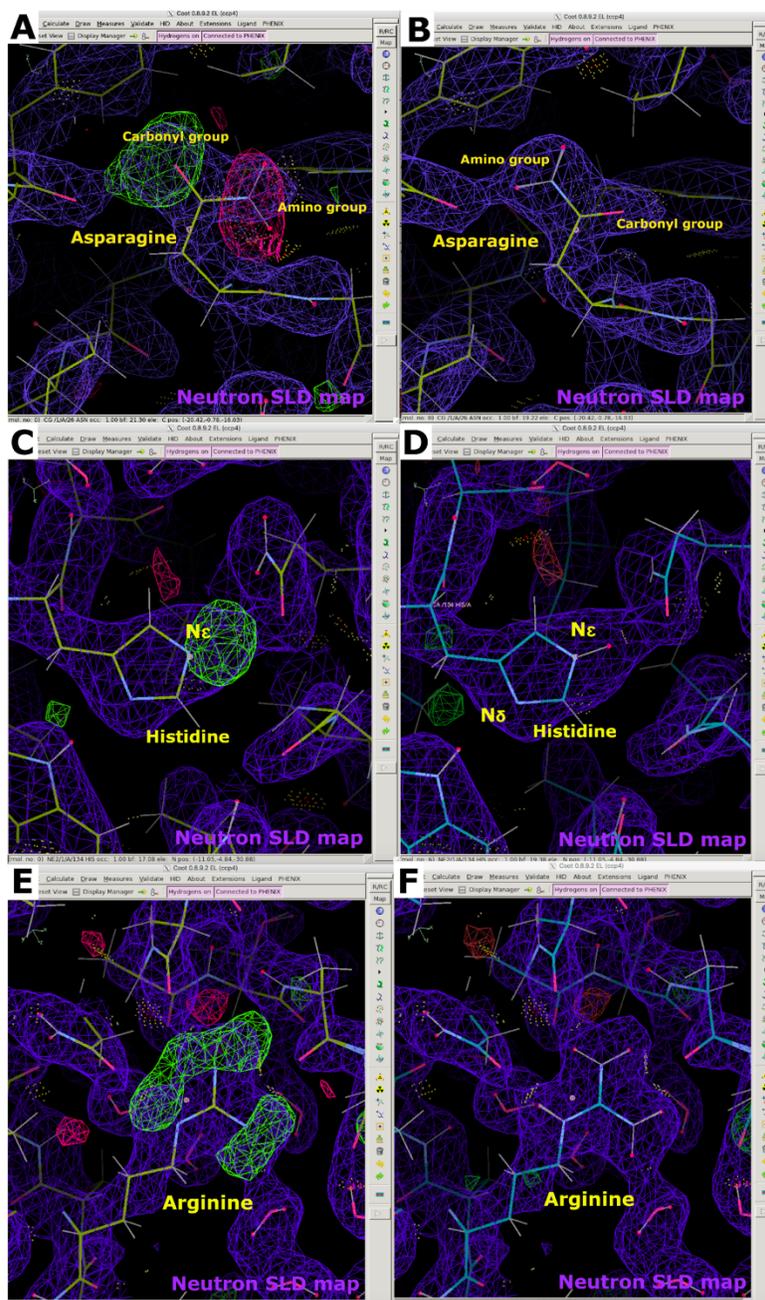


Figure 1.31 Information about amino acid orientation and protonation provided by neutron SLD maps. **(A)** The neutron SLD $F_o - F_c$ map peaks (green) indicate incorrect orientation of an asparagine residue. The $2F_o - F_c$ neutron SLD map is displayed in purple and $2F_o - F_c$ electron density map is displayed in blue. **(B)** $2F_o - F_c$ neutron SLD map (purple) of the correct asparagine orientation. **(C)** The neutron SLD $F_o - F_c$ map peak (green) indicates single protonation of the histidine at N ϵ . **(D)** $2F_o - F_c$ neutron SLD map (purple) of histidine N ϵ -protonation. **(E)** Neutron SLD $F_o - F_c$ omit map peaks (green) confirm the positive charge of arginine. **(F)** $2F_o - F_c$ neutron SLD map (purple) of positively charged arginine.

When analyzing neutron SLD maps, it will become apparent that density cancellation due to the negative neutron scattering length of H will occur for hydrogenated proteins that were subjected to vapor exchange with D₂O-containing crystallization buffer. Due to this reason, neutron SLD maps in which non-exchangeable H atoms are attached to carbon appear incomplete when compared to their electron density map counterpart (Figure 1.32 A). The effect of cancellation is often more apparent at poorer resolutions, making it imperative to obtain protein crystals of a high quality. It is therefore preferable to perform a joint refinement of a sample with both X-ray and neutron data in which the X-ray data can be used to determine the position of the protein backbone (Figure 1.32 B). Furthermore, sulfur atoms in cysteine and methionine may be poorly visible, requiring X-ray data for exact atom placement (Figure 1.32 C and 1.32 D). Metals with weak neutron scattering lengths may also be challenging to model in neutron SLD maps, as is apparent in our LPMO9D maps. Collection of a low dose X-ray dataset on the same crystal is therefore useful, since it permits metal atom positioning using electron density maps (Figure 1.32 E and 1.32 F).

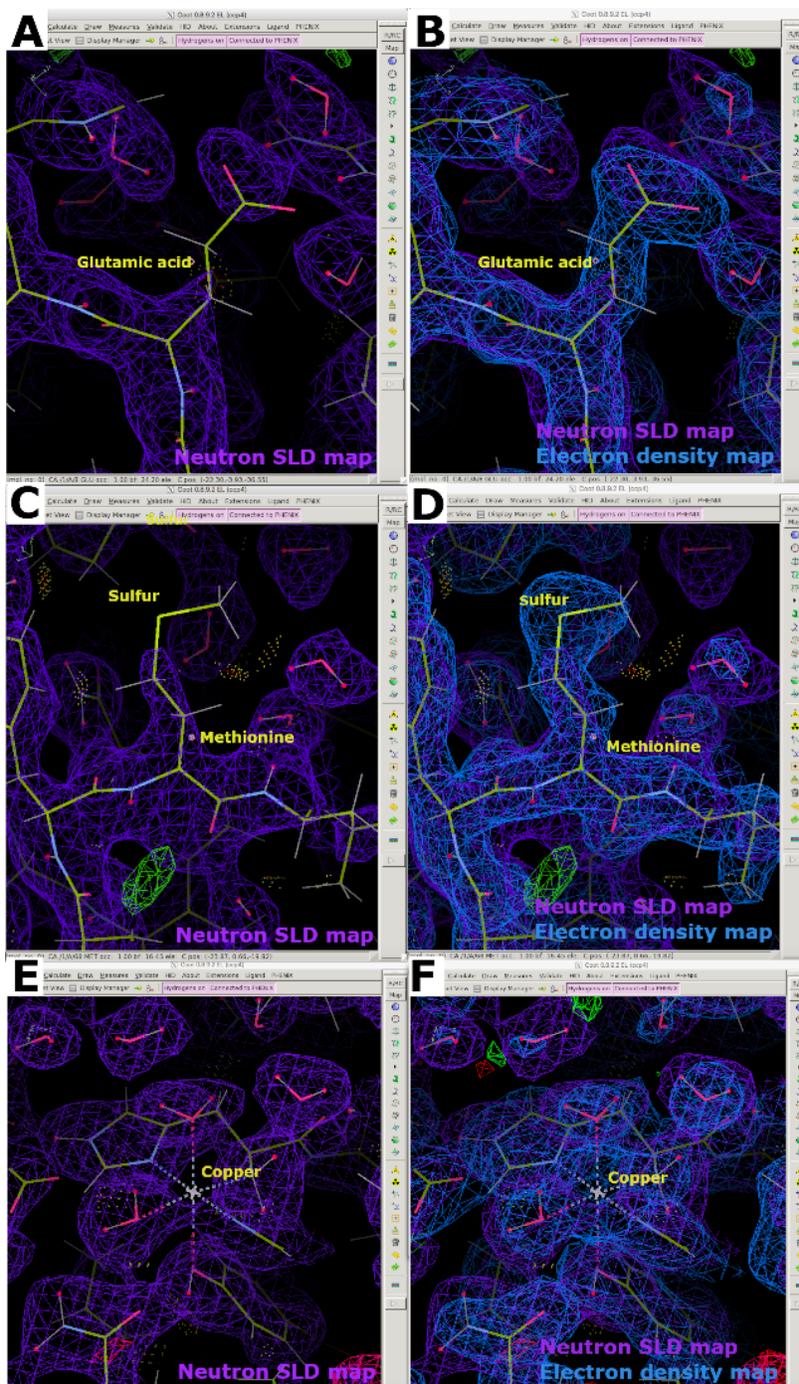


Figure 1.32 Discontinuous neutron SLD maps. (A) $2F_o - F_c$ neutron SLD map (purple) of a hydrogenated, vapor H/D exchanged protein. Glutamic acid displays neutron SLD map cancellation due to the negative scattering length of non-exchangeable H atoms. (B) An overlaid $2F_o - F_c$ electron density map (blue) clearly displays the density of the glutamic acid. (C) Sulfur atom in methionine is poorly visible in $2F_o - F_c$ neutron SLD maps (purple). (D) An overlaid electron density map clearly displays the density of the methionine. (E) Metal atoms, here copper, are poorly visible in neutron $2F_o - F_c$ SLD maps (purple). (F) An overlaid $2F_o - F_c$ electron density map (blue) clearly displays the density of the coordinated copper atom.

1.3.6 Discussion

Neutron protein crystallography is a highly sensitive technique to probe protonation states and water molecule orientation in proteins. This information sheds light on protein catalytic mechanisms since changes in protonation and hydrogen bonding interactions are often central to enzyme chemistry¹⁰. Neutron protein crystallography, albeit an informative technique, has a number of factors that should be taken into consideration before planning to conduct a neutron diffraction experiment, namely:

1. The requirement for large protein crystals for data collection.
2. The scattering properties of hydrogen and other elements, such as metal ions.
3. Limitations in the structure refinement and model building software when working with deuterated samples.

Neutron protein crystallography is a flux limited technique. In contrast to X-ray diffraction datasets, higher R-factors and lower completeness, redundancy and signal-to-noise ratios are expected for neutron datasets due to the technique inherent limitations (flux limited, quasi-Laue, longer wavelengths). Data collection of a single frame is typically 12 – 18 hours. Success of an experiment is highly dependent on sample size and quality with crystals of 0.1 mm³ often being the minimum requirement³. Neutron diffraction requires production of large amounts of protein to set up crystallization drops ranging from 10 to 800 µL. The minimum volume for growing sufficiently large crystals can be estimated using a Volume Calculator given the crystal and sample parameters (<https://neutrons.ornl.gov/imagine>). Growth of large crystals has most prevalently been accomplished by vapor diffusion³. Hanging drop crystallization permits growth of crystals in large drops ranging from 10-25 µL, while larger drops ranging up to ~ 50µL can be set up using commercially available sitting drop equipment^{15,50}. Siliconized nine-well glass plates can be used to set up very large drops, with volumes up to 800 µL. These glass plates are placed in “sandwich boxes” commercially available from Hampton Research (<https://hamptonresearch.com/>). Further crystallization

techniques include batch crystallization, in which the limit of the drop size is dictated by the vessel. Batch crystallization experiment set up can range from microliters to milliliters⁵¹. Crystallization can also be performed using the dialysis technique in which the protein is equilibrated with the precipitant *via* a dialysis membrane or by counter-diffusion along a precipitant concentration gradient or through a porous plug such as agarose^{52,53}. Seeding offers another alternative to obtain crystals of the desired volume. Micro- and macroseeding have been successfully employed for large crystal growth, including large crystal of *NcLPMO9D*⁴⁶. Some knowledge of the protein phase diagram, including the influence of temperature on solubility, aid in large crystal growth.

When planning a neutron diffraction experiment, optimization of the protein preparation to maximize signal-to-noise ratio during diffraction data collection is essential⁷. To circumvent density cancellation and high incoherent scattering caused by H atoms, neutron SLD maps can be improved by exchanging H atoms for its isotope D, which possesses a positive coherent scattering length and low incoherent scattering length. To accomplish this, vapor exchange of the hydrogenated protein crystal against deuterated crystallization buffer is performed. This ensures H/D exchange of solvent molecules and the labile, titratable protein H atoms²⁴. Vapor exchange is performed by mounting the hydrogenated crystal in a quartz capillary with D₂O-based deuterated crystallization buffer “plugs” and it represents an effective, gentle technique that is most often applied^{15,24,36}. The exchange can take several weeks and preferably requires the deuterated buffer to be frequently changed to ensure maximum H/D exchange. H/D exchange can also be performed by directly soaking the crystal in deuterated buffer. To avoid placing the crystal under stress due to D₂O exposure, the soaking process should be performed gradually by incrementally increasing the D₂O:H₂O ratio⁵⁴. In addition to this, crystallization of hydrogenated protein can also be performed in deuterated buffer for H/D exchange at labile H sites^{23,55}. It should be noted, however, that D₂O-based buffer has an effect on protein solubility requiring further adjustment of the known H₂O-based conditions^{3,55}. D₂O-based buffers have also been observed to lead to smaller crystals in some cases⁵⁵. Full exchange of titratable and carbon-bound H atoms to D can be achieved by expressing proteins in deuterated

media to generate a perdeuterated sample²¹. The resulting neutron SLD maps of the perdeuterated sample will be significantly improved, no longer displaying the density cancellation of the hydrogenated sample counterpart. This is beneficial when characterizing H/D bound at non-exchangeable sites in a protein or cofactor. However, expression of perdeuterated protein is both high in cost and low in yield⁵⁶. The Oak Ridge National Laboratory (ORNL) Center for Structural Molecular Biology (CSMB) offers a deuteration facility for users seeking to generate a perdeuterated sample (<https://www.ornl.gov/facility/csmb>). Perdeuterated expression is typically performed in a bioreactor on the 1 L scale yielding ~50 mg of purified protein⁵⁷.

Following the collection of neutron diffraction data, refinement and interactive model building is performed. Refinement can be run using multiple software suites including *phenix.refine*, nCNS or SHELXL^{29,32-34}. The Phenix suite is the most commonly utilized software for refinement of neutron diffraction data in conjunction with Coot which is used to manually build the model from the neutron SLD maps³⁵. Although both Phenix and Coot allow for the processing of neutron diffraction data, they may lack certain features necessary to process the idiosyncrasies associated with neutron data and deuterated samples. For example, Coot does not contain geometry optimization for deuterated residues, which can lead to complications during model building since the “Real Space Refine” feature results in “exploding” residues⁵⁸. This can be resolved by generating restraint files for all the deuterated residues. However, this is an intensive process and such libraries are not currently publicly available. When performing refinements in Phenix, exchangeable H/D sites will initially be set to 0.50 occupancy for H and D. As refinements are performed, the occupancy of H and D will be refined according to the neutron SLD maps. During interactive model building, difference density F_o-F_c maps are very informative in assessing H/D occupancies. Maps can be used to determine which sites possess high D occupancy, which is particularly informative at the active site where protonation states are catalytically relevant⁵⁹. Ambiguous situations do arise, however, when the H:D occupancy is close to 0.70:0.30 which results in complete signal cancellation in neutron SLD maps⁶⁰. It should also be taken into account that quasi-Laue neutron datasets often have a completeness of

~80%, which is lower than the routinely observed $\geq 98\%$ for X-ray diffraction data. When refining neutron diffraction data in Phenix, the missing observed amplitudes (F_o) are therefore calculated from the model to complete the reflection list, thus introducing model bias. To account for this potential bias “no_fill” maps should be examined during interactive model building as opposed to “fill” maps.

Users can choose to perform a joint X-ray/neutron data refinement of their structure, or a neutron-data only refinement. Visualizing neutron SLD maps, particularly at lower resolution, may initially be disconcerting especially for a hydrogenated protein in which H is still present at non-exchangeable sites despite H/D vapor exchange. This results in neutron density map cancellations, giving the impression of discontinuous maps^{61,62}. Collecting a corresponding X-ray dataset advantageously complements these cancellations in a joint refinement (Figure 1.32 A and 1.32 B). A joint-refinement strategy typically involves refining the protein backbone coordinates against the X-ray data, while the neutron diffraction data is used to refine the position and occupancy of the H/D atoms at exchangeable sites²⁹. Since introduction of joint H/D occupancy at exchangeable sites increases the number of parameters being refined, a joint refinement with X-ray data also increases the data-to-parameter ratio. A joint refinement requires a corresponding X-ray dataset to be collected at the same temperature on the same crystal or a crystal grown under the same conditions. For neutron diffraction data collected at room temperature (300K), the corresponding X-ray dataset should be collected at room temperature using a low-dose data collection strategy to limit radiation damage. Perdeuterated samples, in contrast, provide improved and continuous neutron SLD maps since they do not possess the same magnitude of H/D signal cancellation. However, the neutron scattering length of certain elements including metals and sulfur make them poorly visible in neutron SLD maps, even if the protein has been perdeuterated (Figure 1.32 C-F)¹⁹. If a metal needs to be characterized, it is best to utilize X-ray diffraction in a joint refinement or apply spectroscopic techniques to complement diffraction experiments. Neutron-only data refinements are often performed when the neutron dataset has high resolution or if a perdeuterated protein was used. In addition, neutron-only data refinement is particularly useful if a protein highly sensitive to radiation damage is being studied, since an

X-ray derived structure may possess radiation-induced artefacts. If a neutron-data-only refinement is to be performed, it must be ascertained whether the corresponding neutron dataset has sufficient completeness and resolution.

ORNL offers two facilities for collection of neutron diffraction data: the IMAGINE beamline at the HFIR as well as the MaNDi beamline at the SNS^{37,63}. While both instruments provide effective means for collecting a neutron diffraction dataset employing similar principles, each instrument has unique specifications that should be taken into account when applying for beam time. IMAGINE collects quasi-Laue data and is optimized for room temperature data collection on crystals with unit cells up to ~ 100 Å. MaNDi can be used for the collection of room temperature and cryo-temperature data employing TOF-Laue collection on crystals with unit cells up to ~ 300 Å. Prior to collecting a complete dataset, a test is performed on the crystal to evaluate the quality of the obtained diffraction pattern in which the crystal is exposed to the neutron beam for a single frame. If the crystal is of sufficient quality, a full neutron diffraction dataset will be collected, indexed, integrated, scaled and merged in a process analogous to X-ray data processing. IMAGINE makes use of Lauegen and Lscale and MaNDi utilizes the Mantid package and employs three-dimensional profile fitting^{48,64-68}. Scientists who become users at either of these facilities will be provided with a dataset in MTZ or HKL format for further analysis.

Neutron diffraction is a non-destructive, highly sensitive technique for probing the protonation state and hydrogen bond interactions of biological macromolecules. It is particularly useful for photo-sensitive proteins and metalloproteins. Several considerations regarding the technique as well as the processing of the data must be taken into account before conducting an experiment, however the outcome yields results which may give valuable insight into the catalytic mechanism of the protein of interest. Neutron protein crystallography complements computational, structural, biochemical and spectroscopic studies, making it a valuable tool in the biologist's toolbox of techniques used to characterize biological macromolecules.

1.3.7 Acknowledgements

Protein expression, purification and crystallization experiments were conducted at the Center for Structural Molecular Biology (CSMB), a U.S. Department of Energy Biological and Environmental Research User Facility at Oak Ridge National Laboratory. Neutron diffraction data was collected at BL-11B MaNDi at the Spallation Neutron Source (SNS) at ORNL which is sponsored by the Scientific User Facilities Division, Office of Basic Energy Sciences, U.S. Department of Energy. The authors thank Brendan Sullivan for assistance with data reduction. X-ray diffraction data was collected at the Molecular Education, Technology, and Research Innovation Center (METRIC) facilities at North Carolina State University, which is supported by the State of North Carolina. GCS acknowledges support in part from the National Research Foundation (NRF), South Africa and the Graduate Opportunities (GO!) program at ORNL. FM acknowledges support from USDA NIFA Hatch 211001.

1.3.8 Disclosures

The authors have nothing to disclose.

1.3.9 References

1. Neumann, P. & Tittmann, K. Marvels of enzyme catalysis at true atomic resolution: distortions, bond elongations, hidden flips, protonation states and atom identities. *Curr. Opin. Struct. Biol.* **29**, 122–133 (2014).
2. Pynn, R. Neutron Scattering—A Non-destructive Microscope for Seeing Inside Matter. in *Neutron Applications in Earth, Energy and Environmental Sciences* (eds. Liang, L., Rinaldi, R. & Schober, H.) 15–36 (Springer US, 2009). doi:10.1007/978-0-387-09416-8_2
3. O'Dell, W. B., Bodenheimer, A. M. & Meilleur, F. Neutron protein crystallography: A complementary tool for locating hydrogens in proteins. *Arch. Biochem. Biophys.* **602**, 48–60 (2016).
4. Niimura, N. & Podjarny, A. *Neutron Protein Crystallography: Hydrogen, Protons, and Hydration in Bio-macromolecules*. (Oxford University Press, 2011). doi:10.1093/acprof:oso/9780199578863.001.0001
5. Blakeley, M. P. P. Neutron macromolecular crystallography. *Crystallogr. Rev.* **15**, 157–218 (2009).
6. Blakeley, M. P., Cianci, M., Helliwell, J. R. & Rizkallah, P. J. Synchrotron and neutron techniques in biological crystallography. *Chem. Soc. Rev.* **33**, 548–557 (2004).
7. Ashkar, R. *et al.* Neutron scattering in the biological sciences: progress and prospects. *Acta Crystallographica Section D: Structural Biology* **74**, (2018).
8. Teixeira, S. C. M. *et al.* New sources and instrumentation for neutrons in biology. *Chem. Phys.* **345**, 133–151 (2008).
9. Furrer, A., Mesot, J. & Strässle, T. *Neutron Scattering in Condensed Matter Physics*. (World Scientific Publishing Company, 2009).
10. Meilleur, F., Coates, L., Cuneo, M. J., Kovalevsky, A. & Myles, D. A. A. The neutron macromolecular crystallography instruments at Oak Ridge national laboratory: Advances, challenges, and opportunities. *Crystals* **8**, 1–10 (2018).
11. Coates, L. *et al.* The Macromolecular Neutron Diffractometer MaNDi at the Spallation Neutron Source. *J. Appl. Crystallogr.* **48**, 1302–1306 (2015).
12. Coates, L., Stoica, A. D., Hoffmann, C., Richards, J. & Cooper, R. The macromolecular neutron diffractometer (MaNDi) at the Spallation Neutron Source, Oak Ridge: enhanced optics design, high-resolution neutron detectors and simulated diffraction. *J. Appl. Crystallogr.* **43**, 570–577 (2010).
13. Schröder, G. C., O'Dell, W. B., Myles, D. A. A., Kovalevsky, A. & Meilleur, F. IMAGINE: Neutrons reveal enzyme chemistry. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, (2018).
14. Koetzle, T. F., Piccoli, P. M. B. & Schultz, A. J. Single-crystal neutron diffraction studies of hydrogen-bonded systems: Two recent examples from IPNS. *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* **600**, 260–262 (2009).
15. Ng, J. D. *et al.* Large-volume protein crystal growth for neutron macromolecular crystallography. *Acta Crystallogr. Sect. F Structural Biol. Commun.* **71**, 358–370 (2015).
16. Blakeley, M. P., Langan, P., Niimura, N. & Podjarny, A. Neutron crystallography: opportunities, challenges, and limitations. *Curr. Opin. Struct. Biol.* **18**, 593–600 (2008).
17. Sears, V. F. Neutron scattering lengths and cross sections. *Neutron News* **3**, 26–37 (1992).
18. Weik, M., Patzelt, H., Zaccari, G. & Oesterhelt, D. Localization of glycolipids in membranes by in vivo labeling and neutron diffraction. *Mol. Cell* **1**, 411–419 (1998).
19. Helliwell, J. R. *Fundamentals of neutron crystallography in structural biology. Methods in Enzymology* **634**, (Elsevier Inc., 2020).
20. Niimura, N. & Bau, R. Neutron protein crystallography: Beyond the folding structure of biological macromolecules. *Acta Crystallogr. Sect. A Found. Crystallogr.* **64**, 12–22 (2008).
21. Hazemann, I. *et al.* High-resolution neutron protein crystallography with radically small crystal volumes: Application of perdeuteration to human aldose reductase. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **61**, 1413–1417 (2005).
22. Niimura, N., Chatake, T., Ostermann, A., Kurihara, K. & Tanaka, I. High resolution neutron protein crystallography. Hydrogen and hydration in proteins. *Zeitschrift für Krist. - Cryst. Mater.* **218**,

- (2003).
23. Meilleur, F., Contzen, J., Myles, D. A. A. & Jung, C. Structural stability and dynamics of hydrogenated and perdeuterated cytochrome P450cam (CYP101). *Biochemistry* **43**, 8744–8753 (2004).
 24. Bennett, B. C., Gardberg, A. S., Blair, M. D. & Dealwis, C. G. On the determinants of amide backbone exchange in proteins: A neutron crystallographic comparative study. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **64**, 764–783 (2008).
 25. Meilleur, F., Kovalevsky, A. & Myles, D. A. A. IMAGINE: The neutron protein crystallography beamline at the high flux isotope reactor. in *Methods in enzymology* **634**, 69–85 (Elsevier Inc., 2020).
 26. Wang, X. P. *et al.* A suite-level review of the neutron single-crystal diffraction instruments at Oak Ridge National Laboratory. *Rev. Sci. Instrum.* **89**, 092802 (2018).
 27. Wlodawer, A. & Hendrickson, W. A. A procedure for joint refinement of macromolecular structures with X-ray and neutron diffraction data from single crystals. *Acta Crystallogr. Sect. A* **38**, 239–247 (1982).
 28. Halle, B. Biomolecular cryocrystallography: Structural changes during flash-cooling. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 4793–4798 (2004).
 29. Adams, P. D., Mustyakimov, M., Afonine, P. V. & Langan, P. Generalized X-ray and neutron crystallographic analysis: More accurate and complete structures for biological macromolecules. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **65**, 567–573 (2009).
 30. Berman, H. M. *et al.* The Protein Data Bank. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **58**, 899–907 (2002).
 31. Liebschner, D., Afonine, P. V., Moriarty, N. W., Langan, P. & Adams, P. D. Evaluation of models determined by neutron diffraction and proposed improvements to their validation and deposition. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 800–813 (2018).
 32. Afonine, P. V. *et al.* Joint X-ray and neutron refinement with phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 1153–1163 (2010).
 33. Brünger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54**, 905–921 (1998).
 34. Gruene, T., Hahn, H. W., Luebben, A. V., Meilleur, F. & Sheldrick, G. M. Refinement of macromolecular structures against neutron data with SHELXL2013. *J. Appl. Crystallogr.* **47**, 462–466 (2014).
 35. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
 36. Lakey, J. H. Neutrons for biologists: A beginner’s guide, or why you should consider using neutrons. *J. R. Soc. Interface* **6**, (2009).
 37. Schröder, G. C., O’Dell, W. B., Myles, D. A. A., Kovalevsky, A. & Meilleur, F. IMAGINE: Neutrons reveal enzyme chemistry. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 778–786 (2018).
 38. Halsted, T. P. *et al.* Catalytically important damage-free structures of a copper nitrite reductase obtained by femtosecond X-ray laser and room-temperature neutron crystallography. *IUCrJ* **6**, 761–772 (2019).
 39. Meier, K. K. *et al.* Oxygen Activation by Cu LPMOs in Recalcitrant Carbohydrate Polysaccharide Conversion to Monomer Sugars. *Chem. Rev.* **118**, 2593–2635 (2018).
 40. Beeson, W. T., Vu, V. V., Span, E. A., Phillips, C. M. & Marletta, M. A. Cellulose Degradation by Polysaccharide Monooxygenases. *Annu. Rev. Biochem.* **84**, 923–946 (2015).
 41. Walton, P. H. & Davies, G. J. On the catalytic mechanisms of lytic polysaccharide monooxygenases. *Curr. Opin. Chem. Biol.* **31**, 195–207 (2016).
 42. Bertini, L. *et al.* Catalytic Mechanism of Fungal Lytic Polysaccharide Monooxygenases Investigated by First-Principles Calculations. *Inorg. Chem.* **57**, 86–97 (2018).
 43. Hedegård, E. D. & Ryde, U. Molecular mechanism of lytic polysaccharide monooxygenases. *Chem. Sci.* **9**, 3866–3880 (2018).

44. Hangasky, J. A., Detomasi, T. C. & Marletta, M. A. Glycosidic Bond Hydroxylation by Polysaccharide Monooxygenases. *Trends Chem.* **1**, 198–209 (2019).
45. Bacik, J. P. *et al.* Neutron and Atomic Resolution X-ray Structures of a Lytic Polysaccharide Monooxygenase Reveal Copper-Mediated Dioxygen Binding and Evidence for N-Terminal Deprotonation. *Biochemistry* **56**, 2529–2532 (2017).
46. O'Dell, W. B., Agarwal, P. K. & Meilleur, F. Oxygen Activation at the Active Site of a Fungal Lytic Polysaccharide Monooxygenase. *Angew. Chemie - Int. Ed.* **56**, 767–770 (2017).
47. O'Dell, W. B., Swartz, P. D., Weiss, K. L. & Meilleur, F. Crystallization of a fungal lytic polysaccharide monooxygenase expressed from glycoengineered *Pichia pastoris* for X-ray and neutron diffraction. *Acta Crystallogr. Sect. Struct. Biol. Commun.* **73**, 70–78 (2017).
48. Sullivan, B. *et al.* Improving the accuracy and resolution of neutron crystallographic data by three-dimensional profile fitting of Bragg peaks in reciprocal space. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 1085–1095 (2018).
49. Sullivan, B. *et al.* BraggNet: integrating Bragg peaks using neural networks. *J. Appl. Crystallogr.* **52**, 854–863 (2019).
50. Blum, M. M. *et al.* Preliminary time-of-flight neutron diffraction study on diisopropyl fluorophosphatase (DFPase) from *Loligo vulgaris*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **63**, 42–45 (2007).
51. Tomanicek, S. J. *et al.* Neutron and X-ray crystal structures of a perdeuterated enzyme inhibitor complex reveal the catalytic proton network of the Toho-1 β -lactamase for the acylation reaction. *J. Biol. Chem.* **288**, 4715–4722 (2013).
52. Metcalfe, C. *et al.* The tuberculosis prodrug isoniazid bound to activating peroxidases. *J. Biol. Chem.* **283**, 6193–6200 (2008).
53. Hughes, R. C. *et al.* Inorganic pyrophosphatase crystals from *Thermococcus thio-reducens* for X-ray and neutron diffraction. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **68**, 1482–1487 (2012).
54. Bennett, B. C., Meilleur, F., Myles, D. A. A., Howell, E. E. & Dealwis, C. G. Preliminary neutron diffraction studies of *Escherichia coli* dihydrofolate reductase bound to the anticancer drug methotrexate. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **61**, 574–579 (2005).
55. Snell, E. H. *et al.* Optimizing crystal volume for neutron diffraction: D-xylose isomerase. *Eur. Biophys. J.* **35**, 621–632 (2006).
56. Golden, E., Attwood, P. V., Duff, A. P., Meilleur, F. & Vrielink, A. Production and characterization of recombinant perdeuterated cholesterol oxidase. *Anal. Biochem.* **485**, 102–108 (2015).
57. Munshi, P. *et al.* Rapid visualization of hydrogen positions in protein neutron crystallographic structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 35–41 (2012).
58. Logan, D. T. Interactive model building in neutron macromolecular crystallography. in *Methods in Enzymology* **634**, 201–224 (Elsevier Inc., 2020).
59. Koruza, K., Lafumat, B., Végvári, Knecht, W. & Fisher, S. Z. Deuteration of human carbonic anhydrase for neutron crystallography: Cell culture media, protein thermostability, and crystallization behavior. *Arch. Biochem. Biophys.* **645**, 26–33 (2018).
60. Fisher, S. J. *et al.* Perdeuteration: Improved visualization of solvent structure in neutron macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **70**, 3266–3272 (2014).
61. Chen, J. C.-H., Hanson, B. L., Fisher, S. Z., Langan, P. & Kovalevsky, a. Y. Direct observation of hydrogen atom dynamics and interactions by ultrahigh resolution neutron protein crystallography. *Proc. Natl. Acad. Sci.* **109**, 15301–15306 (2012).
62. Cuyppers, M. G. *et al.* Near-atomic resolution neutron crystallography on perdeuterated *Pyrococcus furiosus* rubredoxin: Implication of hydronium ions and protonation state equilibria in redox changes. *Angew. Chemie - Int. Ed.* **52**, 1022–1025 (2013).
63. Coates, L. & Sullivan, B. *The macromolecular neutron diffractometer at the spallation neutron source. Methods in Enzymology* **634**, (Elsevier Inc., 2020).
64. Ren, Z., Kingman, N. G., Borgstahl, G. E. O., Getzoff, E. D. & Moffat, K. Quantitative Analysis of

- Time-Resolved Laue Diffraction Patterns. *J. Appl. Crystallogr.* **29**, 246–260 (1996).
65. Campbell, J. W., Hao, Q., Harding, M. M., Nguti, N. D. & Wilkinson, C. LAUEGEN version 6.0 and INTLDM. *J. Appl. Crystallogr.* **31**, 496–502 (1998).
 66. Arnold, O. *et al.* Mantid - Data analysis and visualization package for neutron scattering and μ SR experiments. *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* **764**, 156–166 (2014).
 67. Helliwell, J. R. *et al.* The recording and analysis of synchrotron X-radiation Laue diffraction photographs. *J. Appl. Crystallogr.* **22**, 483–497 (1989).
 68. Arzt, S., Campbell, J. W., Harding, M. M., Hao, Q. & Helliwell, J. R. LSCALE - The new normalization, scaling and absorption correction program in the Daresbury Laue software suite. *J. Appl. Crystallogr.* **32**, 554–562 (1999).

1.4 Overview of research

Metalloproteins possess a broad “chemical toolbox”, catalyzing numerous essential reactions in living organisms. The metal cofactor expands the physiochemical properties of proteins, often playing a central role in activation and stabilization during the catalytic cycle. The research performed in this dissertation sought to further characterize the structure and reaction mechanism of metalloproteins, with particular emphasis on metal-coordinated oxidoreductases. The systems examined in this thesis are the copper center lytic polysaccharide monooxygenases (LPMOs) and the heme-iron center cytochrome P450s. **Chapter 1** provides background on how structural studies, particularly neutron protein crystallography, can be used as a powerful tool to elucidate the metalloprotein reaction mechanism. In addition to this chapter includes a technical overview of performing neutron protein crystallography under room temperature and cryo-conditions.

LPMOs catalyze the insertion of oxygen into the glycosidic bond of recalcitrant crystalline cellulose, and while much research has been performed on the reaction intermediates involved in catalysis, most of this remains theoretical due to the insoluble nature of the polysaccharide substrate. This work aims to provide a structural basis for the theoretically proposed catalytic mechanisms since fundamental aspects of the LPMO reaction mechanism remain poorly understood. **Chapter 2** provides insight into the structural studies performed to characterize the initial steps of dioxygen activation by the fungal *Neurospora crassa* NcLPMO9D as well as the role of second shell residues in catalysis. To characterize the activated dioxygen intermediate, neutron protein crystallography was performed on a reduced NcLPMO9D in which an activated dioxygen species was cryogenically trapped. A further neutron protein crystal structure of NcLPMO9D at low pH conditions characterizes the protonation state of second shell residues under these functionally relevant conditions.

Cytochrome P450s (CYP450) enzymes comprise a superfamily of heme-containing proteins ubiquitous in most life forms that catalyze the highly regio- and stereoselective hydroxylation of the sp^3 -

hybridized carbons in non-activated hydrocarbons. In order to achieve catalysis, cleavage of the O-O bond in molecular dioxygen bond is necessary to form the catalytically active form of CYP450cam, the high-valent iron-oxo species termed Compound I. The cleavage of this O-O bond requires the leaving oxygen to be protonated. In related enzyme systems, such as peroxidases, there are suitably positioned groups in the active site (Arg and His) capable of mediating this function. There are no similar groups found at the active site of CYP450s. Instead, a complicated proton delivery pathway involving active site residues and water molecules is likely to be involved, but the exact nature of the pathway is still debated. In **Chapter 3**, the work performed to purify and crystallize CYP450cam from *Pseudomonas putida* which is involved in camphor hydroxylation is described. The aim of these crystallization trials was to grow crystals of sufficient volume ($> 0.1 \text{ mm}^3$) for studies by neutron protein crystallography. Growth of such crystals would permit determination of the protonation pathway by determination of protonation state, water molecule orientation and hydrogen bond interactions near the active site using neutron protein crystallography.

To achieve catalysis CYP450cam requires the addition of molecular oxygen and two electrons obtained from reduced cofactors. The first reduction of CYP450cam can be achieved by non-native redox partners including small molecules, however Putidaredoxin (Pdx), a Fe_2S_2 ferredoxin redox partner protein, is essential for the second reduction. Pdx binds to CYP450cam on the proximal side of the heme and is believed to play an important effector role, inducing conformational changes essential for substrate hydroxylation. **Chapter 4** examines the effector role of Pdx using small angle X-ray scattering (SAXS). This chapter additionally studies the conformational state of substrate-free CYP450cam and conformational changes induced by binding of the substrate camphor using SAXS.

Chapter 5 provides a discussion of the insights gained from the structural characterization of LPMOs and CYP450s using X-ray and neutron scattering, how this can be placed into context with current knowledge on these systems and future avenues of research.

Chapter 2 : Elucidating the reaction mechanism of lytic polysaccharide monooxygenases with neutron protein crystallography

2.1 Introduction to catalysis by lytic polysaccharide monooxygenases

2.1.1 Lignocellulosic biomass

Biomass is composed of photosynthetically derived organic materials, encompassing wood, grass, crops, algae and animal waste ¹. Given the production of biomass from sunlight, water and CO₂, it is inherently renewable and represents a crucial energy source with additional applications for the production of fine chemicals ². Biomass is composed of a large number of organic molecules such as lipids, proteins, nucleotides and steroids, however it is carbohydrates in particular that serve as an energy repository ³. Carbohydrates also play an important structural role by forming polymers from cellulose and hemicellulose in plant cell walls ⁴. The cellulose and hemicellulose polymers associate with lignin, a polymer made of phenylpropanoid subunits, to form lignocellulosic biomass, the most abundant form of biomass on earth ⁵. Lignocellulose is highly organized into a hierarchical structure highly recalcitrant to degradation (Figure 2.1) ⁶. Lignocellulose is composed primarily of cellulose, a linear homopolymer made up of D-glucopyranose linked by β -(1,4)-glycosidic bonds, with the “true” repeating unit represented by the disaccharide cellobiose composed of two glucose units (1- β -D-glucopyranosyl-4- β -D-glucopyranose) ⁷. The linear cellulose strands assemble into microfibrils with a very stable crystalline core held together by inter- and intramolecular hydrogen bonds as well as Van der Waals interactions ⁸. Hemicellulose, a polysaccharide that forms irregular and amorphous structures, is the second component of the microfibrils in lignocellulose ⁹. The short-chain branched architecture of hemicellulose results in a lower molecular weight than cellulose. It is composed of the pentoses D-xylose and L-arabinose, the hexoses D-glucose, D-mannose, D-galactose, D-rhamnose as well as 4-O-methyl-glucuronic, D-galacturonic, and D-glucuronic acids which are linked by β -1,3- and β -1,4-glycosidic bonds ¹⁰. The hemicellulose heteropolymers form a complex network that intertwines with the cellulose microfibrils, providing structural stability ¹¹. Lignin is the final constituent of lignocellulose and provides a structural support for the hemicellulose and cellulose

microfibrils as well as forming a protective barrier¹². Lignin forms a three dimensional heteropolymer that acts as a “glue” by crosslinking with hemicellulose via its phenylpropanoid units¹³. Lignin is composed of three phenolic components: *p*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol which form *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S), respectively during polymerization¹⁴.

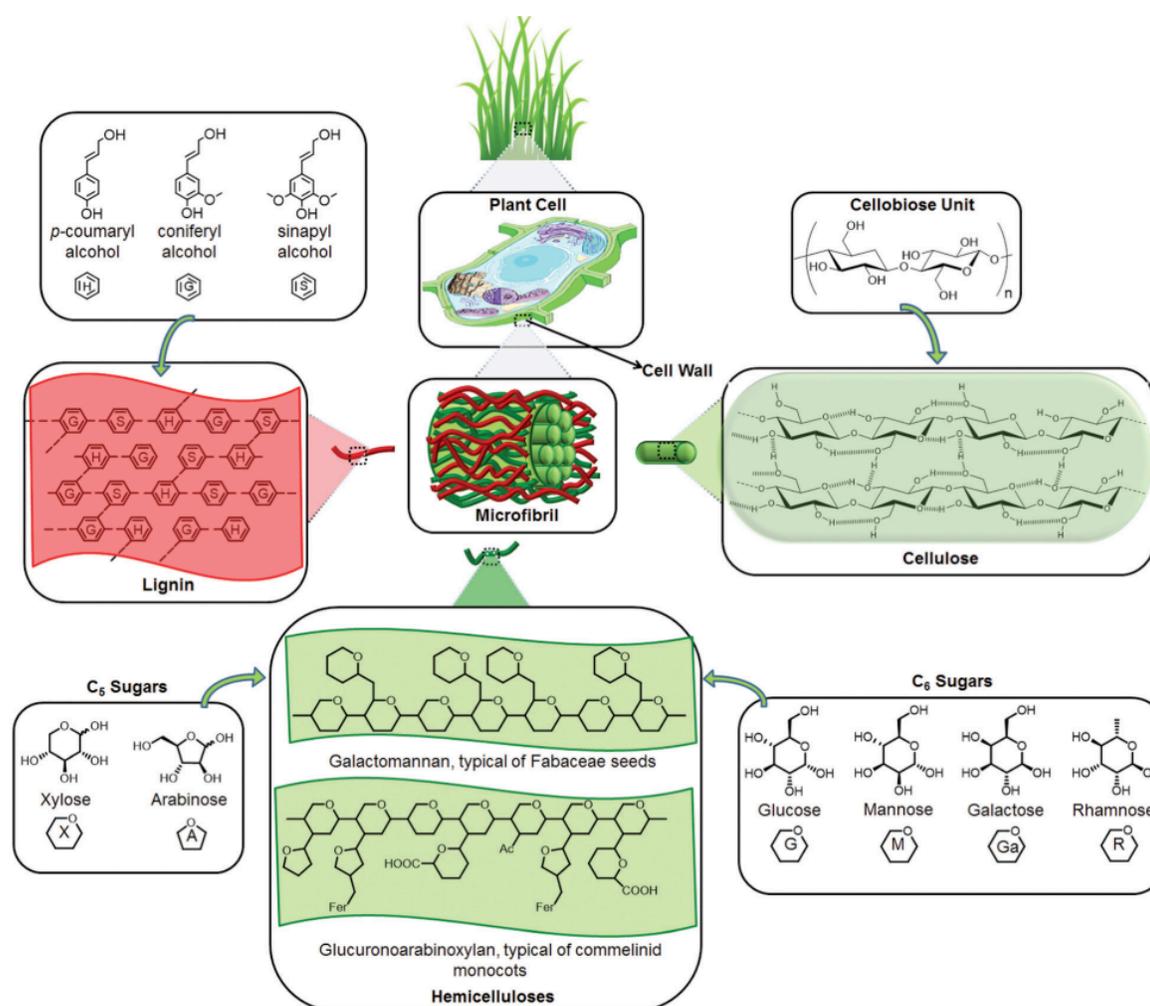


Figure 2.1 The hierarchical structure of lignocellulose. The plant cell wall is composed of three major components: cellulose, hemicellulose and lignin. The cellulose forms microfibrils that are embedded in a lignin and hemicellulose matrix. Within the microfibrils, there are crystalline and amorphous sections of cellulose. Cellulose is composed repeating cellobiose units which consists of two glucose units (1-β-D-glucopyranosyl-4-β-D-glucopyranose). Hemicellulose is primarily composed of the five-carbon pentoses D-xylose and L-arabinose and the six-carbon hexoses D-glucose, D-mannose, D-galactose, D-rhamnose. Lignin is composed of the phenols *p*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol. Figure adapted from Isikgor *et al.*⁶ and reprinted with permission of the Royal Society of Chemistry.

Lignocellulose has a unique composition higher in oxygen and lower in hydrogen and carbon than fossil fuel derived petroleum making it a multi-faceted precursor for the production of value-added compounds¹⁵. As non-renewable fuel resources from petroleum become limited and demand continues to increase, however, the use of renewable biomass such as lignocellulose for biofuel production is gaining increasing importance^{16,17}. In order to harness the potential of lignocellulose for fuel and fine chemical production, it is necessary to break it down into its primary constituents: cellulose, hemicellulose and lignin¹⁸. Deconstruction of lignocellulose is a multistep process given its inherent recalcitrance due to its complex hierarchical structure that involves pretreatment via mechanical, chemical, physicochemical or biological methods before further conversion and processing methods to obtain biofuels and valued-added compounds¹⁹. Pretreatments are designed to disrupt the lignocellulose structure, reduce crystallinity and polymer size, allowing fractionation and solubilization of the cellulose, hemicellulose and lignin using heat, milling, acid hydrolysis, steam, microwave irradiation and ammonia fiber expansion amongst others²⁰. Additionally, pretreatment can also be accomplished by biological methods, particularly with the use of ligninolytic enzymes that depolymerize lignin to yield holocellulose^{21,22}. Following pretreatment, cellulose, hemicellulose and lignin act as feedstocks that can undergo conversion for downstream applications^{23,24}. Production of cellulosic ethanol from the fractionated products of lignocellulose pretreatment requires saccharification of cellulose to yield its monosaccharide components for fermentation²⁵. The cellulose glycosidic bond can be disrupted chemically by acid-catalyzed hydrolysis. However acid catalysis has the drawback of resulting in incomplete hydrolysis and producing unwanted furan derivatives²⁶. Enzymatic hydrolysis of cellulose, in contrast, is highly specific thereby avoiding undesirable side-reactions and can be performed under comparatively mild conditions²⁷. Cellulases (EC 3.2.1.4) are classified as O-glycoside hydrolases (GH) according to the Carbohydrate-Active enZYme Database (CAZy) that classifies enzymes according to their structural similarity²⁸. GH catalyze hydrolysis of β -1,4-glycosidic bonds in cellulose and cellulases of both fungal and bacterial origin have biotechnological applications in the production of soluble sugars²⁹. A forerunner in biofuel applications has been the fungus *Trichoderma reesei*, whose secreted cellulases have been widely used in industrial biorefineries³⁰, however several strains of *Penicillium spp.*,

Aspergillus spp. and *Clostridium* spp. have also emerged as alternatives for biofuel production^{31,32}. Industrial biofuel production from cellulose makes use of a cellulolytic cocktail of enzymes to maximize cellulose deconstruction³³. Fungi are the preferred source of enzymes for cellulolytic cocktails, since they secrete the large amounts of cellulases extracellularly, making them ideal for industrial applications, while bacterial secreted cellulases are limited and remain associated with the cell wall³⁴.

2.1.2 Fungal enzymatic cellulose breakdown

Fungi secrete a number of cellulases that function simultaneously and synergistically for the breakdown of cellulose³⁵. The primary cellulases secreted by fungi include endoglucanases (EG) that act on amorphous cellulose, cellobiohydrolases (CBH) that act on cellulose chain ends and β -glucosidases (β -G) that breaks down cellobiose into D-glucose (Figure 2.2)³⁶⁻³⁸. EGs (GH5, GH6, GH7, GH9, GH12, GH45 and GH74) act on the amorphous regions of cellulose by randomly cleaving internal β -1,4-glycosidic bonds and releasing fragments with reducing and non-reducing chain ends³⁹. CBHs (GH6 and GH7) are also known as exoglucanases and act on the cellulose chain ends to release cellobiose subunits⁴⁰. CBH are classified into two classes: CBHI that acts on reducing ends and CBHII that acts on non-reducing ends⁴¹. β -Gs (GH1 and GH3) hydrolyze cellobiose and soluble gluco-oligosaccharides to D-glucose⁴². These cellulases function in concert to breakdown cellulose, and the synergism between different components including endoglucanases (endo-exo), cellobiohydrolases (exo-exo), cellobiohydrolases and β -glucosidases has been studied⁴³⁻⁴⁵.

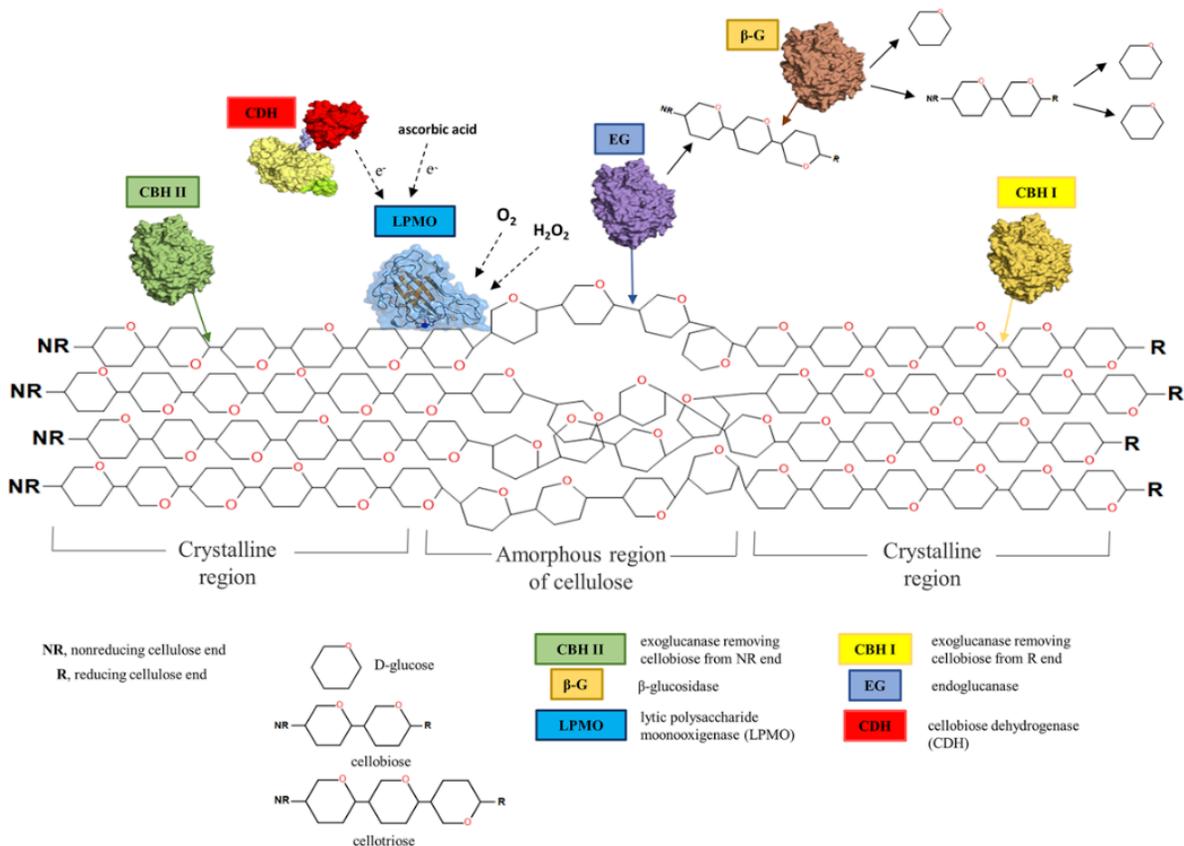


Figure 2.2 Enzymes involved in the degradation of cellulose. Fungi contain a host of enzymes to break down the cellulose component of microfibrils. Crystalline cellulose composed of repeating cellobiose. Exoglucanases such as cellobiohydrolases (CBHI – reducing-end active and CBH II – non-reducing end active) act on chain ends to generate soluble sugars. Endoglucanases (EG) increase the accessibility of cellulose chain ends to exoglucanases by creating new chain ends from internal regions of amorphous cellulose. β-glucosidases (β-G) break down cellobiose into D-glucose. Lytic polysaccharide monooxygenases (LPMO) oxidatively cleaves the glycosidic bond with cellobiose dehydrogenase providing reducing equivalents for LPMO catalysis. Adapted from Andlar *et al.*⁴⁶ and reprinted with permission of the Engineering in Life Sciences Journal.

Although these GH are able to degrade cellulase, a significant hurdle remained for the efficient degradation of crystalline cellulose, which due to its regular packing and strong hydrogen bond interactions is highly recalcitrant to depolymerization, leading to postulation of an external factor to assist in making crystalline cellulose more accessible⁴⁷. There was a new dawn in enzymatic degradation of lignocellulose, when it was found that an enzyme was able to act on the glycosidic bond via an oxidative mechanism, thereby disrupting cellulose crystallinity and making it more accessible⁴⁸. These lignocellulose degradation-enhancing enzymes were originally classified as GH61 in the CAZy database due to a low

level of endoglucanase activity^{49,50}. A structure by Harris *et al.* on a GH61 from *Thielavia terrestris* showed that this enzyme enhanced cellulose breakdown by *T. reesei* and the X-ray crystal structure revealed a fold significantly different to the canonical GH family, with a Zn²⁺ or Mg²⁺ coordinated to histidine residues (Figure 2.3)⁵¹. It was found that the stimulated cellulose deconstruction was eliminated in the presence of a chelating agent, while mutagenesis of residues near the coordinating metal negatively affected cellulose degradation.

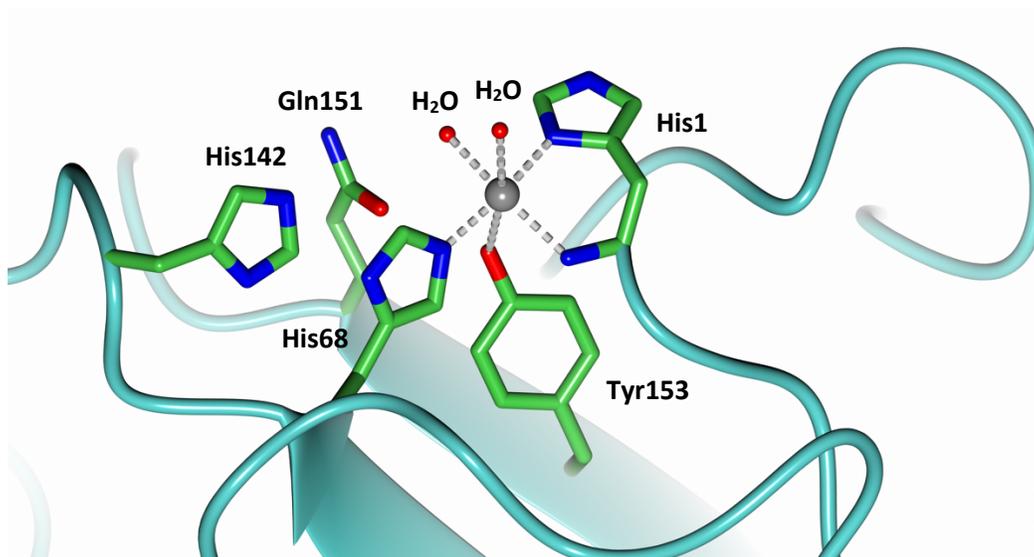


Figure 2.3 Structure of *TtGH61E* with Zn²⁺ modelled in the active site (PDB 3EJJ). Harris *et al.* found that mutation of His1, His68 and Gln151 eliminated cellulose deconstruction while the rate of cellulose deconstruction was decreased by mutation of His142 and Tyr153⁵¹.

Studies on chitin breakdown by *Serratia marcescens* SmAA10A by Vaaje-Kolstad *et al.*, subsequently found that a GH61-like protein was responsible for the oxidative breakdown of crystalline chitin⁵². This oxygenase activity resulted in chain disruption by formation of oxidized ends which were then accessible to chitinases for further breakdown. Quinlan *et al.* demonstrated that a GH61 from *Thermoascus aurantiacus* performed oxidative cellulose chain cleavage and demonstrated that GH61 enzymes are copper binding enzymes by performing metal binding assays with a range of metal ions including Ca²⁺, Co²⁺, Cu²⁺, Mg²⁺, Mn²⁺ and Zn²⁺⁵³. It was postulated that the discrepancy of GH61 activity when binding to Mg²⁺ and Zn²⁺ could be attributed to the fact that residual amounts of Cu²⁺ remained,

enabling catalysis. Given the monooxygenase activity of these GH61 enzymes, they were designated polysaccharide monooxygenases (PMO) or lytic polysaccharide monooxygenases (LPMO) given their ability to instigate crystalline cellulose chain lysis⁵⁴. Dynamic imaging on an LPMO from *Neurospora crassa* using fluorescence from confocal laser scanning microscopy showed that the LPMO introduces carboxyl groups on the surface of crystalline cellulose, while atomic force microscopy confirmed that disruption of crystallinity by LPMO provided access to cellulases for improved degradation⁵⁵. Following their discovery and characterization, LPMOs were utilized in commercial cocktails for cellulose processing for bioethanol production⁵⁶. The commercially produced Novozymes Inc. enzyme cocktail Cellic Ctec2 which contains LPMOs resulted in a 85% cellulose conversion of hydrothermally pretreated wheat straw to fermentable sugars in comparison to a 60% cellulose conversion for the cellulase-only cocktail Celluclast/Novozym 188⁵⁷. The use of LPMOs in commercial enzyme cocktails to function as “cellulase boosters” has become prevalent, with several new commercially available cocktails such as Cellic Ctec3 and Spirizyme (Novozymes Inc.) as well as Accellerase 1500 (Genencor Inc.)⁵⁸⁻⁶⁰. Industrial biotechnological applications of enzymatic degradation of cellulose employ either a saccharification and separate hydrolysis and fermentation (SHF) or a simultaneous saccharification and fermentation approach (SSF)⁶¹. During SHF, soluble sugars are first produced by enzymatic degradation of pretreated biomass, after which yeast is added for fermentation, providing the advantage of giving each process the optimal operation conditions⁶². In SSF, enzymatic deconstruction of cellulose from pretreated biomass and fermentation of these sugars by yeast take place in the same reactor, providing the advantage of efficiency and preventing product inhibition⁶³. While enzymatic degradation of lignocellulose to produce fermentable sugars has the advantages of requiring low energy input, since enzymes require moderate temperature and pH to function, and does not produce toxic or environmentally harmful byproducts as may be the case for chemical treatments, there are several aspects that must still be taken into consideration⁶⁴. Enzymatic cellulose conversion results in lower yields than non-biological methods, is more time consuming, requires large sterile areas and constant monitoring during catalysis⁶⁵. Furthermore, the cost of enzymatic deconstruction of cellulose should also be taken into consideration, since it has been considered a significant

contributor to the final cost of cellulosic ethanol ⁶⁶. The cellulase cost for the production of 1 gallon of ethanol was estimated to be \$0.34, however this decreased to \$0.23 per gallon if soluble sugars from biomass deconstruction were used as a feed source as opposed to pure glucose ⁶⁷. To mitigate costs, future research will be necessary to develop optimized enzymatic cocktail tailored to the biomass substrate composition ^{33,68}. Production of bioethanol from a renewable source such as lignocellulose is important in an era of ever-depleting fossil fuel reserves and the European Council has committed to 0.5% of the fuel supply to be from bioethanol by 2030, while the US seeks to include 36 billion gallons of biofuel in transport fuel consumption by 2022 ⁶⁹.

2.1.3 Lytic polysaccharide monoxygenases

2.1.3.1 Classification of LPMOs

Following their discovery and initial classification as glycoside hydrolases, LPMOs were assigned “auxiliary activities” (AA) enzyme category in the CAZy database ⁷⁰. The AA classification created a category for enzymes that assist glycoside hydrolases (GH), polysaccharide lyases (PL) and carbohydrate esterases (CE) in gaining access to the carbohydrate substrate. LPMOs are classified into the AA9, AA10, AA11, AA13, AA14, AA15 and AA16 families, originating from a range of organisms including bacteria, fungi, algae as well as animals ⁷¹. The various AA families as well as their organism of origin and substrate are specified in Table 2.1. The AA9 family comprises LPMOs of primarily fungal origin that degrade cellulose, galactoglucomannan, xylan or xyloglucan, with some fungi possessing multiple genes ^{53,54,72–75}. LPMOs in the AA10 family are mainly present in bacteria and are involved in chitin and cellulose oxidation ^{52,76,77}. The AA11 family which is active on chitin was established in 2014 in with the characterization of *AoLPMO11* from the fungus *Aspergillus oryzae* RIB40, which thus far remains the only characterized member ⁷⁸. The AA13 family is present in fungi and is involved in starch degradation ^{79,80}. Family AA14 are also of primarily fungal origin and degrade xylan fibers ⁸¹. The AA15 family is present mostly in insects and algae and involved in the degradation of chitin and cellulose, with both characterized members isolated from the insect *Thermobia domestica* ⁸². The newest AA family member is AA16, which is present in fungi

and oomycetes, with a single characterized cellulase-active enzyme, *AaAA16*, isolated from *Aspergillus aculeatus*⁸³.

Table 2.1 Classification of LPMO families.

LPMO	Organism	Characterized enzymes	Structures solved	Polysaccharide substrate
AA9 (GH61)	Fungi	32	16	Cellulose, galactoglucomannan, xylan or xyloglucan (C1 and/or C4)
AA10 (CBM33)	Bacteria, viruses, plants	30	23	Chitin or cellulose (C1)
AA11	Fungi	1	1	Chitin (C1)
AA13	Fungi	4	1	Amylose, amylopectin or starch (C1)
AA14	Fungi	2	1	Xylan (C1)
AA15	Insects, Metazoa, Oomycota, Alveolata, Rhodophyta, Chlorophyta, Haptophyta, Ichthyosporea, Phaeophyceae, Bacillariophyceae, viruses	2	1	Cellulose or chitin (C1)
AA16	Fungi, Oomycota	1	-	Cellulose (C1)

2.1.3.2 Structure of LPMOs

The structure of LPMOs is characterized by a immunoglobulin G-like β -sandwich fold⁸⁴. The LPMO catalytic copper active site is solvent accessible and located on the surface of the protein, forming a flat surface complementary to the planar cellulose substrate⁸⁵. Comparison of the of protein surface of the LPMO families shows a distinct complementarity of the substrate binding surface to its polysaccharide substrate (Figure 2.4)⁸⁶. LPMOs from the family AA9, AA10 and AA11 exhibit a planar binding surface around the copper active site, corresponding to the linear fibrils of their cellulose and chitin substrates^{87,88}. In contrast, the AA13 active site contains a groove above the copper active site that is shaped to house the amylose double helix present in starch and bring the glycosidic bond linkages in proximity to the

coordinated copper ^{80,89}. The xylan-active AA14 displays a contoured surface which may provide complementarity to its polysaccharide substrate shape ⁸¹, while the cellulose-active AA15 appears to have a planar active site surface ⁸².

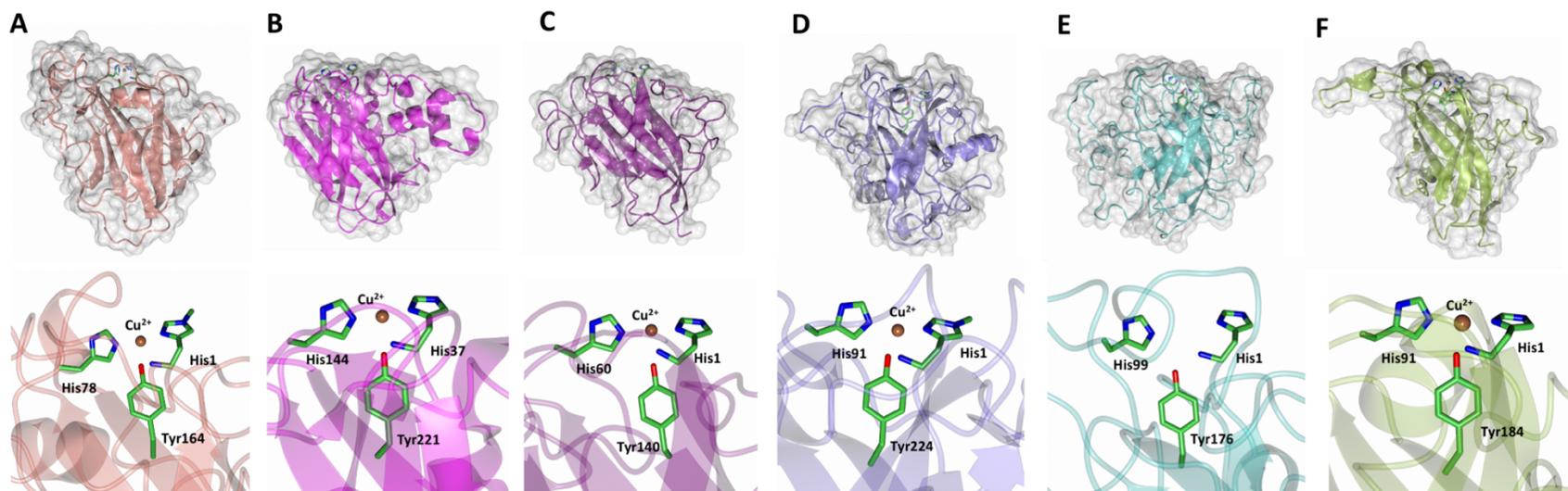


Figure 2.4 Common architecture of the LPMO conserved copper-coordinating histidine brace across families. The copper-center active site is located on the solvent accessible surface with the surface topology complementary to the polysaccharide substrate. Illustrated are (A) AA9 (PDB 5ACH) (B) AA10 (PDB 5OPF) (C) AA11 (PDB 4MAI) (D) AA13 (4OPB) (E) AA14 (5NO7) with putative copper site (F) AA15 (5MSZ).

LPMO substrate specificity and catalytic regioselectivity of LPMOs is governed by loops connecting the β -strands of the substrate binding surface. The loops can be grouped as follows: L2 is the loop between the $\beta 1$ and $\beta 2$ strands ($\beta 1$ and $\beta 3$ in AA10), L3 is the loop between the $\beta 3$ and $\beta 4$ strands, LS for “loop short” is a loop located opposite L2 on the protein and LC is a C-terminal loop⁸⁶. The amino acids within L2, L3, LS and LC interact directly with the polysaccharide substrate as has been shown in X-ray crystallographic studies of *Lentinus similis* LsAA9A and *Collariella virescens* (CvAA9A) bound to oligosaccharide substrates^{87,90}. In AA9 LPMOs, the loops contain a number of aromatic residues that can interact with the oligosaccharide via pyranose ring stacking interactions^{91,92}. On the other hand, AA10 only has one aromatic residue that interacts with the polysaccharide substrate and AA11 and AA13 interact with their polysaccharide substrates via polar residues instead^{78,93}. In addition to ensuring substrate binding, interaction of the LPMO binding surface with the polysaccharide substrate chain ensures positioning of C1 or C4 for oxidation by position of the carbon near the active site copper⁹⁴.

2.1.3.3 Active site

The mono-copper active site, located on the LPMO surface, is coordinated in a characteristic coordination termed the “histidine-brace” which was coined by Quinlan *et al.*⁵³. The histidine brace is composed of the N-terminal histidine that coordinates copper in a bidentate manner through its amino terminal nitrogen and the N_{δ} nitrogen of the imidazole group as well as the N_{ϵ} nitrogen of a second downstream histidine. This results in a T-shaped N_3 coordination geometry in the copper equatorial plane, with Cu-N distances of 1.9 – 2.4 Å^{91,95-97}. Fungal LPMOs, including AA9 and AA13 families, display methylation of the N-terminal N_{ϵ} nitrogen^{53,80,98}. This post-translational modification is not present in LPMOs after heterologous expression in *Pichia pastoris*^{91,99}. It has been proposed that N-terminal methylation contributes to fungal LPMO stability by protecting it from oxidative damage, however absence of the methylation does not negatively impact oxygen activation or substrate binding¹⁰⁰. In the Cu^{2+} form, LPMO AA9, AA11 and AA13 possess four ligands in the equatorial plane composed of the three nitrogen histidine brace atoms and a water molecule, resulting in an approximate square planar geometry. Upon

reduction to the Cu^{1+} form, a transition from a four- to three-coordinate system takes place as the equatorial water is eliminated as indicated by Extended X-Ray Absorption Fine Structure (EXAFS), X-ray Absorption Near Edge Spectroscopy (XANES) and DTF studies on *T. aurantiacus* TaLPMO9A¹⁰¹. While it has been shown that the reduced Cu^{1+} LPMO possesses three equatorial ligands in a T-shaped geometry, there is variation in the ligand identity of the remaining axial and equatorial positions in the Cu^{2+} resting state. LPMOs from the AA9, AA11 and AA13 have a tyrosine residue in the axial position 2.5 – 3.0 Å from the copper center^{78,80,91}. LPMO AA10 families possess a phenylalanine residue in place of the axial tyrosine residue¹⁰². In the Cu^{2+} form, LPMO AA9, AA11 and AA13 exhibit tetragonal geometry with Jahn-Teller distortion resulting in axial elongation^{53,96,101}. Both AA9 and AA10 have two waters coordinated to Cu^{2+} , however while AA9 exhibits a tetragonal geometry with one water in the axial and one in the equatorial position⁸⁴, AA10 has a trigonal bipyramidal structure in which the equatorial plane is formed by the two water ligands and the and the N-terminal nitrogen (Figure 2.5)¹⁰².

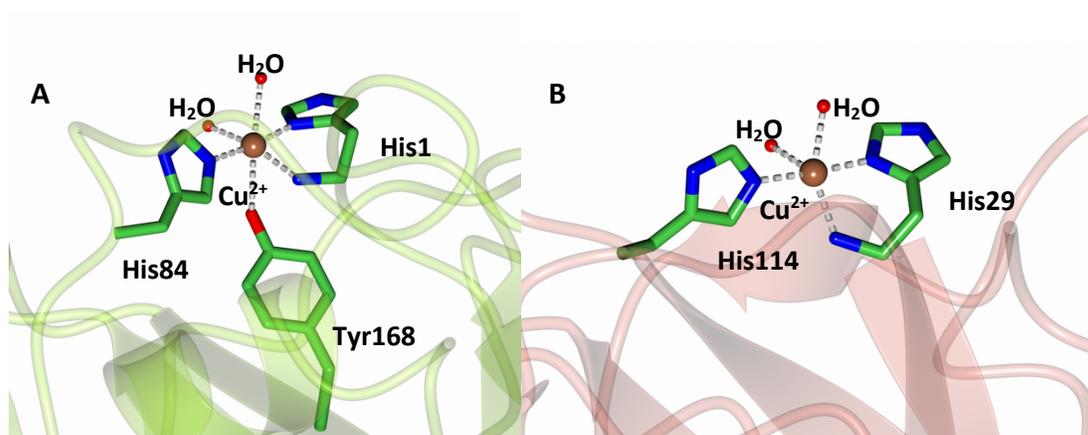


Figure 2.5 Tetragonal and trigonal geometry of AA9 and AA10. **(A)** *Nc*LPMO9D displays a tetragonal geometry in the Cu^{2+} resting state (PDB 5TKG). **(B)** *Ef*AA10 (PDB 4ALC) displays trigonal geometry in the Cu^{2+} resting state.

2.1.3.4 Mechanism of LPMOs

The LPMO active site is responsible for the oxidative cleavage of the glycosidic bond by hydroxylation of the C1 or C4 carbon and the overall reaction can be summarized as follows:



During catalysis oxygen is inserted at the C1 or C4 position which destabilizes the glycosidic bond and results in bond breakage via an elimination reaction (Figure 2.6)^{54,103}. Hydroxylation of the C1 position results in formation of a lactone which is in equilibrium with a reducing-end aldonic acid species via a hydration equilibrium¹⁰⁴. Hydroxylation of the C4 position results in formation of a 4-ketoaldose which hydrates to form a gemdiol with which it is in equilibrium.

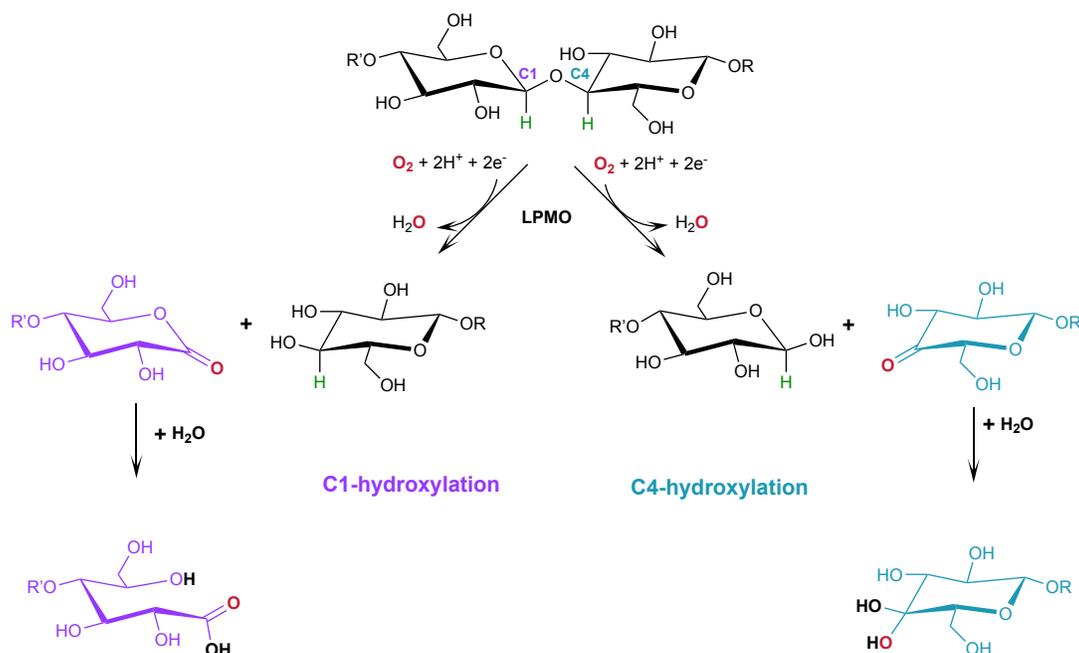


Figure 2.6 Overview of the oxidative reaction catalyzed by LPMOs. LPMOs oxidize either the C1 or C4 of the glycosidic bond which results in polysaccharide substrate chain cleavage. C1 oxidation results in the formation of a lactone which hydrates to form a reducing-end aldonic acid. C4 oxidation results in formation of a ketoaldose which hydrates to form a gem-diol.

Given that LPMOs are responsible for the hydroxylation of the C1 or C4 carbon, but that the resulting bond breakage can be considered spontaneous, it has been reasoned that LPMOs are therefore not inherently “lytic” and should be classified solely as polysaccharide monooxygenases (PMOs)⁵⁴.

The LPMO reaction mechanism is initiated by single electron reduction of the Cu^{2+} center to Cu^{1+} , with this reduced form designated as the activated state^{54,98,103}. Copper center reduction additionally

increases the affinity for the polysaccharide substrate and it has been shown that O₂ binding to Cu¹⁺ is thermoneutral¹⁰⁵⁻¹⁰⁸. Oxygen subsequently binds to the reduced LPMO and is reduced to a superoxo species via an inner sphere pathway resulting in a superoxide Cu²⁺-(O₂⁻) complex as has been demonstrated by DFT calculations informed by (EXAFS), X-ray Absorption Near Edge Spectroscopy (XANES) and DTF studies on *T. aurantiacus* TaLPMO9A by Kjaegaard *et al.*¹⁰¹. Structural studies have sought to characterize the copper-oxygen adduct with Li *et al.* proposing that oxygen coordinates in the copper axial position with end-on positioning in early studies of *Neurospora crassa* NcLPMO9D and NcLPMO9M⁹⁸. The LPMO9D structure indicated a oxygen species with an unusually short O-O bond length of 1.16 Å and a Cu-O distance of ~2.94 Å, while the LPMO9M structure indicated a peroxide species with a O-O bond length of 1.49 Å and a Cu-O distance of 3.44 Å. Both Cu-O bond lengths are longer than the coordination length observed in copper complexes, even when the Jahn-Teller bond length distortion is taken into consideration¹⁰⁹. This may indicate that the observed oxygen species and their coordination was a result of radiation-damage induced artefacts. A further X-ray crystallographic study on an ascorbate reduced NcLPMO9D by O'Dell *et al.* indicated the presence of a peroxy species with a O-O bond length of 1.44 Å in the equatorial plane with side-on positioning¹¹⁰. An X-ray and neutron diffraction study on *Jonesia denitrificans* JdLPMO10A by Bacik *et al.* indicated an oxygen species in the equatorial plane¹¹¹. The X-ray structure contained two molecules in the asymmetric unit, with Molecule A showing a peroxy species with side-on coordination and Molecule B a peroxy species with a O-O distance of 1.5 Å and end-on coordination, both in the equatorial plane. The neutron structure indicated an end-on equatorial dioxygen species, however resolution did not permit a chemical assignment. Caldararu *et al.* later re-refined the data, using the neutron and X-ray data to perform a joint refinement¹¹². Analysis of a joint quantum refinement led the authors to conclude that the Molecule A dioxygen was in an end-on position and QM/MM calculation suggested that it may be a superoxo species. Ambiguity of the identity and coordination of the dioxygen species remains however, given the weak nuclear density and the discrepancy of binding modes. In addition to Kjaegaard *et al.*, QM/MM calculations by Hedegård and Ryde also indicated an equatorial Cu²⁺-(O₂⁻) with end-on geometry^{113,114}. Studies performed using artificial copper complexes have also provided support for

the formation of a superoxide species following oxygen binding to the Cu^{1+} active site with end-on coordination ^{115,116}.

2.1.3.5 H_2O_2 as the LPMO co-substrate

In addition to utilizing oxygen as a co-substrate, it has been found that LPMOs possess peroxygenase activity with H_2O_2 as the co-substrate ¹¹⁷. Studies on the bacterial *Streptomyces coelicolor* ScLPMO10C by Bissaro *et al.* showed that addition of H_2O_2 results in much higher catalytic rates and labelled oxygen atoms in conjunction with mass spectrometry indicated that H_2O_2 was the preferred co-substrate ¹¹⁷. Furthermore, it was found that reducing equivalents are not consumed in stoichiometric amounts during the catalytic cycle, suggesting that the copper reduction occurs once to “prime” the LPMO after which the LPMO remains in this reduced active state for several rounds of catalysis. If H_2O_2 is the primary co-substrate of LPMO, this would result in a “streamlined” catalytic pathway, since following the initial priming reduction, further supply of electrons and protons is not necessary. The identity of the co-substrate remains debated, however, with Hangasky *et al.* having performed analysis of O_2 and H_2O_2 reactivity ¹¹⁸. Kinetic assays and high-resolution mass spectrometry analysis on the fungal *Thermothelomyces thermophila* MtPMO9E (formerly *Myceliophthora thermophila*) indicated that while both O_2 and H_2O_2 are co-substrates, reactions with H_2O_2 results in non-specific oxidation of the polysaccharide substrate, while reactions with O_2 was regioselective for the C4 carbon of the glycosidic bond of the polysaccharide substrate. Hangasky *et al.* posited that under low substrate conditions, when many LPMO active sites remain unbound, off-pathway reactions result in H_2O_2 formation which can outcompete O_2 . When LPMO is saturated with substrate, however, O_2 is the principal co-substrate. A recent study by Hedison *et al.* also found that H_2O_2 is an efficient co-substrate with spectroscopic electron paramagnetic resonance (EPR) studies showing that following an initial priming reduction, copper remains in a Cu^{1+} state for several catalytic cycles ¹¹⁹. Given these findings, it is clear that both O_2 and H_2O_2 can act as the co-substrate, however the LPMO research community remains divided about which is the main co-substrate during catalysis ¹²⁰.

2.1.4 Proposed hydrogen atom abstraction (HAA) species

During LPMO catalysis, it is necessary for an H atom to be abstracted from the C1 or C4 carbon to be hydroxylated in the glycosidic bond ¹²¹. This step presents a significant energy barrier since the bond dissociation energy (BDE) of the glycosidic C-H bond is approximately 100 kcal mol⁻¹ ^{122,123}. Intensive research has been dedicated to the identification of this oxidizing species responsible for hydrogen atom abstraction (HAA), although currently most of these studies span synthetic complexes and computational analysis since it has thus far not been possible to trap the HAA intermediate and characterize it spectroscopically or structurally ¹²⁴. The unique configuration of the histidine brace plays a central role in the LPMO mechanism, facilitating the formation and stabilization of reaction intermediates ¹²⁵. Furthermore, second shell residues may play an important role in regulation of catalysis and intermediate formation ^{114,126,127}. Potential species for HAA include superoxo, hydroperoxo and oxyl intermediates, which will be discussed further below.

2.1.4.1 Superoxo Cu²⁺-(O₂^{•-})

The Cu²⁺-(O₂^{•-}) superoxo species was established as a reaction intermediate in the LPMO mechanism following spectroscopic and DFT studies by Kjaergaard *et al.* ¹⁰¹. They showed that the superoxo species is formed by single electron reduction of O₂ by Cu¹⁺ to form a radical species O₂^{•-}. Additionally, studies on synthetic cupric superoxo complexes suggested a role of the superoxo intermediate in C-H bond activation ¹²⁸. While these synthetic studies indicated that it is energetically feasible for the superoxo species to abstract hydrogen from the substrate C-H bond, the C-H bonds studied were considerably weaker than those of the LPMO polysaccharide substrate at ~72 kcal mol⁻¹ ¹²⁹. In a computational study of the *Thermoascus aurantiacus* TaAA9A structure, Kim *et al.* studied a limited active site composed of the histidine brace, the axial tyrosine, the second shell residues glutamine and histidine, three crystallographic waters and a truncated disaccharide substrate ¹³⁰. It was found that HAA by the superoxo species is less favorable when compared to HAA by a Cu²⁺-(O[•]) oxyl species. It must be noted however, that these calculations were performed with the assumption that oxygen binds in the axial position,

which is not experimentally supported since X-ray crystallographic, spectroscopic and DFT studies have found that oxygen binds to the copper active site in the equatorial position ¹¹⁰.

2.1.4.3 Hydroperoxo Cu^{2+} -(OOH) or Cu^{3+} -(OOH)

The Cu^{2+} -(OOH) hydroperoxo species is formed by the addition of one electron and one proton to the Cu^{2+} -($\text{O}_2^{\cdot-}$) superoxo species ¹³¹. This Cu^{2+} -(OOH) species has been shown to be stabilized by hydrogen bond interactions in a synthetic copper protein with a histidine-like brace ¹³². In a study of the potential of a hydroperoxo species to perform HAA, Trammell *et al.* used a synthetic Cu^{2+} -(OOR) complex as a model for monooxygenase activity ¹³³. It was found that following homolytic cleavage, a generated RO \cdot radical performs HAA and subsequent hydroxylation via a radical rebound mechanism. Alternatively, a Cu^{3+} -(OOH) species can be formed by protonation of the Cu^{2+} -($\text{O}_2^{\cdot-}$) superoxo species. This Cu^{3+} -(OOH) species has been shown to be responsible for substrate HAA in the bicopper peptidylglycine α -hydroxylating monooxygenase (PHM) and it is postulated that LPMOs may follow a similar mechanism ¹³⁴. The Cu^{3+} -(OOH) species can also be formed by reaction of Cu^{2+} with H_2O_2 to form Cu^{3+} -OH species that reacts with H_2O_2 to form an OOH radical which can perform HAA as shown by Garcia-Bosch *et al.* using a synthetic copper complex ¹³⁵. During such catalysis, it can be presumed that a Cu^{3+} -(OOH) intermediate is formed.

2.1.4.4 Hydroxyl Cu^{3+} -(OH)

It is also feasible that HAA occurs following cleavage of the O-O bond ¹²⁴. A model complex of the Cu^{3+} -(OH) hydroxide species with a copper ion coordinated to three nitrogen atoms in T-shaped geometry was used to study C-H bond activation ¹³⁶. The Cu^{3+} -(OH) species was found to have high rates of HAA from phenols and dihydroanthracene, making it a candidate as an intermediate in LPMO catalysis. A hydroxide intermediate has been suggested to be involved in C-H bond activation in particulate methane monooxygenases (pMMO) ¹³⁷. However these were recently shown to be mononuclear copper enzymes, with a single copper responsible for methane oxidation, in contrast to the binuclear copper site that was previously used for mechanistic model studies ¹³⁸. It has been postulated that a ligand is necessary to

stabilize the high oxidation state of the $\text{Cu}^{3+}\text{-(OH)}$ species and assist in driving HAA ¹³⁹. An observed deprotonated coordinating N-terminal nitrogen of the histidine brace in a study by Bacik *et al.* has been put forward as evidence of such a strong electron donating ligand, however subsequent re-refinement of that data using a joint quantum refinement approach did not support the deprotonated nature of the N-terminus ^{112,140}. A further case against the $\text{Cu}^{3+}\text{-(OH)}$ intermediate as the HAA species, lies in the fact that following HAA a $\text{Cu}^{2+}\text{-(OH}_2\text{)}$ species would result which is incapable of undergoing a rebound reaction to form the hydroxylated product ¹²⁵.

2.1.4.5 Oxyl $\text{Cu}^{2+}\text{-(O}^\bullet\text{)}$ and oxo $\text{Cu}^{3+}\text{-(O)}$

Early mass spectrometric studies coupled with computational analysis were used to show that a $\text{Cu}^{2+}\text{-(O}^\bullet\text{)}$ oxyl species can be used for C-H bond activation ¹⁴¹⁻¹⁴³. In their DFT study of *TaAA9A*, Kim *et al.* also investigated whether the formation of a $\text{Cu}^{2+}\text{-(O}^\bullet\text{)}$ oxyl radical could perform substrate HAA ¹³⁰. Their calculations led them to put forward the mechanism that the $\text{Cu}^{2+}\text{-(O}^\bullet\text{)}$ species is responsible for HAA from the substrate, and hydroxylation subsequently occurs via a rebound mechanism. Although this mechanism was energetically favorable, the axial placement of the oxyl species was not supported by additional structural studies ¹¹⁰. Hedegård and Ryde confirmed that the O-O bond is broken before HAA in a QM/MM study of *LsAA9A* with a truncated polysaccharide substrate bound ¹²³. The bond dissociation energy for the $\text{Cu}^{2+}\text{-(O}^\bullet\text{)}$ was found to be -25 kJ mol^{-1} , while the $\text{Cu}^{3+}\text{-(O)}$ species displayed a bond dissociation energy of -34 kJ mol^{-1} , making both these intermediates candidates for HAA. The $\text{Cu}^{2+}\text{-(O}_2^\bullet\text{)}$, $\text{Cu}^{2+}\text{-(OOH)}$ and $\text{Cu}^{3+}\text{-(OH)}$ intermediates were all shown to have endothermic bond dissociation energies, making them less feasible as HAA species. A late follow-up study by Hedegård and Ryde on the same substrate bound *LsAA9A*, confirmed that following Cu^{2+} reduction to Cu^{1+} , oxygen binds and is reduced to the $\text{Cu}^{2+}\text{-(O}_2^\bullet\text{)}$ superoxo species ¹¹⁴. This complex is protonated by a second-shell histidine and further electron reduction results in O-O bond cleavage and formation of a $\text{Cu}^{2+}\text{-(O}^\bullet\text{)}$ oxyl intermediate, the most energetically favorable candidate for HAA. A tour de force study by Bertini *et al.* used QM/MM to evaluate the energetics of *NcLPMO9M* bound to a celloheptaose oligosaccharide to mimic substrate binding ¹⁴⁴.

They concluded that the most energetically feasible mechanism is formation of $\text{Cu}^{2+}\text{-(OOH)}$, followed by one electron reduction and loss of water to form the $\text{Cu}^{2+}\text{-(O}^\bullet\text{'})$ species that performs HAA with the hydroxylated product formed by a rebound mechanism.

2.1.4.6 Oxyl $\text{Cu}^{2+}\text{-(O}^\bullet\text{'})$ derived from H_2O_2

The $\text{Cu}^{2+}\text{-(O}^\bullet\text{'})$ oxyl intermediate has also been shown to be a probable HAA in the H_2O_2 -dependent pathway demonstrated by Bissaro *et al.*¹¹⁷. It was proposed from these findings that the most likely pathway would involve elimination of a water molecule to form a $\text{Cu}^{2+}\text{-(O}^\bullet\text{'})$ oxyl that would perform HAA abstraction and form the hydroxylated product through a rebound mechanism. The utilization of H_2O_2 to form a $\text{Cu}^{2+}\text{-(O}^\bullet\text{'})$ oxyl was additionally confirmed by the QM/MM studies of Bertini *et al.* and Hedegård and Ryde who found that either a $\text{Cu}^{2+}\text{-(O}^\bullet\text{'})$ oxyl or $\text{Cu}^{3+}\text{-(OH)}$ hydroxide species can be formed from H_2O_2 and perform HAA^{114,144}. Wang *et al.* sought to address the H_2O_2 catalytic mechanism and performed QM/MM studies on *LsAA9A* bound to a trisaccharide substrate¹⁴⁵. Wang *et al.* proposed that one electron reduction of H_2O_2 by Cu^{1+} results in O-O bond breakage to form $\text{Cu}^{2+}\text{-(OH)}$ and OH^\bullet . This reactive OH^\bullet radical abstracts a hydrogen from the $\text{Cu}^{2+}\text{-(OH)}$ resulting in formation of a $\text{Cu}^{2+}\text{-(O}^\bullet\text{'})$ oxyl that is responsible for HAA. The proposed formation of a OH^\bullet follows a Fenton-like mechanism, in which H_2O_2 reacts with a reduced metal to form a free hydroxyl radical¹⁴⁶. While the OH^\bullet radical is thought to be highly reactive and result in non-specific reactions, it has been observed in lignocellulose degradation by brown-rot fungi^{147,148}. To this end, Wang *et al.* argue that the hydrogen bonding interactions at the active site may function to stabilize the OH^\bullet radical, preventing off-pathway reactions. A later QM/MM study by Wang *et al.* also found that the $\text{Cu}^{2+}\text{-(O}^\bullet\text{'})$ oxyl functions as the HAA species, even when O_2 is the co-substrate¹⁴⁹. During O_2 activation, Wang *et al.* propose that the $\text{Cu}^{2+}\text{-(O}_2^{\bullet-}\text{'})$ superoxo species formed after O_2 reduction by Cu^{1+} abstracts a hydrogen atom from the ascorbate reductant to form $\text{Cu}^{2+}\text{-(OOH}^\bullet\text{'})$. This $\text{Cu}^{2+}\text{-(OOH}^\bullet\text{'})$ is protonated on its proximal oxygen atom which is followed or coupled to a single electron reduction to form $\text{Cu}^{1+}\text{-(H}_2\text{O}_2)$ which proceeds through the Fenton-like mechanism proposed for H_2O_2 . The authors conclude that this rationalizes the observation that LPMO can utilize both O_2 and H_2O_2 as their co-substrate.

This further explains the formation of H₂O₂ in the absence of substrate, since the formed H₂O₂ is not barricaded in the active site where it can form the reactive oxyl species but diffuses into the bulk-solvent leading to reaction uncoupling.

2.1.5 Electron donors

In order for O₂ to be activated at the copper active site for LPMO catalysis, it is necessary for the resting state Cu²⁺ active site be reduced to Cu¹⁺^{52,103}. The reduction potential of LPMOAA9 is in the range of 150 – 330 mV⁹⁶, while that of AA10 is 220 – 370 mV at pH 6 when compared to a standard hydrogen electrode⁹⁷. The catalytic performance of LPMO can be correlated to the reduction potential of the reductants as shown by Kracher *et al.* using a range of small molecules as well as reductases and dehydrogenases¹⁵⁰. It was further found that pH can be used to modify the reduction potential of reductants, in this case ascorbic acid and 3-methylcatechol¹⁵¹.

2.1.5.1 Small molecule electron donors

Numerous small molecules can function to reduce LPMO, including ascorbic acid and glutathione shown in the early characterization studies by Vaaje-Kolstad on *Sm*LPMO10A⁵². Further small molecules include gallic acid, hydroquinone and catechin¹⁵², however it has been demonstrated that reducing agents are not universally effective and many LPMOs display preference for certain reducing agents¹⁵³. Lignin and lignin derived compounds, which are present in biomass feedstocks, can also act as reductants¹⁵⁴. Lignin present as a high molecular weight polymer can further function as a “reservoir” of electrons driving catalysis in synergy with LPMO action¹⁵⁵.

2.1.5.2 Cellobiose dehydrogenase (CDH)

In addition to reduction by small molecules, LPMO is reduced by several enzymatic electron donors¹²⁴. Cellobiose dehydrogenase (CDH) has been the most extensively characterized^{54,150,156}. CDH enzymes are classified as flavocytochromes that possess a heme-b binding cytochrome domain (CYT) which is connected to a flavin adenine dinucleotide (FAD)-binding dehydrogenase (DH) domain through a flexible linker region¹⁵⁷. CDHs can be divided into various subclasses depending on their structural organization

and organism of origin¹⁵⁸. Class I CDHs are found in basidiomycetes and Class II is found in ascomycetes with Class IIA containing a carbohydrate-binding module (CBM) which is absent in Class IIB¹⁵⁹. The DH domain is responsible for the oxidation of cellobiose and other oligosaccharides resulting in two electron reduction of the FAD cofactor¹⁶⁰. These electrons are transferred to the CYT domain by sequential interdomain electron transfer, after which the electrons are transferred to an electron acceptor^{161,162}. In addition to LPMO acting as the electron acceptor of CDH, CDH can also reduce O₂, resulting in the formation of H₂O₂^{163,164}. The formation of H₂O₂ has led to the proposal that the catalytic role of CDH in cellulose degradation is to generate H₂O₂ for LPMO catalysis, however, the kinetic studies have shown that CDH is more efficient at reducing the LPMO copper active site than O₂^{150,165,166}. NMR and docking studies have suggested that CDH interacts with LPMO through direct contact with the copper active site^{160,161}. The direct interaction of CDH with the LPMO active site required for catalysis in the O₂-dependent mechanism does not address the impediment that bound substrate presents, since the substrate would block access to the active site for reduction by CDH^{121,167}. To circumvent this steric obstacle, the active site tyrosine residue has been proposed to play a role in catalysis, acting as an electron source and forming a radical species quenched by a second electron donation from CDH following substrate hydroxylation and release¹²⁷. Electron transfer from CDH, or indeed small molecules, may proceed via long range electron transfer along LPMO surface aromatic residues^{98,121,155,168}. A very recent study by Wang et al. using QM/MM as well as MD, showed that electron transfer can proceed along the heme propionates of the CYT domain and a network of water molecules and is enhanced by the presence of O₂ at the copper active site¹⁶⁹.

2.1.5.3 Pyrroloquinoline quinone (PQQ)-dependent pyranose dehydrogenase (PDH)

The pyrroloquinoline quinone (PQQ)-dependent pyranose dehydrogenase (PDH) present in *Coprinopsis cinerea*, CcPDH, can also reduce LPMOs from the AA9 family¹⁷⁰. The PDH structure is composed of a CBM domain, a PQQ-dependent dehydrogenase domain and a CYT domain¹⁷¹. PDH has been shown to oxidize D-glucosone, L-fucose and certain pyranoses.

2.1.5.4 Glucose-methanol choline (GMC) Oxidases and Polyphenol Oxidases

Glucose-methanol-choline (GMC) oxidoreductases can reduce LPMO via a plant derived diphenol intermediate¹⁵⁰. The single-domain flavoenzymes glucose dehydrogenase (GDH) and aryl-alcohol quinone oxidoreductase (AAQO) may also function as compatible electron donors to LPMO¹⁷². LPMOs catalysis can also be driven by reducing equivalents derived from the action of polyphenol oxidases that convert lignin-derived monophenols into diphenols¹⁷³. These diphenols were shown to be efficient electron donors of *Myceliophthora thermophila* MtLPMO9B.

2.1.5.5 Light-driven reduction

Cannella *et al.* showed that LPMOs can be reduced by photosynthetic pigments such as chlorophyllin^{174,175}. It was found that the combination of these pigments with reducing agents such as ascorbic acid and light greatly improved catalytic efficiency. This finding has been put into question since it may be the generation of H₂O₂ that is driving the reaction as opposed to direct donation of electrons from the photosynthetic pigment^{117,176}.

2.1.6 Structural studies on oxygen activation in *Neurospora crassa* LPMO9D

The AA9 family of LPMOs is particularly abundant, currently containing 32 characterized enzymes and 16 solved structures according to the CAZy database¹⁷⁷. *Neurospora crassa* is a well-characterized filamentous fungus whose growth on dead plant matter has made it an ideal candidate for cellulose degradation applications^{178,179}. In addition to secreting numerous cellulases, *N. crassa* encodes 14 LPMO genes from the AA9 family and one from the AA13 family¹⁸⁰. Nine of the *N. crassa* LPMOs have been characterized, namely the C1-oxidizing NcLPMO9E, NcLPMO9F, NcLPMO9G and NCU02344; C4-oxidizing NcLPMO9A, NcLPMO9C and NcLPMO9D; mixed C1/C4-oxidizing NcLPMO9B and NcLPMO9M^{72,99,181}. The cellulose regiospecificity of AA9 cellulose glycosidic bond carbon oxygen insertion can be categorized into four classes: Type 1 oxidizes C1; Type 2 oxidizes C4; Type 3 oxidizes C1 or C4 and Type 3* represents a subgroup with structural differences from Type 3 that oxidizes C1¹⁸¹.

The studies described in this chapter center on *NcLPMO9D* and the structural studies that have been performed to elucidate the LPMO mechanism.

The structure of *NcLPMO9D* (initially designated PMO-2) was first elucidated by Li *et al.* with the determination of a 1.10 Å X-ray structure (PDB 4EIR)⁹⁸. The structure revealed a characteristic histidine brace composed of the N-terminal histidine and His84 with an axially coordinated Tyr168. The *NcLPMO9D* active site was present in a planar surface substrate binding region with aromatic residues Tyr25 and Tyr206 postulated to be involved in polysaccharide substrate interactions. The Asn60 residue displayed N-linked glycosylation with two N-acetylglucosamine residues present in the deposited X-ray structure. Due to the fact that this *NcLPMO9D* was purified from the *N. crassa* secretome, the N-terminal His1 displays methylation on the N_ε atom. Interestingly, the structure solved by Li *et al.* displayed a dioxygen species bound in the axial position with a Cu-O distance of 2.94 Å. Given the bond length of 1.16 Å following unrestrained refinement, this species was designated as a superoxo species weakly coordinated to Cu²⁺, with the proximal oxygen atom at an occupancy of 0.80 and the distal oxygen atom at an occupancy of 0.71. The observed Cu-O distance of 2.94 Å is however longer than would be expected for coordination to be feasible¹⁰⁹. It can be presumed that the protein underwent X-ray-induced radiation damage during data collection, which is known to cause metalcenter reduction, lengthening of metal to ligand bonds and formation or loss of ligands^{182–185}. The presence of an activated dioxygen species, while probably an artefact, demonstrated the feasibility of reactive intermediate capture, paving the way for future structural studies^{121,186}.

O'Dell *et al.* subsequently crystallized *NcLPMO9D* and collected a X-ray diffraction structure as well as a neutron diffraction structure¹⁸⁷. To achieve this, *NcLPMO9D* was heterologously expressed in a glycoengineered strain of *Pichia pastoris*, which ensured limited glycosylation of two N-acetylglucosamine subunits linked to five mannose subunits at the Asn60 position. The reduced glycosylation pattern allowed the authors to grow large high-quality crystals following multiple rounds of micro- and macroseeding.

Determination of the Cu^{2+} resting state structure at a resolution of 1.20 Å revealed a copper coordinated in the histidine brace composed of His1 and His84, with an axial Tyr168¹¹⁰. In contrast to Li *et al.*, the N_ϵ atom of His1 was not methylated since expression was performed heterologously in yeast^{87,96,98}. The copper was additionally coordinated to two water molecules in the axial and equatorial position as expected for LPMO AA9 Cu^{2+} coordination geometry⁸⁴. O'Dell *et al.* additionally observed an O_2 molecule in a “pre-binding” position in a pocket delimited by the residues His157 and Gln166 which are conserved in LPMOs in the AA9 family¹¹⁰.

Following reduction of the Cu^{2+} site to Cu^{1+} using the small molecule ascorbate, the *NcLPMO9D* was flash-frozen and a 1.20 Å X-ray crystal structure solved. This structure revealed an activated dioxygen species present in the copper equatorial plane with end-on geometry, which agrees with experimental and computational findings (Figure 2.7)^{101,113,114,123}. The activated dioxygen species had an O-O bond length of 1.44 ± 0.06 Å and a Cu-O bond length of 1.90 ± 0.05 Å, which did not permit definitive discrimination between the superoxo or peroxy species.

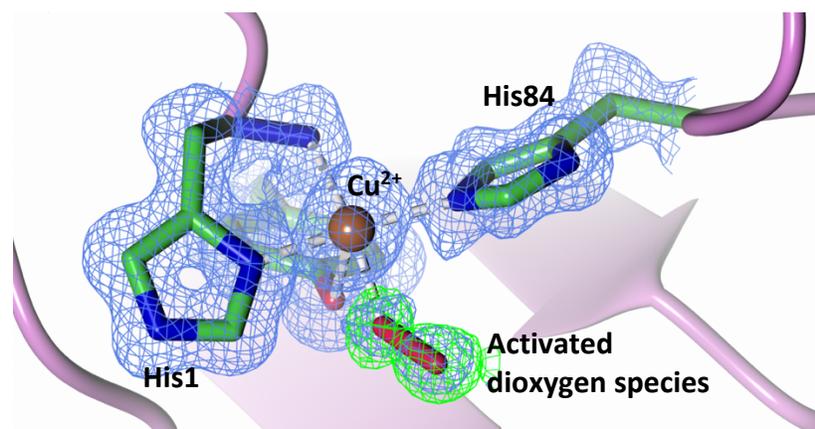


Figure 2.7 The active site of the ascorbate-reduced *NcLPMO9D* (PDB 5TKH). Copper reduction of the LPMO9D crystals resulted in the binding of molecular oxygen at the equatorial plane and activation to a superoxo/peroxy species. $2F_o - F_c$ density ($\sigma = 1.00$) and $F_o - F_c$ density ($\sigma = 3.00$) is displayed as blue and green mesh, respectively.

A 2.12 Å room-temperature neutron diffraction structure was also collected on a large high-quality crystal in the Cu^{2+} resting state. This structure was solved using joint refinement with a 1.50 Å X-ray

diffraction dataset collected on the same crystal. Analysis of the neutron scattering length density maps (NSLD) of the neutron protein structure permitted identification of the protonation states and hydrogen bond interactions of active site and second-shell residues (Figure 2.8). In particular, the authors noted that the His157 residue, which forms part of the O₂ pre-binding pocket, was singly N_ε-protonated.

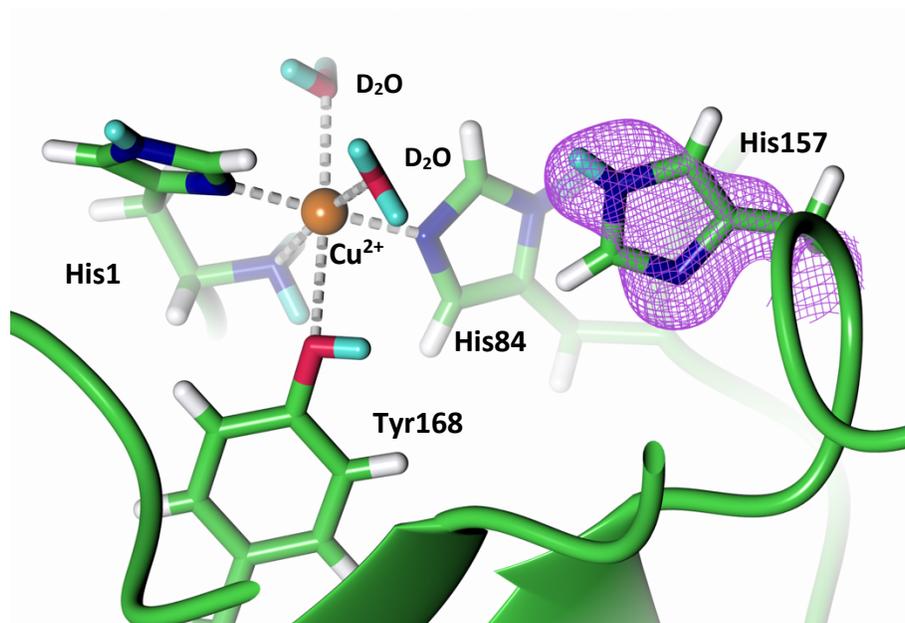


Figure 2.8 Neutron structure of the resting state active site of *NcLPMO9D* (PDB 5TKI). The second-shell His157 is neutral in the Cu²⁺ resting state. 2F_o-F_c NSLD map (σ = 1.50) is displayed as purple mesh; H and D atoms displayed in white and turquoise respectively.

2.1.7 Aim of research

It stands to reason that since the optimal pH for cellulase function^{148,188,189}, along which LPMOs perform catalysis in a synergistic manner, is acidic, it could be expected that this His157 may be doubly N_ε- and N_δ-protonated and positively charged. It was postulated that double protonation of His157 may make pre-binding of O₂ for catalysis more likely. DFT calculations of the limited active site including His157 and Gln166 indeed showed that double His157 protonation makes oxygen pre-binding thermodynamically favorable with $\Delta G = -17.2 \text{ kcal mol}^{-1}$, in comparison to a $\Delta G = -1.1 \text{ kcal mol}^{-1}$. These structural studies on *NcLPMO9D* lay the groundwork for the examination of oxygen activation and protonation states described herein. Using structural interrogation of the *NcLPMO9D* active site and second-shell residues, we aim to further characterize the reaction intermediates and mechanism of LPMO9D.

2.1.8 References

1. McKendry, P. Energy production from biomass (part 1): Overview of biomass. *Bioresour. Technol.* **83**, 37–46 (2002).
2. Ragauskas, A. J. *et al.* The path forward for biofuels and biomaterials. *Science* **311**, 484–9 (2006).
3. Zhou, C. H., Xia, X., Lin, C. X., Tong, D. S. & Beltramini, J. Catalytic conversion of lignocellulosic biomass to fine chemicals and fuels. *Chem. Soc. Rev.* **40**, 5588–5617 (2011).
4. Somerville, C. *et al.* Toward a systems approach to understanding plant cell walls. *Science*. **306**, 2206–2211 (2004).
5. Carrier, M. *et al.* Thermogravimetric analysis as a new method to determine the lignocellulosic composition of biomass. *Biomass and Bioenergy* **35**, 298–307 (2011).
6. Isikgor, F. H. & Becer, C. R. Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers. *Polym. Chem.* **6**, 4497–4559 (2015).
7. Cocinero, E. J., Gamblin, D. P., Davis, B. G. & Simons, J. P. The building blocks of cellulose: The intrinsic conformational structures of cellobiose, its epimer, lactose, and their singly hydrated complexes. *J. Am. Chem. Soc.* **131**, 11117–11123 (2009).
8. Sánchez, C. Lignocellulosic residues: Biodegradation and bioconversion by fungi. *Biotechnol. Adv.* **27**, 185–194 (2009).
9. Houfani, A. A., Anders, N., Spiess, A. C., Baldrian, P. & Benallaoua, S. Insights from enzymatic degradation of cellulose and hemicellulose to fermentable sugars— a review. *Biomass and Bioenergy* **134**, 105481 (2020).
10. Ebringerová, A., Hromádková, Z. & Heinze, T. Hemicellulose. in *Polysaccharides I* 1–67 (Springer-Verlag, 2005). doi:10.1007/b136816
11. Pauly, M. *et al.* Hemicellulose biosynthesis. *Planta* **238**, 627–642 (2013).
12. Del Río, J. C., Marques, G., Rencoret, J., Martínez, Á. T. & Gutiérrez, A. Occurrence of naturally acetylated lignin units. *J. Agric. Food Chem.* **55**, 5461–5468 (2007).
13. Terrett, O. M. & Dupree, P. Covalent interactions between lignin and hemicelluloses in plant secondary cell walls. *Curr. Opin. Biotechnol.* **56**, 97–104 (2019).
14. Chen, H. *Biotechnology of Lignocellulose. Biotechnology of Lignocellulose: Theory and Practice* (Springer Netherlands, 2014). doi:10.1007/978-94-007-6898-7
15. Taarning, E. *et al.* Zeolite-catalyzed biomass conversion to fuels and chemicals. *Energy Environ. Sci.* **4**, 793–804 (2011).
16. Bilal, M., Asgher, M., Iqbal, H. M. N., Hu, H. & Zhang, X. Biotransformation of lignocellulosic materials into value-added products—A review. *Int. J. Biol. Macromol.* **98**, 447–458 (2017).
17. Li, P., Sakuragi, K. & Makino, H. Extraction techniques in sustainable biofuel production: A concise review. *Fuel Process. Technol.* **193**, 295–303 (2019).
18. Arevalo-Gallegos, A., Ahmad, Z., Asgher, M., Parra-Saldivar, R. & Iqbal, H. M. N. Lignocellulose: A sustainable material to produce value-added products with a zero waste approach—A review. *Int. J. Biol. Macromol.* **99**, 308–318 (2017).
19. Baruah, J. *et al.* Recent trends in the pretreatment of lignocellulosic biomass for value-added products. *Front. Energy Res.* **6**, 1–19 (2018).
20. Taherzadeh, M. J. & Karimi, K. *Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: A review. International Journal of Molecular Sciences* **9**, (2008).
21. Silva, I. S., de Menezes, C. R., Franciscon, E., dos Santos, E. da C. & Durrant, L. R. Degradation of lignosulfonic and tannic acids by ligninolytic soil fungi cultivated under microaerobic conditions. *Brazilian Arch. Biol. Technol.* **53**, 693–699 (2010).
22. Vasco-Correa, J., Ge, X. & Li, Y. Biological Pretreatment of Lignocellulosic Biomass. *Biomass Fractionation Technol. a Lignocellul. Feed. Based Biorefinery* 561–585 (2016). doi:10.1016/B978-0-12-802323-5.00024-4
23. Fernando, S., Adhikari, S., Chandrapal, C. & Murali, N. Biorefineries: Current status, challenges, and future direction. *Energy and Fuels* **20**, 1727–1737 (2006).

24. Parajuli, R. *et al.* Biorefining in the prevailing energy and materials crisis: A review of sustainable pathways for biorefinery value chains and sustainability assessment methodologies. *Renew. Sustain. Energy Rev.* **43**, 244–263 (2015).
25. Axelsson, L. *et al.* Perspective: Jatropha cultivation in southern India: Assessing farmers' experiences. *Biofuels, Bioprod. Biorefining* **6**, 246–256 (2012).
26. Rinaldi, R. & Schüth, F. Acid hydrolysis of cellulose as the entry point into biorefinery schemes. *ChemSusChem* **2**, 1096–1107 (2009).
27. Casas-Godoy, L. *et al.* *Biofuels. Biobased Products and Industries* **108**, (Springer Berlin Heidelberg, 2007).
28. Cantarel, B. I. *et al.* The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycomics. *Nucleic Acids Res.* **37**, 233–238 (2009).
29. Morana, A., Maurelli, L., Ionata, E., Cara, F. La & Rossi, M. Cellulases from fungi and bacteria and their biotechnological applications. in *Cellulose: Types and Action, Mechanism and Uses* (ed. Golan, A. E.) 1–79 (Nova Science Publishers, Inc., 2011).
30. Bischof, R. H., Ramoni, J. & Seiboth, B. Cellulases and beyond: The first 70 years of the enzyme producer *Trichoderma reesei*. *Microb. Cell Fact.* **15**, 1–13 (2016).
31. Gusakov, A. V. Alternatives to *Trichoderma reesei* in biofuel production. *Trends Biotechnol.* **29**, 419–425 (2011).
32. Chandel, A. K. *et al.* Comparative analysis of key technologies for cellulosic ethanol production from Brazilian sugarcane bagasse at a commercial scale. *Biofuels, Bioprod. Biorefining* **13**, 994–1014 (2019).
33. Adsul, M. *et al.* Designing a cellulolytic enzyme cocktail for the efficient and economical conversion of lignocellulosic biomass to biofuels. *Enzyme Microb. Technol.* **133**, 109442 (2020).
34. Singh, N. *et al.* Cellulosic ethanol production via consolidated bioprocessing by a novel thermophilic anaerobic bacterium isolated from a Himalayan hot spring. *Biotechnol. Biofuels* **10**, 1–18 (2017).
35. Panchapakesan, A. & Shankar, N. Fungal Cellulases: An Overview. in *New and Future Developments in Microbial Biotechnology and Bioengineering* 9–18 (Elsevier, 2016). doi:10.1016/B978-0-444-63507-5.00002-2
36. Haldar, D., Sen, D. & Gayen, K. A review on the production of fermentable sugars from lignocellulosic biomass through conventional and enzymatic route—a comparison. *Int. J. Green Energy* **13**, 1232–1253 (2016).
37. Talamantes, D., Biabini, N., Dang, H., Abdoun, K. & Berlemont, R. Natural diversity of cellulases, xylanases, and chitinases in bacteria. *Biotechnol. Biofuels* **9**, 1–11 (2016).
38. Wang, M. & Lu, X. Exploring the synergy between cellobiose dehydrogenase from *Phanerochaete chrysosporium* and cellulase from *Trichoderma reesei*. *Front. Microbiol.* **7**, 1–10 (2016).
39. Cohen, R., Suzuki, M. R. & Hammel, K. E. Processive endoglucanase active in crystalline cellulose hydrolysis by the brown rot basidiomycete *Gloeophyllum trabeum*. *Appl. Environ. Microbiol.* **71**, 2412–2417 (2005).
40. Couturier, M. & Berrin, J.-G. The Saccharification Step: The Main Enzymatic Components. in *Lignocellulose Conversion* 93–110 (Springer Berlin Heidelberg, 2013). doi:10.1007/978-3-642-37861-4_5
41. Hiras, J. *et al.* Comparative Community Proteomics Demonstrates the Unexpected Importance of Actinobacterial Glycoside Hydrolase Family 12 Protein for Crystalline Cellulose Hydrolysis. *MBio* **7**, (2016).
42. Estela, R. & Luis, J. Hydrolysis of Biomass Mediated by Cellulases for the Production of Sugars. in *Sustainable Degradation of Lignocellulosic Biomass - Techniques, Applications and Commercialization* (eds. Chandel, A. & Da Silva, S. S.) 119–155 (InTech, 2013). doi:10.5772/53719
43. Teeri, T. T. Crystalline cellulose degradation: New insight into the function of cellobiohydrolases. *Trends Biotechnol.* **15**, 160–167 (1997).

44. Jalak, J., Kurašin, M., Teugjas, H. & Väljamä, P. Endo-exo synergism in cellulose hydrolysis revisited. *J. Biol. Chem.* **287**, 28802–28815 (2012).
45. Yang, M. *et al.* Synergistic cellulose hydrolysis dominated by a multi-modular processive endoglucanase from *Clostridium cellulosi*. *Front. Microbiol.* **7**, 1–8 (2016).
46. Andlar, M. *et al.* Lignocellulose degradation: An overview of fungi and fungal enzymes involved in lignocellulose degradation. *Eng. Life Sci.* **18**, 768–778 (2018).
47. Reese, E. T. Enzymatic Hydrolysis of Cellulose. *Appl. Microbiol.* **4**, 39–45 (1956).
48. Arfi, Y., Shamshoum, M., Rogachev, I., Peleg, Y. & Bayer, E. A. Integration of bacterial lytic polysaccharide monoxygenases into designer cellulosomes promotes enhanced cellulose degradation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 9109–9114 (2014).
49. Karlsson, J. *et al.* Homologous expression and characterization of Cel61A (EG IV) of *Trichoderma reesei*. *Eur. J. Biochem.* **268**, 6498–6507 (2001).
50. Koseki, T. *et al.* Biochemical characterization of a glycoside hydrolase family 61 endoglucanase from *Aspergillus kawachii*. *Appl. Microbiol. Biotechnol.* **77**, 1279–1285 (2008).
51. Harris, P. V. *et al.* Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: Structure and function of a large, enigmatic family. *Biochemistry* **49**, 3305–3316 (2010).
52. Vaaje-Kolstad, G. *et al.* An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* **330**, 219–222 (2010).
53. Quinlan, R. J. *et al.* 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 (2011).
54. Phillips, C. M., Beeson, W. T., Cate, J. H. & Marletta, M. A. Cellobiose dehydrogenase and a copper-dependent polysaccharide monoxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem. Biol.* **6**, 1399–1406 (2011).
55. Eibinger, M. *et al.* Cellulose surface degradation by a lytic polysaccharide monoxygenase and its effect on cellulase hydrolytic efficiency. *J. Biol. Chem.* **289**, 35929–35938 (2014).
56. Harris, P. V., Xu, F., Kreel, N. E., Kang, C. & Fukuyama, S. New enzyme insights drive advances in commercial ethanol production. *Curr. Opin. Chem. Biol.* **19**, 162–170 (2014).
57. Cannella, D., Hsieh, C. W. C., Felby, C. & Jørgensen, H. Production and effect of aldonic acids during enzymatic hydrolysis of lignocellulose at high dry matter content. *Biotechnol. Biofuels* **5**, 1–10 (2012).
58. Dixit, P. *et al.* A screening approach for assessing lytic polysaccharide monoxygenase activity in fungal strains. *Biotechnol. Biofuels* **12**, 1–16 (2019).
59. Barbosa, F. C., Silvello, M. A. & Goldbeck, R. Cellulase and oxidative enzymes: new approaches, challenges and perspectives on cellulose degradation for bioethanol production. *Biotechnol. Lett.* **42**, 875–884 (2020).
60. Rani Singhania, R. *et al.* Role and significance of lytic polysaccharide monoxygenases (LPMOs) in lignocellulose deconstruction. *Bioresour. Technol.* **335**, 125261 (2021).
61. Dahnum, D., Tasum, S. O., Triwahyuni, E., Nurdin, M. & Abimanyu, H. Comparison of SHF and SSF processes using enzyme and dry yeast for optimization of bioethanol production from empty fruit bunch. *Energy Procedia* **68**, 107–116 (2015).
62. Gupta, R., Sharma, K. K. & Kuhad, R. C. Separate hydrolysis and fermentation (SHF) of *Prosopis juliflora*, a woody substrate, for the production of cellulosic ethanol by *Saccharomyces cerevisiae* and *Pichia stipitis*-NCIM 3498. *Bioresour. Technol.* **100**, 1214–1220 (2009).
63. Cantarella, M., Cantarella, L., Gallifuoco, A., Spera, A. & Alfani, F. Comparison of different detoxification methods for steam-exploded poplar wood as a substrate for the bioproduction of ethanol in SHF and SSF. *Process Biochem.* **39**, 1533–1542 (2004).
64. Capolupo, L. & Faraco, V. Green methods of lignocellulose pretreatment for biorefinery development. *Appl. Microbiol. Biotechnol.* **100**, 9451–9467 (2016).
65. Balan, V. Current Challenges in Commercially Producing Biofuels from Lignocellulosic Biomass. *ISRN Biotechnol.* **2014**, 1–31 (2014).

66. Olofsson, J., Barta, Z., Börjesson, P. & Wallberg, O. Integrating enzyme fermentation in lignocellulosic ethanol production: Life-cycle assessment and techno-economic analysis. *Biotechnol. Biofuels* **10**, 1–14 (2017).
67. Dragone, G. *et al.* Innovation and strategic orientations for the development of advanced biorefineries. *Bioresour. Technol.* **302**, 122847 (2020).
68. Agrawal, R. *et al.* Current understanding of the inhibition factors and their mechanism of action for the lignocellulosic biomass hydrolysis. *Bioresour. Technol.* **332**, 125042 (2021).
69. Kargbo, H., Harris, J. S. & Phan, A. N. “Drop-in” fuel production from biomass: Critical review on techno-economic feasibility and sustainability. *Renew. Sustain. Energy Rev.* **135**, 110168 (2021).
70. Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M. & Henrissat, B. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol. Biofuels* **6**, 1–14 (2013).
71. Manavalan, T., Stepnov, A. A., Hegnar, O. A. & Eijsink, V. G. H. Sugar oxidoreductases and LPMOs – two sides of the same polysaccharide degradation story? *Carbohydr. Res.* **505**, (2021).
72. Agger, J. W. *et al.* 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 (2014).
73. Frommhagen, M. *et al.* Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase. *Biotechnol. Biofuels* **8**, 4–15 (2015).
74. Hüttner, S. *et al.* Specific Xylan Activity Revealed for AA9 Lytic Polysaccharide Monooxygenases of the Thermophilic Fungus *Malbranchea cinnamomea* by Functional Characterization. *Appl. Environ. Microbiol.* **85**, 1–13 (2019).
75. Lenfant, N. *et al.* A bioinformatics analysis of 3400 lytic polysaccharide oxidases from family AA9. *Carbohydr. Res.* **448**, 166–174 (2017).
76. Vaaje-Kolstad, G., Horn, S. J., Van Aalten, D. M. F., Synstad, B. & Eijsink, V. G. H. The non-catalytic chitin-binding protein CBP21 from *Serratia marcescens* is essential for chitin degradation. *J. Biol. Chem.* **280**, 28492–28497 (2005).
77. Forsberg, Z. *et al.* Cleavage of cellulose by a CBM33 protein. *Protein Sci.* **20**, 1479–1483 (2011).
78. 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.* **10**, 122–126 (2014).
79. Vu, V. V., Beeson, W. T., Span, E. A., Farquhar, E. R. & Marletta, M. A. A family of starch-active polysaccharide monooxygenases. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 13822–13827 (2014).
80. Lo Leggio, L. *et al.* Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase. *Nat. Commun.* **6**, (2015).
81. Couturier, M. *et al.* Lytic xylan oxidases from wood-decay fungi unlock biomass degradation. *Nat. Chem. Biol.* **14**, 306–310 (2018).
82. Sabbadin, F. *et al.* An ancient family of lytic polysaccharide monooxygenases with roles in arthropod development and biomass digestion. *Nat. Commun.* **9**, (2018).
83. Filiatrault-Chastel, C. *et al.* AA16, a new lytic polysaccharide monooxygenase family identified in fungal secretomes. *Biotechnol. Biofuels* **12**, 1–15 (2019).
84. Karkehabadi, S. *et al.* The First Structure of a Glycoside Hydrolase Family 61 Member, Cel61B from *Hypocrea jecorina*, at 1.6 Å Resolution. *J. Mol. Biol.* **383**, 144–154 (2008).
85. Aachmann, F. L., Sørli, M., Skjåk-Bræk, G., Eijsink, V. G. H. & Vaaje-Kolstad, G. NMR structure of a lytic polysaccharide monooxygenase provides insight into copper binding, protein dynamics, and substrate interactions. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 18779–18784 (2012).
86. Zhou, X. & Zhu, H. Current understanding of substrate specificity and regioselectivity of LPMOs. *Bioresour. Bioprocess.* **7**, (2020).
87. Frandsen, K. E. H. *et al.* The molecular basis of polysaccharide cleavage by lytic polysaccharide monooxygenases. *Nat. Chem. Biol.* **12**, 298–303 (2016).
88. Courtade, G., Forsberg, Z., Heggset, E. B., Eijsink, V. G. H. & Aachmann, F. L. The carbohydrate-binding module and linker of a modular lytic polysaccharide monooxygenase promote localized cellulose oxidation. *J. Biol. Chem.* **293**, 13006–13015 (2018).

89. Vu, V. V. & Marletta, M. A. Starch-degrading polysaccharide monooxygenases. *Cell. Mol. Life Sci.* **73**, 2809–2819 (2016).
90. Simmons, T. J. *et al.* Structural and electronic determinants of lytic polysaccharide monooxygenase reactivity on polysaccharide substrates. *Nat. Commun.* **8**, (2017).
91. Wu, M. *et al.* Crystal structure and computational characterization of the lytic polysaccharide monooxygenase GH61D from the basidiomycota fungus *Phanerochaete chrysosporium*. *J. Biol. Chem.* **288**, 12828–12839 (2013).
92. Liu, B. *et al.* Structural and molecular dynamics studies of a C1-oxidizing lytic polysaccharide monooxygenase from *Heterobasidion irregulare* reveal amino acids important for substrate recognition. *FEBS J.* **285**, 2225–2242 (2018).
93. Forsberg, Z. *et al.* 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 (2014).
94. Frommhagen, M., Westphal, A. H., van Berkel, W. J. H. & Kabel, M. A. Distinct Substrate Specificities and Electron-Donating Systems of Fungal Lytic Polysaccharide Monooxygenases. *Front. Microbiol.* **9**, 1–22 (2018).
95. Mekasha, S. *et al.* Structural and functional characterization of a small chitin-active lytic polysaccharide monooxygenase domain of a multi-modular chitinase from *Jonesia denitrificans*. *FEBS Lett.* **590**, 34–42 (2016).
96. Borisova, A. S. *et al.* Structural and functional characterization of a lytic polysaccharide monooxygenase with broad substrate specificity. *J. Biol. Chem.* **290**, 22955–22969 (2015).
97. Hemsworth, G. R. *et al.* The copper active site of CBM33 polysaccharide oxygenases. *J. Am. Chem. Soc.* **135**, 6069–6077 (2013).
98. Li, X., Beeson, W. T., Phillips, C. M., Marletta, M. A. & Cate, J. H. D. Structural Basis for Substrate Targeting and Catalysis by Fungal Polysaccharide Monooxygenases. *Structure* **20**, 1051–1061 (2012).
99. Kittl, R., Kracher, D., Burgstaller, D., Haltrich, D. & Ludwig, R. Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol. Biofuels* **5**, 79 (2012).
100. Petrović, D. M. *et al.* Methylation of the N-terminal histidine protects a lytic polysaccharide monooxygenase from auto-oxidative inactivation. *Protein Sci.* **27**, 1636–1650 (2018).
101. Kjaergaard, C. H. *et al.* Spectroscopic and computational insight into the activation of O₂ by the mononuclear Cu center in polysaccharide monooxygenases. *Proc. Natl. Acad. Sci.* **111**, 8797–8802 (2014).
102. Gudmundsson, M. *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).
103. Beeson, W. T., Phillips, C. M., Cate, J. H. D. & Marletta, M. A. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J. Am. Chem. Soc.* **134**, 890–892 (2012).
104. Silva, C. D. O. G. *et al.* Combination of MALDI-TOF MS and UHPLC-ESI-MS for the characterization of lytic polysaccharide monooxygenase activity. *Anal. Methods* **12**, 149–161 (2020).
105. Kracher, D., Andlar, M., Furtmüller, P. G. & Ludwig, R. Active-site copper reduction promotes substrate binding of fungal lytic polysaccharide monooxygenase and reduces stability. *J. Biol. Chem.* **293**, 1676–1687 (2018).
106. Hangasky, J. A. & Marletta, M. A. A Random-Sequential Kinetic Mechanism for Polysaccharide Monooxygenases. *Biochemistry* **57**, 3191–3199 (2018).
107. Lanci, M. P. *et al.* Isotopic Probing of Molecular Oxygen Activation at Copper(I) Sites. *J. Am. Chem. Soc.* **129**, 14697–14709 (2007).
108. Kunishita, A. *et al.* Active Site Models for the Cu A Site of Peptidylglycine α -Hydroxylating Monooxygenase and Dopamine β -Monooxygenase. *Inorg. Chem.* **51**, 9465–9480 (2012).

109. Solomon, E. I. *et al.* Copper active sites in biology. *Chem. Rev.* **114**, 3659–3853 (2014).
110. O'Dell, W. B., Agarwal, P. K. & Meilleur, F. Oxygen Activation at the Active Site of a Fungal Lytic Polysaccharide Monooxygenase. *Angew. Chemie - Int. Ed.* **56**, 767–770 (2017).
111. Michalczyk, R. *et al.* Joint neutron crystallographic and NMR solution studies of Tyr residue ionization and hydrogen bonding: Implications for enzyme-mediated proton transfer. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 5673–5678 (2015).
112. Caldararu, O., Oksanen, E., Ryde, U. & Hedegård, E. D. Mechanism of hydrogen peroxide formation by lytic polysaccharide monooxygenase. *Chem. Sci.* **10**, 576–586 (2019).
113. Hedegård, E. D. & Ryde, U. Multiscale Modelling of Lytic Polysaccharide Monooxygenases. *ACS Omega* **2**, 536–545 (2017).
114. Hedegård, E. D. & Ryde, U. Molecular mechanism of lytic polysaccharide monooxygenases. *Chem. Sci.* **9**, 3866–3880 (2018).
115. Komiyama, K. *et al.* Dioxygen Reactivity of Copper(I) Complexes with Tetradentate Tripodal Ligands Having Aliphatic Nitrogen Donors: Synthesis, Structures, and Properties of Peroxo and Superoxo Complexes. *Bull. Chem. Soc. Jpn.* **77**, 59–72 (2004).
116. Lee, J. Y. *et al.* Mechanistic insights into the oxidation of substituted phenols via hydrogen atom abstraction by a cupric-superoxo complex. *J. Am. Chem. Soc.* **136**, 9925–9937 (2014).
117. Bissaro, B. *et al.* Oxidative cleavage of polysaccharides by monocopper enzymes depends on H₂O₂. *Nat. Chem. Biol.* **13**, 1123–1128 (2017).
118. Hangasky, J. A., Iavarone, A. T. & Marletta, M. A. Reactivity of O₂ versus H₂O₂ with polysaccharide monooxygenases. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 4915–4920 (2018).
119. Hedison, T. M. *et al.* Insights into the H₂O₂-driven catalytic mechanism of fungal lytic polysaccharide monooxygenases. *FEBS J.* **288**, 4115–4128 (2021).
120. Bissaro, B., Várnai, A., Røhr, Å. K. & Eijsink, V. G. H. Oxidoreductases and Reactive Oxygen Species in Conversion of Lignocellulosic Biomass. *Microbiol. Mol. Biol. Rev.* **82**, 1–51 (2018).
121. Beeson, W. T., Vu, V. V., Span, E. A., Phillips, C. M. & Marletta, M. A. Cellulose Degradation by Polysaccharide Monooxygenases. *Annu. Rev. Biochem.* **84**, 923–946 (2015).
122. Gao, J., Thomas, D. A., Sohn, C. H. & Beauchamp, J. L. Biomimetic reagents for the selective free radical and acid-base chemistry of glycans: Application to glycan structure determination by mass spectrometry. *J. Am. Chem. Soc.* **135**, 10684–10692 (2013).
123. Hedegård, E. D. & Ryde, U. Targeting the reactive intermediate in polysaccharide monooxygenases. *J. Biol. Inorg. Chem.* **22**, 1029–1037 (2017).
124. Meier, K. K. *et al.* Oxygen Activation by Cu LPMOs in Recalcitrant Carbohydrate Polysaccharide Conversion to Monomer Sugars. *Chem. Rev.* **118**, 2593–2635 (2018).
125. Ciano, L., Davies, G. J., Tolman, W. B. & Walton, P. H. Bracing copper for the catalytic oxidation of C–H bonds. *Nat. Catal.* **1**, 571–577 (2018).
126. Span, E. A., Suess, D. L. M., Deller, M. C., Britt, R. D. & Marletta, M. A. The Role of the Secondary Coordination Sphere in a Fungal Polysaccharide Monooxygenase. *ACS Chem. Biol.* **12**, 1095–1103 (2017).
127. Mcevoy, A., Creutzberg, J., Singh, R. K., Bjerrum, M. J. & Hedegård, E. D. The role of the active site tyrosine in the mechanism of lytic polysaccharide monooxygenase. *Chem. Sci.* **12**, 352–362 (2021).
128. Peterson, R. L. *et al.* Cupric superoxo-mediated intermolecular C-H activation chemistry. *J. Am. Chem. Soc.* **133**, 1702–1705 (2011).
129. Liu, J. J., Diaz, D. E., Quist, D. A. & Karlin, K. D. Copper(I)-Dioxygen Adducts and Copper Enzyme Mechanisms. *Isr. J. Chem.* **56**, 738–755 (2016).
130. Kim, S., Ståhlberg, J., Sandgren, M., Paton, R. S. & Beckham, G. T. Quantum mechanical calculations suggest that lytic polysaccharide monooxygenases use a copper-oxyl, oxygen-rebound mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 149–154 (2014).
131. Elwell, C. E. *et al.* Copper-Oxygen Complexes Revisited: Structures, Spectroscopy, and Reactivity. *Chem. Rev.* **117**, 2059–2107 (2017).

132. Mann, S. I., Heinisch, T., Ward, T. R. & Borovik, A. S. Peroxide Activation Regulated by Hydrogen Bonds within Artificial Cu Proteins. *J. Am. Chem. Soc.* **139**, 17289–17292 (2017).
133. Trammell, R. *et al.* Decoding the Mechanism of Intramolecular Cu-Directed Hydroxylation of sp³ C-H Bonds. *J. Org. Chem.* **82**, 7887–7904 (2017).
134. Abad, E., Rommel, J. B. & Kästner, J. Reaction mechanism of the bicopper enzyme peptidylglycine α -hydroxylating monooxygenase. *J. Biol. Chem.* **289**, 13726–13738 (2014).
135. Garcia-Bosch, I. & Sieglar, M. A. Copper-Catalyzed Oxidation of Alkanes with H₂O₂ under a Fenton-like Regime. *Angew. Chemie Int. Ed.* **55**, 12873–12876 (2016).
136. Donoghue, P. J. *et al.* Rapid C–H Bond Activation by a Monocopper(III)–Hydroxide Complex. *J. Am. Chem. Soc.* **133**, 17602–17605 (2011).
137. Himes, R. A. & Karlin, K. D. Copper-dioxygen complex mediated C-H bond oxygenation: relevance for particulate methane monooxygenase (pMMO). *Curr. Opin. Chem. Biol.* **13**, 119–131 (2009).
138. Ross, M. O. *et al.* Particulate methane monooxygenase contains only mononuclear copper centers. *Science*. **364**, 566–570 (2019).
139. Dhar, D. *et al.* Perturbing the Copper(III)-Hydroxide Unit through Ligand Structural Variation. *J. Am. Chem. Soc.* **138**, 356–368 (2016).
140. Bacik, J. P. *et al.* Neutron and Atomic Resolution X-ray Structures of a Lytic Polysaccharide Monooxygenase Reveal Copper-Mediated Dioxygen Binding and Evidence for N-Terminal Deprotonation. *Biochemistry* **56**, 2529–2532 (2017).
141. Schröder, D., Holthausen, M. C. & Schwarz, H. Radical-like activation of alkanes by the ligated copper oxide cation (phenanthroline)CuO⁺. *J. Phys. Chem. B* **108**, 14407–14416 (2004).
142. Yoshizawa, K., Kihara, N., Kamachi, T. & Shiota, Y. Catalytic mechanism of dopamine β -monooxygenase mediated by Cu(III)-oxo. *Inorg. Chem.* **45**, 3034–3041 (2006).
143. Dietl, N., Van Der Linde, C., Schlangen, M., Beyer, M. K. & Schwarz, H. Diatomic [CuO]⁺ and its role in the spin-selective hydrogen- and oxygen-atom transfers in the thermal activation of methane. *Angew. Chemie - Int. Ed.* **50**, 4966–4969 (2011).
144. Bertini, L. *et al.* Catalytic Mechanism of Fungal Lytic Polysaccharide Monooxygenases Investigated by First-Principles Calculations. *Inorg. Chem.* **57**, 86–97 (2018).
145. Wang, B. *et al.* QM/MM Studies into the H₂O₂-Dependent Activity of Lytic Polysaccharide Monooxygenases: Evidence for the Formation of a Caged Hydroxyl Radical Intermediate. *ACS Catal.* **8**, 1346–1351 (2018).
146. Lloyd, R. V., Hanna, P. M. & Mason, R. P. The origin of the hydroxyl radical oxygen in the fenton reaction. *Free Radic. Biol. Med.* **22**, 885–888 (1997).
147. Arantes, V., Jellison, J. & Goodell, B. Peculiarities of brown-rot fungi and biochemical Fenton reaction with regard to their potential as a model for bioprocessing biomass. *Appl. Microbiol. Biotechnol.* **94**, 323–338 (2012).
148. Cragg, S. M. *et al.* Lignocellulose degradation mechanisms across the Tree of Life. *Curr. Opin. Chem. Biol.* **29**, 108–119 (2015).
149. Wang, B., Walton, P. H. & Rovira, C. Molecular Mechanisms of Oxygen Activation and Hydrogen Peroxide Formation in Lytic Polysaccharide Monooxygenases. *ACS Catal.* **9**, 4958–4969 (2019).
150. Kracher, D. *et al.* Extracellular electron transfer systems fuel cellulose oxidative degradation. *Science*. **352**, 1098–1101 (2016).
151. Frommhagen, M. *et al.* Quantification of the catalytic performance of C1-cellulose-specific lytic polysaccharide monooxygenases. *Appl. Microbiol. Biotechnol.* **102**, 1281–1295 (2018).
152. Isaksen, T. *et al.* A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *J. Biol. Chem.* **289**, 2632–2642 (2014).
153. Frommhagen, M. *et al.* Lytic polysaccharide monooxygenases from *Myceliophthora thermophila* C1 differ in substrate preference and reducing agent specificity. *Biotechnol. Biofuels* **9**, 1–17 (2016).
154. Dimarogona, M., Topakas, E., Olsson, L. & Christakopoulos, P. Lignin boosts the cellulase performance of a GH-61 enzyme from *Sporotrichum thermophile*. *Bioresour. Technol.* **110**, 480–487 (2012).

155. Westereng, B. *et al.* Enzymatic cellulose oxidation is linked to lignin by long-range electron transfer. *Sci. Rep.* **5**, 1–9 (2015).
156. Langston, J. A. *et al.* Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Appl. Environ. Microbiol.* **77**, 7007–7015 (2011).
157. Zamocky, M. *et al.* Cellobiose Dehydrogenase – A Flavocytochrome from Wood-Degrading, Phytopathogenic and Saprotrophic Fungi. *Curr. Protein Pept. Sci.* **7**, 255–280 (2006).
158. Csarman, F., Wohlschlager, L. & Ludwig, R. Cellobiose dehydrogenase. in *Flavin-Dependent Enzymes: Mechanisms, Structures and Applications* (eds. Chaiyen, P. & Fuyuhiko Tamanoi) 457–489 (Academic Press, 2020). doi:10.1016/bs.enz.2020.06.002
159. Harreither, W. *et al.* Catalytic properties and classification of cellobiose dehydrogenases from ascomycetes. *Appl. Environ. Microbiol.* **77**, 1804–1815 (2011).
160. Tan, T. C. *et al.* Structural basis for cellobiose dehydrogenase action during oxidative cellulose degradation. *Nat. Commun.* **6**, (2015).
161. Courtade, G. *et al.* Interactions of a fungal lytic polysaccharide monooxygenase with β -glucan substrates and cellobiose dehydrogenase. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 5922–5927 (2016).
162. Loose, J. S. M. *et al.* Activation of bacterial lytic polysaccharide monooxygenases with cellobiose dehydrogenase. *Protein Sci.* **25**, 2175–2186 (2016).
163. Igarashi, K., Momohara, I., Nishino, T. & Samejima, M. Kinetics of inter-domain electron transfer in flavocytochrome cellobiose dehydrogenase from the white-rot fungus *Phanerochaete chrysosporium*. *Biochem. J.* **365**, 521–6 (2002).
164. Sygmond, C. *et al.* Semi-rational engineering of cellobiose dehydrogenase for improved hydrogen peroxide production. *Microb. Cell Fact.* **12**, 1–10 (2013).
165. Paul M. Wood. Pathways for production of Fenton ' s reagent by wood-rotting fungi. *FEMS Microbiol. Rev.* **13**, 313–320 (1994).
166. Bissaro, B., Várnai, A., Røhr, Å. K. & Eijsink, V. G. H. Oxidoreductases and Reactive Oxygen Species in Conversion of Lignocellulosic Biomass. *Microbiol. Mol. Biol. Rev.* **82**, (2018).
167. Meier, K. K. *et al.* Oxygen Activation by Cu LPMOs in Recalcitrant Carbohydrate Polysaccharide Conversion to Monomer Sugars. *Chem. Rev.* **118**, 2593–2635 (2018).
168. Courtade, G. *et al.* Interactions of a fungal lytic polysaccharide monooxygenase with β -glucan substrates and cellobiose dehydrogenase. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 5922–5927 (2016).
169. Wang, Z., Feng, S., Rovira, C. & Wang, B. How Oxygen Binding Enhances Long-Range Electron Transfer: Lessons From Reduction of Lytic Polysaccharide Monooxygenases by Cellobiose Dehydrogenase. *Angew. Chemie - Int. Ed.* **60**, 2385–2392 (2021).
170. Várnai, A., Umezawa, K., Yoshida, M. & Eijsink, V. G. H. The Pyrroloquinoline-Quinone-Dependent Pyranose Dehydrogenase from *Coprinopsis cinerea* Drives Lytic Polysaccharide Monooxygenase Action. *Appl. Environ. Microbiol.* **84**, (2018).
171. Matsumura, H. *et al.* Discovery of a Eukaryotic Pyrroloquinoline Quinone-Dependent oxidoreductase belonging to a new auxiliary activity family in the database of carbohydrate-active enzymes. *PLoS One* **9**, 4–11 (2014).
172. Garajova, S. *et al.* Single-domain flavoenzymes trigger lytic polysaccharide monooxygenases for oxidative degradation of cellulose. *Sci. Rep.* **6**, 1–9 (2016).
173. Frommhagen, M. *et al.* Boosting LPMO-driven lignocellulose degradation by polyphenol oxidase-activated lignin building blocks. *Biotechnol. Biofuels* **10**, 1–16 (2017).
174. Cannella, D. *et al.* Light-driven oxidation of polysaccharides by photosynthetic pigments and a metalloenzyme. *Nat. Commun.* **7**, 1–8 (2016).
175. Möllers, K. B. *et al.* On the formation and role of reactive oxygen species in light-driven LPMO oxidation of phosphoric acid swollen cellulose. *Carbohydr. Res.* **448**, 182–186 (2017).
176. Chylenski, P. *et al.* Lytic Polysaccharide Monooxygenases in Enzymatic Processing of Lignocellulosic Biomass. *ACS Catal.* **9**, 4970–4991 (2019).
177. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, 490–495 (2014).

178. Eberhart, B. M., Beck, R. S. & Goolsby, K. M. Cellulase of *Neurospora crassa*. *J. Bacteriol.* **130**, 181–186 (1977).
179. Yazdi, M. T., Radford, A., Keen, J. N. & Woodward, J. R. Cellulase production by *Neurospora crassa*: Purification and characterization of cellulolytic enzymes. *Enzyme Microb. Technol.* **12**, 120–123 (1990).
180. Petrović, D. M. *et al.* Comparison of three seemingly similar lytic polysaccharide monooxygenases from *Neurospora crassa* suggests different roles in plant biomass degradation. *J. Biol. Chem.* **294**, 15068–15081 (2019).
181. Vu, V. V., Beeson, W. T., Phillips, C. M., Cate, J. H. D. & Marletta, M. A. Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *J. Am. Chem. Soc.* **136**, 562–565 (2014).
182. Horrell, S. *et al.* Serial crystallography captures enzyme catalysis in copper nitrite reductase at atomic resolution from one crystal. *IUCrJ* **3**, 271–281 (2016).
183. Yano, J. *et al.* X-ray damage to the Mn4Ca complex in single crystals of photosystem II: A case study for metalloprotein crystallography. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 12047–12052 (2005).
184. George, G. N. *et al.* X-ray-induced photo-chemistry and X-ray absorption spectroscopy of biological samples. *J. Synchrotron Radiat.* **19**, 875–886 (2012).
185. Nass, K. *et al.* Indications of radiation damage in ferredoxin microcrystals using high-intensity X-FEL beams. *J. Synchrotron Radiat.* **22**, 225–238 (2015).
186. Span, E. A. & Marletta, M. A. The framework of polysaccharide monooxygenase structure and chemistry. *Curr. Opin. Struct. Biol.* **35**, 93–99 (2015).
187. O'Dell, W. B., Swartz, P. D., Weiss, K. L. & Meilleur, F. Crystallization of a fungal lytic polysaccharide monooxygenase expressed from glycoengineered *Pichia pastoris* for X-ray and neutron diffraction. *Acta Crystallogr. Sect. Struct. Biol. Commun.* **73**, 70–78 (2017).
188. Schülein, M. Enzymatic properties of cellulases from *Humicola insolens*. *J. Biotechnol.* **57**, 71–81 (1997).
189. Boer, H. & Koivula, A. The relationship between thermal stability and pH optimum studied with wild-type and mutant *Trichoderma reesei* cellobiohydrolase Cel7A. *Eur. J. Biochem.* **270**, 841–848 (2003).

2.2 Preliminary results on neutron and X-ray diffraction data collection of a lytic polysaccharide monooxygenase under reduced and acidic conditions

The following work was reprinted with permission from: Schröder, G. C., O'Dell, W. B., Swartz, P. D. & Meilleur, F. Preliminary results of neutron and X-ray diffraction data collection on a lytic polysaccharide monooxygenase under reduced and acidic conditions. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **77**, 128–133 (2021).

2.2.1 Synopsis

Neutron diffraction data collection of *Neurospora crassa* LPMO9D under ascorbate-reduced conditions at cryo-temperature and at low-pH conditions at room temperature are presented here.

2.2.2 Abstract

Lytic polysaccharide monooxygenases (LPMOs) are copper-center enzymes involved in the oxidative cleavage of the glycosidic bond in crystalline cellulose and other polysaccharides. The LPMO reaction is initiated by the addition of a reductant and oxygen to ultimately form an unknown activated copper–oxygen species responsible for polysaccharide substrate hydrogen atom abstraction. Given the sensitivity of metalloproteins to radiation damage, neutron protein crystallography provides a non-destructive technique for structural characterization while also informing on the position of hydrogen atoms. Neutron cryo-crystallography permits trapping of catalytic intermediates, thereby providing insight into protonation states and chemical nature of otherwise short-lived species in the reaction mechanism. To characterize the reaction mechanism intermediates of LPMO9D from *Neurospora crassa*, a cryo-neutron diffraction dataset was collected on an ascorbate-reduced crystal. A second neutron diffraction dataset was collected at room temperature on a LPMO9D crystal exposed to low pH conditions to probe the protonation states of ionizable groups involved in catalysis under acidic conditions.

2.2.3 Keywords

Lytic polysaccharide monooxygenases; *Neurospora crassa*, neutron protein crystallography; reduction; protonation; enzymatic mechanism.

2.2.4 Introduction

Fungi and bacteria employ a number of enzymes functioning in concert to unlock the energetic potential of lignocellulose for their cellular functions^{1,2}. This enzymatic conversion of biomass is valuable for the generation of biofuels. Several commercial cocktails composed of glycoside hydrolases have been commercially employed^{3,4}. However, the inherent recalcitrance of lignocellulose, particularly its crystalline cellulose content, is known to limit the yields of solubilized sugar. This hurdle has been alleviated by addition of lytic polysaccharide monooxygenases (LPMOs) that assist in the oxidative polysaccharide cleavage^{5,6}. LPMOs are a class of copper-dependent enzymes involved in the oxidative cleavage of cellulose at the glycosidic bond, thereby disrupting crystalline cellulose and increasing its accessibility to cellulases for hydrolytic depolymerization^{7,8}. The structure of LPMOs is characterized by an active site located at the planar enzyme–cellulose binding surface and a copper(II) ion coordinated by two conserved histidine residues shown in Figure 2.9⁹. Within this “histidine-brace” (Figure 2.9 B) the copper is coordinated by three nitrogen atoms in the equatorial plane: two from the N-terminal histidine amino group and imidazole ring and one from a second histidine imidazole ring^{9,10}. LPMOs perform glycosidic bond cleavage of the polysaccharide substrate by hydroxylation of a glycosidic carbon at the C1 or C4 position¹¹. In the O₂-dependent pathway, the reaction is initiated by a single electron reduction of the copper(II) center to copper(I), after which molecular dioxygen spontaneously binds and is reduced to the superoxide species¹². The reaction mechanism of LPMO requires further reduction and protonation to complete the mechanism; however the nature of the intermediates and oxidizing species responsible for hydroxylation are an ongoing area of investigation¹³. Central to the characterization of the mechanism is the chemical nature of the hydrogen atom abstracting (HAA) species responsible for abstraction of the hydrogen atom from the substrate. Potential candidates include superoxide^{8,10,14}, oxyl^(15–19), and hydroxy²⁰ and hydroperoxy²¹ species. Several computational studies have interrogated the LPMO mechanism with

quantum mechanical calculations^{12,16,18,22–25}; however structural data to validate these results are still needed. The majority of LPMO crystallographic studies are of the resting state²⁶, with a limited number of oxygen-^{27,28} or substrate-bound^{29–31} structures.

We previously determined the X-ray structure of substrate-free *Neurospora crassa* LPMO9D in its reduced form bound to an activated dioxygen intermediate²⁷. This structure has provided insight onto the initial stages of oxygen activation directly following copper reduction. The activated dioxygen intermediate was present in the equatorial copper coordination site with end-on geometry as had been previously proposed for LPMO and other mononuclear copper enzymes and synthetic complexes^{12,32,33}. However, given the limitations of X-ray crystallography for revealing the position of light atoms such as hydrogen, the exact chemical identity and protonation state of the activated dioxygen species remains to be determined. Neutron protein crystallography represents an effective technique to determine the protonation state and orientation of critical active site residues and water molecules, thereby providing insight into the underlying protein chemistry^{34–36}. It is also a non-destructive technique, particularly suited to metalloproteins sensitive to photoreduction³⁷. To further characterize the chemical nature of the copper-bound, activated dioxygen species and protonation states at the LPMO copper active site, we sought to employ neutron diffraction to solve the full-atom structure of a cryogenically trapped intermediate following LPMO9D reduction. Neutron cryo-crystallography has seen a rise in implementation since, similar to X-ray cryo-crystallography, it allows short-lived reaction intermediates to be captured, providing information on their chemical identity³⁸. Since LPMOs oxidatively degrade lignocellulose alongside cellulases which display optimum activity in acidic conditions (pH 3.5 – pH 5.5), we additionally measured neutron diffraction from LPMO9D under relevant acidic conditions^{4,39,40}. Our previous neutron crystallography findings of LPMO9D suggest that the protonation of second-shell His157 may influence oxygen pre-binding to the active site²⁷, and we sought to further characterize the protonation state of this ionizable group as well as the hydrogen bond interactions of the conserved nearby Gln166 at acidic conditions.

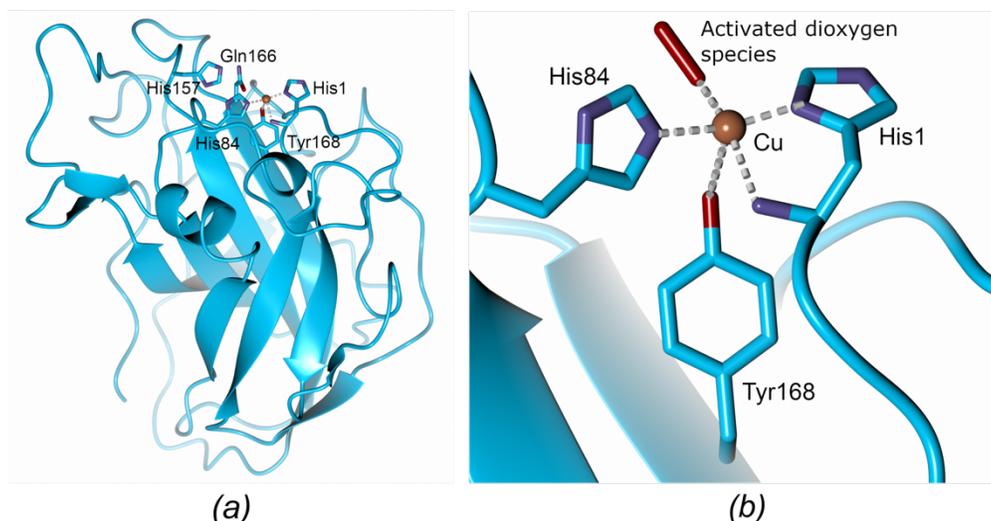


Figure 2.9 The structure of the LPMO active site (PDB code 5TKH; O’Dell, Agarwal *et al.*, 2017). **(A)** Cartoon model of the LPMO structure with the copper active site situated on the surface for interaction with the planar cellulose substrate. Conserved second shell residues His157 and Gln166 are displayed. **(B)** “Histidine-brace” active site of LPMO9D from *Neurospora crassa* composed of a central copper, two planar histidine residues and an axial tyrosine.

Reported here are two neutron crystallography datasets of *N. crassa* LPMO9D under reduced and acidic conditions collected on the Macromolecular Neutron Diffractometer (MaNDi) at the Spallation Neutron Source (SNS) at Oak Ridge National Laboratory⁴¹. Cryo-neutron diffraction data collection of an ascorbate reduced LPMO9D crystal was performed to trap the activated dioxygen intermediate at the LPMO9D active site. Room temperature X-ray and neutron diffraction datasets were furthermore collected on the LPMO9D resting structure at pH 4.4 (pD 4.8). Determination of the protonation states of the intermediates and titratable residues under these conditions will further inform on the oxygen activation mechanism of LPMOs, a topic of active investigation.

2.2.5 Methods and Materials

2.2.5.1 Protein production and crystallization

A glycoengineered form of LPMO9D from *Neurospora crassa* with limited mannose glycosylation was heterologously expressed in *Komagataella phaffii* (previously *Pichia pastoris*) and crystallized as previously reported⁴². In summary, crystals of LPMO9D with sufficient size for neutron protein diffraction were grown by vapor diffusion over two months in 750 μ L sitting drops in 9-well glass plates with a 25 mL reservoir (Hampton Research). The reservoir buffer was composed of 22%(w/v) PEG 3350 and 100 mM HEPES pH 6.0. A crystal measuring \sim 1500 μ m x 900 μ m was mounted in a thin-walled quartz capillary (Hampton Research) with a deuterated reservoir buffer plug to achieve vapor exchange of titratable residues and the crystal solvent from H to D.

2.2.5.2 Data Collection

Initial white beam neutron diffraction tests (Figure 2.10 A) were performed on the IMAGINE beamline at the High Flux Isotope Reactor (HFIR)⁴³. Test frames clearly displayed the superposition of two distinct diffraction patterns, both extending to high resolution indicating that the crystal had cracked during capillary mounting (Figure 2.11 A). Microscopic examination of the crystal revealed a very fine hairline fracture. The crystal was then split in two along this apparent faultline, and each piece was mounted separately (Figure 2.11 B and 2.11 C).

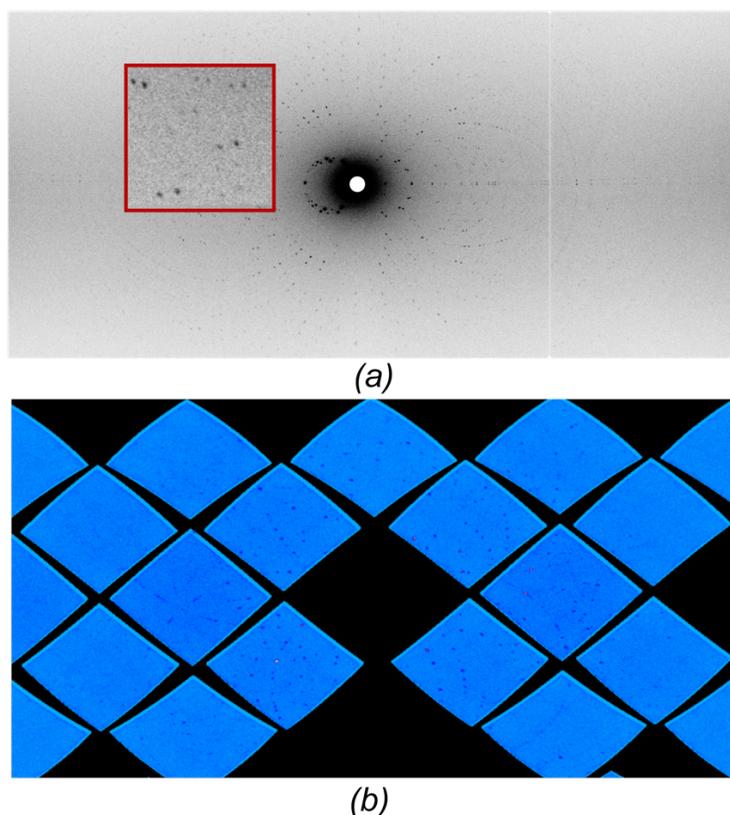


Figure 2.10 Neutron diffraction images of LPMO9D. **(A)** Test Laue neutron diffraction pattern of the LPMO9D crystal on IMAGINE at HFIR displaying superposition of two distinct diffraction patterns. **(B)** Quasi-Laue Neutron diffraction pattern of the low pH room temperature dataset at MaNDi at SNS following manual crystal separation and mounting.

To reduce the resting state copper(II) to copper(I) the crystal was soaked for two hours in deuterated buffer containing 100 mM HEPES pD 6.0 (pH 5.6), 25%(w/v) PEG 3350 and 100mM ascorbic acid pD 6.0 (pH 5.6). The crystal was subsequently flash-cooled in liquid nitrogen using a cryoprotectant composed of the deuterated soaking buffer supplemented with 25%(v/v) glycerol. Neutron time-of-flight diffraction data was collected at 100 K on the MaNDi instrument at the SNS ⁴⁴. An incident neutron wavelength of 2 – 4 Å was used with a $\Delta\phi$ of 10° between frames. A total of 14 diffraction images were collected with an average exposure of 18 hours.

Low pH LPMO9D conditions were achieved in the second crystal via vapor exchange of the quartz capillary mounted crystal in three stages. The deuterated buffer plug was initially composed of 22%(w/v)

PEG 3350 and 100 mM sodium acetate pD 5.6 (pH 5.2) and exchange proceeded for three days. The acetate component of the deuterated buffer plug was then decreased to pD 5.2 (pH 4.8) for six days and finally pD 4.8 (pH 4.2) for 16 days. Neutron diffraction data were collected on MaNDi (Figure 2.10 B) at 298 K with an incident neutron wavelength of 2 – 4 Å and 15 diffraction images with an average exposure of 30 hours and a $\Delta\varphi$ of 10° between frames was used. Following collection of neutron diffraction data, an X-ray dataset was collected on the same crystal at 298 K on a home-source X-ray diffractometer. Additionally, small-size LPMO9D crystals were directly soaked in a low pH buffer for X-ray diffraction data collection. The crystals were harvested and incubated in buffer composed of 22%(w/v) PEG 3350 and 100 mM sodium acetate of which the pH was sequentially lowered to pH 5.6, pH 5.2 and finally pH 4.8 with a 10 minute incubation at each step. These low pH crystals were flash-cooled in liquid nitrogen with the pH 4.8 acidic buffer containing 25%(v/v) glycerol as cryoprotectant.

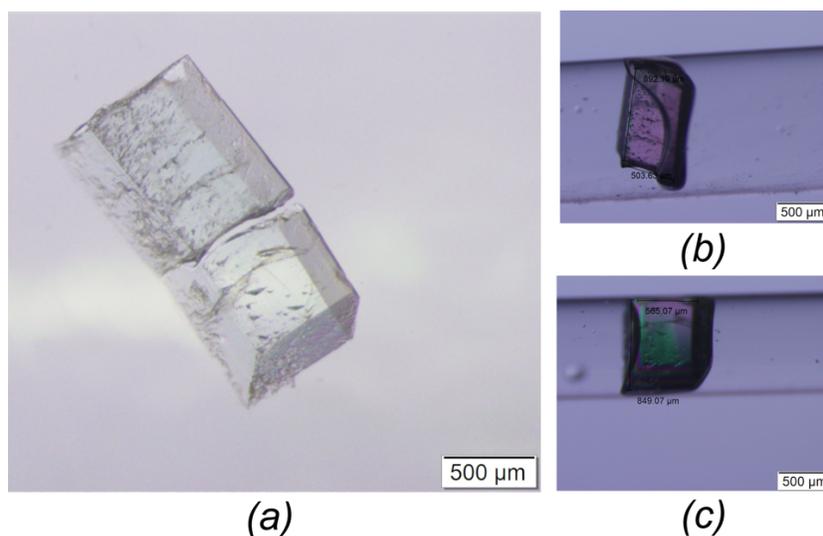


Figure 2.11 LPMO9D crystal for data collection. **(A)** The original large LPMO9D crystal contained a small hairline fracture which affected diffraction quality. This fracture was used to separate the crystal into two pieces as shown **(B)** An individually mounted half of the larger LPMO crystal in a 1.5 mm quartz capillary used to collect cryo-data of the ascorbate reduced form. **(C)** The second half of the original LPMO9D crystal in a 1.0 mm quartz capillary used to collect data under low pH conditions.

Both neutron datasets were reduced using the *Mantid* package⁴⁵, integrated using three dimensional profile fitting⁴⁶ with *LAUENORM* from the *LAUEGEN* suite being used for wavelength normalization^{47,48}.

The X-ray data were indexed, integrated using *CrysAlis PRO* (Rigaku, Woodlands, Texas, USA) and scaled and merged with *AIMLESS* in the *CCP4* suite^{49,50}.

2.2.6 Results and Discussion

The unique interactions of neutrons with light elements makes neutron protein crystallography sensitive to hydrogen and its isotope deuterium which are often invisible in X-ray diffraction⁵¹. Hydrogen atoms are central to enzyme chemistry and intermediate formation, and in order to capitalize on their visibility with neutron protein crystallography we collected two datasets of LPMO9D under reduced and acidic conditions. The ascorbate reduced LPMO9D neutron structure was collected on MaNDi at the SNS at 100 K at a resolution of 2.4 Å (Table 2.2) to probe the protonation state of activated oxygen intermediates.

Table 2.2 Data Collection and Processing.
Values in parentheses are for the outer shell.

	Ascorbate soak (neutron)	Low pH vapor (neutron)	Low pH vapor (X-ray)	Direct soak low pH (X-ray)
Diffraction source	BL-11B, MaNDi, SNS	BL-11B, MaNDi, SNS	MicroMax-007 HF, Rigaku	MicroMax-007 HF, Rigaku
Wavelength (Å)	2.0-4.0	2.0-4.0	1.54	1.54
Temperature (K)	100	Ambient	Ambient	100
Detector	SNS Anger Camera	SNS Anger Camera	Eiger 4M Detector	Eiger 4M Detector
Crystal-to-detector distance (mm)	450	450	77	57
Rotation range per image (°)	0†	0†	1	0.50
Total rotation range (°)	140	150	360	n/a‡
No. of exposures	14	15	360	1284
Exposure time per image (s)	64800	108000	2	10
Space group	<i>P</i> 1 2 ₁ 1			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	67.73, 42.18, 69.76	68.30, 42.27, 70.41	68.30, 42.27, 70.41	67.67, 42.21, 69.72
α , β , γ (°)	90, 99, 90	90, 98.47, 90	90, 98.47, 90	90, 98.90, 90
Resolution range (Å)	14.79 - 2.40 (2.49 - 2.40)	14.65 - 2.14 (2.22 - 2.14)	12.65 - 1.90 (1.97 - 1.90)	12.61 - 1.50 (1.55- 1.50)
Total No. of reflections	45082 (2985)	75277 (6053)	190116 (11672)	512918 (27801)
No. of unique reflections	14168 (1300)	18963 (1655)	30950 (2859)	62507 (6159)
Completeness (%)	91.16 (84.86)	84.65 (74.75)	97.56 (91.75)	99.78 (99.97)
Redundancy	3.2 (2.3)	4.0 (3.3)	6.1 (4.1)	8.2 (4.5)
$\langle I/\sigma(I) \rangle$	7.15 (2.28)	10.01 (3.02)	15.24 (4.40)	62.56 (7.14)
R _{meas}	0.2148 (0.3327)	0.1751 (0.3824)	0.1258 (0.3443)	0.1495 (0.3133)
R _{merge}	0.1851 (0.2737)	0.1532 (0.328)	0.1152 (0.299)	0.142 (0.2775)
CC _{1/2}	0.956 (0.297)	0.97 (0.363)	0.991 (0.88)	0.992 (0.94)
Overall <i>B</i> factor from	31.17	34.32	17.17	11.79
Wilson plot (Å ²)				

† Still quasi-Laue neutron diffraction images were collected with a 10° ϕ rotation between each frame

‡ Data strategy was optimized to maximize completeness using the CrysAlis strategy routines.

Neutron data generally display completeness in the range of 80 – 85%⁵², however the data was very complete at 91.16% due to the high number of frames collected and the spherical arrangement of the 44 MaNDi Anger camera detectors. We further collected a 298 K dataset on a LPMO9D crystal in which the pH had been decreased via vapor exchange to a resolution of 2.14 Å (Table 2.2). Lowering of the pH of a protein crystal by vapor equilibration is most suitable for the large fragile crystals used in neutron protein crystallography and has been demonstrated^{53,54}. An X-ray dataset was collected on the low pH vapor exchange crystal at 298 K following neutron diffraction for joint refinement with the neutron data to a resolution of 1.90 Å. The completeness of the neutron dataset at low pH was 84.65% while that of the X-ray dataset was 97.56%. To verify that the pH of the LPMO9D crystal had been successfully acidified by vapor exchange, we performed a direct soak of a separate smaller LPMO9D crystal at pH 4.8 collected a 100 K X-ray dataset to verify similarity between the structures (Table 2.2). The X-ray dataset collected at cryo-conditions extended to 1.50 Å resolution.

Cryogenic neutron crystallography provides valuable structural and mechanistic insights³⁸ with several recent cryo-studies performed on the MaNDi instrument⁵⁵⁻⁵⁷. Presented here is the first neutron cryogenic dataset of a freeze-trapped active-site intermediate collected on MaNDi at the SNS. We further collected a room temperature neutron diffraction dataset under functionally relevant acidic conditions following vapor-exchange. Alteration of the *in crystallo* pH has the potential to provide information on protein titratable groups and their underlying function, however a direct buffer soak may result in crystal damage. Vapor-exchange therefore represents a gentler technique to modify the protein environment and opens the possibility to collect multiple datasets from a single crystal, ensuring consistency across experiments and mitigating loss of sample. This work further illustrates that well diffracting protein crystals can be salvaged from cracked samples. This is an important consideration since large crystals of sufficient size and quality for neutron diffraction are often challenging to obtain.

2.2.7 Acknowledgements

Protein expression, purification and crystallization experiments were conducted at the Center for Structural Molecular Biology (CSMB), a U.S. Department of Energy Biological and Environmental Research User Facility at Oak Ridge National Laboratory. Neutron diffraction data were collected at BL-11B MaNDi at the Spallation Neutron Source (SNS) and at CG-4D IMAGINE at the High Flux Isotope Reactor (HFIR), both at ORNL and sponsored by the Scientific User Facilities Division, Office of Basic Energy Sciences, U.S. Department of Energy. The authors thank Brendan Sullivan for assistance with data reduction. X-ray diffraction data were collected at the Molecular Education, Technology, and Research Innovation Center (METRIC) facilities at North Carolina State University, which is supported by the State of North Carolina.

2.2.8 Funding Information

GCS acknowledges support in part from the National Research Foundation (NRF), South Africa and from the Oak Ridge National Laboratory Graduate Opportunities (GO!) Program. FM acknowledges support from USDA NIFA Hatch 211001.

2.2.9 References

1. Lynd, L. R., Weimer, P. J., Zyl, W. H. Van & Isak, S. Microbial Cellulose Utilization : Fundamentals and Biotechnology Microbial Cellulose Utilization : Fundamentals and Biotechnology Downloaded from <http://mibr.asm.org/> on February 6 , 2013 by INDIAN INSTITUTE OF TECHNOLOGY MADRAS. *Microbiol. Mol. Biol. Rev.* **66**, 506–577 (2002).
2. Agostoni, M., Hangasky, J. A. & Marletta, M. A. Physiological and Molecular Understanding of Bacterial Polysaccharide Monooxygenases. *Microbiol. Mol. Biol. Rev.* **81**, 1–16 (2017).
3. Harris, P. V., Xu, F., Kreel, N. E., Kang, C. & Fukuyama, S. New enzyme insights drive advances in commercial ethanol production. *Curr. Opin. Chem. Biol.* **19**, 162–170 (2014).
4. Cragg, S. M. *et al.* Lignocellulose degradation mechanisms across the Tree of Life. *Curr. Opin. Chem. Biol.* **29**, 108–119 (2015).
5. Himmel, M. E. *et al.* Biomass recalcitrance: Engineering plants and enzymes for biofuels production. *Science*. **315**, 804–807 (2007).
6. Lopes, A. M., Ferreira Filho, E. X. & Moreira, L. R. S. An update on enzymatic cocktails for lignocellulose breakdown. *J. Appl. Microbiol.* **125**, 632–645 (2018).
7. Vaaje-Kolstad, G. *et al.* An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* **330**, 219–22 (2010).
8. Beeson, W. T., Phillips, C. M., Cate, J. H. D. & Marletta, M. A. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J. Am. Chem. Soc.* **134**, 890–892 (2012).
9. Quinlan, R. J. *et al.* 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 (2011).
10. Li, X., Beeson, W. T., Phillips, C. M., Marletta, M. A. & Cate, J. H. D. Structural Basis for Substrate Targeting and Catalysis by Fungal Polysaccharide Monooxygenases. *Structure* **20**, 1051–1061 (2012).
11. Frommhagen, M., Westphal, A. H., van Berkel, W. J. H. & Kabel, M. A. Distinct Substrate Specificities and Electron-Donating Systems of Fungal Lytic Polysaccharide Monooxygenases. *Front. Microbiol.* **9**, 1–22 (2018).
12. Kjaergaard, C. H. *et al.* Spectroscopic and computational insight into the activation of O₂ by the mononuclear Cu center in polysaccharide monooxygenases. *Proc. Natl. Acad. Sci.* **111**, 8797–8802 (2014).
13. Hangasky, J. A., Detomasi, T. C. & Marletta, M. A. Glycosidic Bond Hydroxylation by Polysaccharide Monooxygenases. *Trends Chem.* **1**, 198–209 (2019).
14. Phillips, C. M., Beeson, W. T., Cate, J. H. & Marletta, M. A. Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem. Biol.* **6**, 1399–1406 (2011).
15. Walton, P. H. & Davies, G. J. On the catalytic mechanisms of lytic polysaccharide monooxygenases. *Curr. Opin. Chem. Biol.* **31**, 195–207 (2016).
16. Kim, S., Ståhlberg, J., Sandgren, M., Paton, R. S. & Beckham, G. T. Quantum mechanical calculations suggest that lytic polysaccharide monooxygenases use a copper-oxyl, oxygen-rebound mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 149–154 (2014).
17. Lee, J. Y. & Karlin, K. D. Elaboration of copper-oxygen mediated CH activation chemistry in consideration of future fuel and feedstock generation. *Curr. Opin. Chem. Biol.* **25**, 184–193 (2015).
18. Wang, B. *et al.* QM/MM Studies into the H₂O₂-Dependent Activity of Lytic Polysaccharide Monooxygenases: Evidence for the Formation of a Caged Hydroxyl Radical Intermediate. *ACS Catal.* **8**, 1346–1351 (2018).
19. Wang, B., Wang, Z., Davies, G. J., Walton, P. H. & Rovira, C. Activation of O₂ and H₂O₂ by Lytic Polysaccharide Monooxygenases. *ACS Catal.* **10**, 12760–12769 (2020).
20. Dhar, D. & Tolman, W. B. Hydrogen atom abstraction from hydrocarbons by a copper(III)-hydroxide complex. *J. Am. Chem. Soc.* **137**, 1322–1329 (2015).
21. Neisen, B. D., Gagnon, N. L., Dhar, D., Spaeth, A. D. & Tolman, W. B. Formally Copper(III)-

- Alkylperoxo Complexes as Models of Possible Intermediates in Monooxygenase Enzymes. *J. Am. Chem. Soc.* **139**, 10220–10223 (2017).
22. Hedegård, E. D. & Ryde, U. Molecular mechanism of lytic polysaccharide monooxygenases. *Chem. Sci.* **9**, 3866–3880 (2018).
 23. Bertini, L. *et al.* Catalytic Mechanism of Fungal Lytic Polysaccharide Monooxygenases Investigated by First-Principles Calculations. *Inorg. Chem.* **57**, 86–97 (2018).
 24. Hedegård, E. D. & Ryde, U. Multiscale Modelling of Lytic Polysaccharide Monooxygenases. *ACS Omega* **2**, 536–545 (2017).
 25. Hedegård, E. D. & Ryde, U. Targeting the reactive intermediate in polysaccharide monooxygenases. *J. Biol. Inorg. Chem.* **22**, 1029–1037 (2017).
 26. Span, E. A. & Marletta, M. A. The framework of polysaccharide monooxygenase structure and chemistry. *Curr. Opin. Struct. Biol.* **35**, 93–99 (2015).
 27. O'Dell, W. B., Agarwal, P. K. & Meilleur, F. Oxygen Activation at the Active Site of a Fungal Lytic Polysaccharide Monooxygenase. *Angew. Chemie - Int. Ed.* **56**, 767–770 (2017).
 28. Bacik, J. P. *et al.* Neutron and Atomic Resolution X-ray Structures of a Lytic Polysaccharide Monooxygenase Reveal Copper-Mediated Dioxygen Binding and Evidence for N-Terminal Deprotonation. *Biochemistry* **56**, 2529–2532 (2017).
 29. Frandsen, K. E. H. *et al.* The molecular basis of polysaccharide cleavage by lytic polysaccharide monooxygenases. *Nat. Chem. Biol.* **12**, 298–303 (2016).
 30. Simmons, T. J. *et al.* Structural and electronic determinants of lytic polysaccharide monooxygenase reactivity on polysaccharide substrates. *Nat. Commun.* **8**, (2017).
 31. Tandrup, T. *et al.* Oligosaccharide Binding and Thermostability of Two Related AA9 Lytic Polysaccharide Monooxygenases. *Biochemistry* **59**, 3347–3358 (2020).
 32. Schatz, M. *et al.* Combined spectroscopic and theoretical evidence for a persistent end-on copper superoxo complex. *Angew. Chemie - Int. Ed.* **43**, 4360–4363 (2004).
 33. Prigge, S. T., Eipper, B. A., Mains, R. E. & Amzel, L. M. Dioxygen binds end-on to mononuclear copper in a precatalytic enzyme complex. *Science* **304**, 864–7 (2004).
 34. O'Dell, W. B., Bodenheimer, A. M. & Meilleur, F. Neutron protein crystallography: A complementary tool for locating hydrogens in proteins. *Arch. Biochem. Biophys.* **602**, 48–60 (2016).
 35. Schröder, G. C., O'Dell, W. B., Myles, D. A. A., Kovalevsky, A. & Meilleur, F. IMAGINE: Neutrons reveal enzyme chemistry. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, (2018).
 36. Schröder, G. C. & Meilleur, F. Neutron Crystallography Data Collection and Processing for Modelling Hydrogen Atoms in Protein Structures. *J. Vis. Exp.* 1–38 (2020). doi:10.3791/61903
 37. Helliwell, J. R. *Fundamentals of neutron crystallography in structural biology. Methods in Enzymology* **634**, (Elsevier Inc., 2020).
 38. Kwon, H., Langan, P. S., Coates, L., Raven, E. L. & Moody, P. C. E. The rise of neutron cryocrystallography. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 792–799 (2018).
 39. Schüle, M. Enzymatic properties of cellulases from *Humicola insolens*. *J. Biotechnol.* **57**, 71–81 (1997).
 40. Boer, H. & Koivula, A. The relationship between thermal stability and pH optimum studied with wild-type and mutant *Trichoderma reesei* cellobiohydrolase Cel7A. *Eur. J. Biochem.* **270**, 841–848 (2003).
 41. Meilleur, F., Coates, L., Cuneo, M. J., Kovalevsky, A. & Myles, D. A. A. The neutron macromolecular crystallography instruments at Oak Ridge national laboratory: Advances, challenges, and opportunities. *Crystals* **8**, 1–10 (2018).
 42. O'Dell, W. B., Swartz, P. D., Weiss, K. L. & Meilleur, F. Crystallization of a fungal lytic polysaccharide monooxygenase expressed from glycoengineered *Pichia pastoris* for X-ray and neutron diffraction. *Acta Crystallogr. Sect. Struct. Biol. Commun.* **73**, 70–78 (2017).
 43. Meilleur, F., Kovalevsky, A. & Myles, D. A. A. IMAGINE: The neutron protein crystallography beamline at the high flux isotope reactor. in *Methods in enzymology* **634**, 69–85 (Elsevier Inc., 2020).

44. Coates, L. & Sullivan, B. *The macromolecular neutron diffractometer at the spallation neutron source. Methods in Enzymology* **634**, (Elsevier Inc., 2020).
45. Arnold, O. *et al.* Mantid - Data analysis and visualization package for neutron scattering and μ SR experiments. *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* **764**, 156–166 (2014).
46. Sullivan, B. *et al.* Improving the accuracy and resolution of neutron crystallographic data by three-dimensional profile fitting of Bragg peaks in reciprocal space. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 1085–1095 (2018).
47. Helliwell, J. R. *et al.* The recording and analysis of synchrotron X-radiation Laue diffraction photographs. *J. Appl. Crystallogr.* **22**, 483–497 (1989).
48. Campbell, J. W., Hao, Q., Harding, M. M., Nguti, N. D. & Wilkinson, C. LAUEGEN version 6.0 and INTLDM. *J. Appl. Crystallogr.* **31**, 496–502 (1998).
49. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
50. Winn, M. D. *et al.* Overview of the CCP 4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 235–242 (2011).
51. Ashkar, R. *et al.* Neutron scattering in the biological sciences: progress and prospects. *Acta Crystallographica Section D: Structural Biology* **74**, (2018).
52. Logan, D. T. Interactive model building in neutron macromolecular crystallography. in *Methods in Enzymology* **634**, 201–224 (Elsevier Inc., 2020).
53. Michalczyk, R. *et al.* Joint neutron crystallographic and NMR solution studies of Tyr residue ionization and hydrogen bonding: Implications for enzyme-mediated proton transfer. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 5673–5678 (2015).
54. Gerlits, O. *et al.* Long-Range Electrostatics-Induced Two-Proton Transfer Captured by Neutron Crystallography in an Enzyme Catalytic Site. *Angew. Chemie - Int. Ed.* **55**, 4924–4927 (2016).
55. Li, L. *et al.* Neutron crystallographic studies of T4 lysozyme at cryogenic temperature. *Protein Sci.* **26**, 2098–2104 (2017).
56. Harp, J. M., Coates, L., Sullivan, B. & Egli, M. Cryo-neutron crystallographic data collection and preliminary refinement of left-handed Z-DNA d(CGCGCG). *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **74**, 603–609 (2018).
57. Hiromoto, T. *et al.* Towards cryogenic neutron crystallography on the reduced form of [NiFe]-hydrogenase. *Acta Crystallogr. Sect. D Struct. Biol.* **76**, 946–953 (2020).

2.3 Neutron cryo-crystallography captures activated dioxygen intermediates at the copper-active site of a lytic polysaccharide monooxygenase

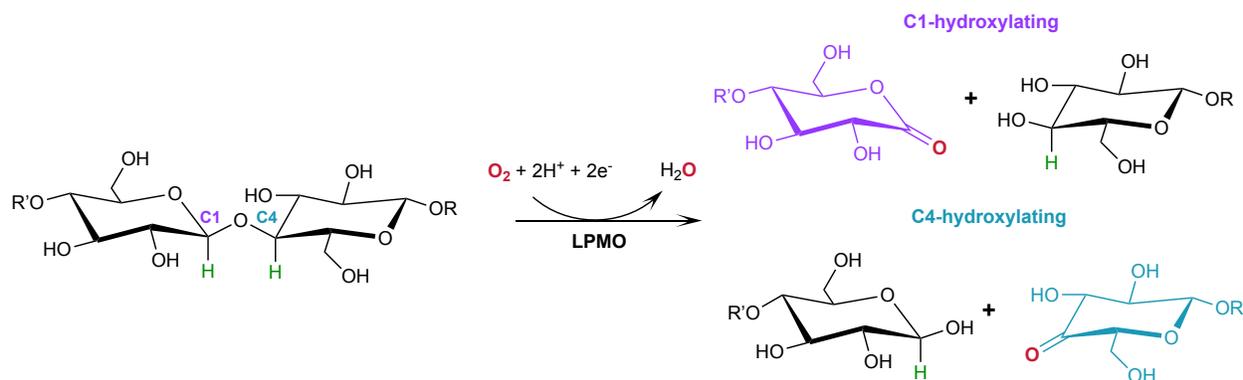
G.C. Schröder, W.B. O'Dell, S.P. Webb, P.K. Aggarwal and F. Meilleur

2.3.1 Abstract

Metalloproteins perform a diverse array of redox-related reactions facilitated by the increased chemical functionality afforded by their metallocofactors. Lytic polysaccharide monooxygenases (LPMOs) are a class of copper-dependent enzymes that are responsible for the breakdown of recalcitrant polysaccharides *via* oxidative cleavage at the glycosidic bond. The activated copper-oxygen intermediates and their mechanism of formation remains to be established. Neutron protein crystallography which permits direct visualization of protonation states was used to investigate the initial steps of oxygen activation directly following active site copper reduction in *Neurospora crassa* LPMO9D. Herein, we cryo-trap a stabilized activated dioxygen intermediate in a mixture of superoxo and hydroperoxo states, and we identify the conserved second coordination shell residue His157 as the proton donor. Density functional theory calculations indicate that both active site states are feasible. The hydroperoxo formed is potentially an intermediate in the mechanism of hydrogen peroxide formation in the absence of substrate. We establish that the N-terminal amino group of the copper coordinating His1 remains doubly protonated directly following molecular oxygen reduction by copper, ruling it out as proton donor during early intermediate formation. Aided by mining minima free energy calculations we establish His157 conformational flexibility in solution that is abolished in the crystal due to the close packing of the two molecules in the asymmetric unit. A neutron crystal structure of *Nc*LPMO9D at low pH supports occlusion of the active site which prevents protonation due to binding-face to binding-face crystal packing by indicating a neutral His157 at these conditions.

2.3.2 Introduction

Lytic polysaccharide monooxygenases (LPMOs; also designated PMOs) are copper metalloenzymes that perform unique redox chemistry resulting in disruption of polysaccharide chains by oxygen atom insertion at the glycosidic bond.¹ The LPMO active site contains a mononuclear Cu^{2+} center within a “histidine brace” motif in which the copper is coordinated by the N-terminal histidine amino group, the N_δ atom of His1 and the N_ϵ atom of a second conserved histidine.² Of particular commercial and biotechnological interest is the ability of fungal LPMOs to catalyze the oxidative cleavage of recalcitrant crystalline cellulose at the C1 and/or C4 position, thereby enhancing polysaccharide depolymerization by increasing glucan access to hydrolytic enzymes via the overall reaction shown in Scheme 1.³⁻⁵



Scheme 1 Oxidation of the C1/C4 carbon in the glycosidic bond catalyzed by LPMOs with oxygen as the co-substrate. H atoms to be abstracted by the reactive oxygen species shown in green.

While H_2O_2 has also been shown to be a LPMO co-substrate,⁶ the work presented here focuses on the activation of O_2 . One-electron reduction of the copper active site to Cu^{1+} and binding of the molecular oxygen co-substrate initiates the reaction cycle with the generally accepted formation of a Cu^{2+} superoxide, $[\text{Cu}-\text{O}_2]^+$.^{3,4,7,8} In the current consensus for the O_2 -based mechanism, following the rapid reduction of molecular dioxygen to the superoxide, the LPMO reaction mechanism requires further delivery of electrons and protons to form a reactive oxygen intermediate which oxidizes the polysaccharide substrate.⁹ The identity of the activated oxygen species responsible for hydrogen atom abstraction (HAA) from the polysaccharide substrate leading to subsequent substrate oxidation, is however still a matter of ongoing

research.¹⁰ Intensive studies encompassing X-ray crystallographic, computational and spectroscopic examinations of mononuclear copper monooxygenases as well as biomimetic copper complexes have proposed dioxygen intermediates such as the superoxide $[\text{Cu-O}_2]^+$ and hydroperoxo $[\text{CuOOH}]^+$ species where the O–O bond remains intact.^{8,11–14} Furthermore, formation of the hydroxide $[\text{CuOH}]^{2+}$ and oxyl $[\text{CuO}]^+$ intermediates after O–O bond scission, with subsequent involvement in HAA, have been proposed.^{15–23} Formation of a hydroperoxo, hydroxy or oxyl HAA species would require further protonation and reduction of the superoxide however the identity of the proton donor is still debated. Several second-shell residues including histidine (His147 in *LsLPMO9A*/His 161 in *MtLPMO3**) as well as Gln 167 in *MtLPMO3** and Glu 201 from *JdLPMO10A* have been proposed to be involved in intermediate protonation.^{22,24,25} Quantum mechanics/molecular mechanics (QM/MM) calculations have also shown that a hydrogen atom could be abstracted by the superoxide species from an exogenous reducing agent such as ascorbic acid.²⁶

Neutron protein crystallography is particularly well suited to the determination of protonation states within a protein due to the technique's unique sensitivity to hydrogen atoms and protons often invisible to X-ray diffraction.^{27,28} We report here two neutron models of a carbohydrate-free C4-hydroxylating *Neurospora crassa* LPMO9D: a structure of the activated, ascorbate-reduced form and a structure of the resting state at pH 4.4 (pD 4.8) to further address open questions regarding initial molecular oxygen activation. The neutron crystallographic structure of the reduced *NcLPMO9D* shows cryo-capture of activated dioxygen at the copper active site modelled as a superoxo and protonated hydroperoxo species with a population ratio of 0.30:0.70. The neutron scattering length density maps of the conserved secondary coordination shell His157 is best represented as 70% imidazolate and a 30% singly protonated form. The copper coordinating N-terminal amino group His1 remains in a doubly protonated ND_2 state following copper reduction and molecular oxygen activation. Mining minima energy calculations reveal His157 flexibility in solution which is absent in the crystal structure due to binding-face to binding-face crystal

packing. Furthermore, the neutron and X-ray crystal structures of the resting state *NcLPMO9D* at low pH show that His157 remains singly protonated when exposed to acidic conditions.

2.3.3 Results and discussion

2.3.3.1 Crystallographic features of the *NcLPMO9D* neutron structures

Both the ascorbate-reduced and low pH neutron crystals of *NcLPMO9D* belonged to the monoclinic $P2_1$ space group. The crystal structure includes two protein molecules per asymmetric unit (Molecule A and Molecule B) related by non-crystallographic symmetry (NCS). Structural analysis of the packing in our *NcLPMO9D* protein crystal shows that the two NCS related molecules pack with their planar substrate binding interfaces against each other in a binding-face to binding-face configuration also seen in PDB structures 4EIR, 5TKG, 5TKH and 5TKI (Figure 2.12 A).^{7,29,30} While present in a dimeric conformation in the crystal structure, LPMOs including *NcLPMO9D* function as single-domain monomeric proteins in solution, where they associate with planar crystalline cellulose as confirmed by early LPMO structural studies,^{7,31,32} and single-molecule observations of LPMO using atomic force microscopy (AFM).³³ Furthermore, within this binding-face to binding-face packing architecture, the two coordinated coppers are ~ 12 Å apart, precluding direct interaction. Notably, the symmetry of the two molecules in the asymmetric unit is broken by Tyr25, Tyr206 and Trp206, which have asymmetric conformations in Molecule A and B, contributing to different microenvironments near the active site (Figure 2.12 B). Structural alignment of our *NcLPMO9D* Molecule A or Molecule B with the oligosaccharide-bound form of *LsLPMO9A* (PDB 5ACI),³⁴ indicates that the substrate pyranose rings of the *LsLPMO9A* structure superimpose with the proline rings Pro43 and Pro163 for both Molecule A or Molecule B of *NcLPMO9D* (Figure 2.12 C). The alignment further indicates that Tyr206 of Molecule A and Tyr162 of Molecule B are positioned to interact with substrate *via* their phenol rings. This suggests that the binding-face to binding-face packing in our crystal mimics surface residue interaction with the pyranose ring of the glucose units and creates an occluded active site, possibly analogous to the bulk solvent-protected environment created when the planar active site of LPMO binds to its crystalline linear polysaccharide substrate. Such crystal packing in which

active site residues mimic carbohydrate substrate pyranose rings has also been observed in the crystal packing of *Nc*LPMO-3 (PDB 4EIS).⁷ Representation of the protein surface additionally illustrates the active site occlusion due to crystal packing (Figure 2.13).

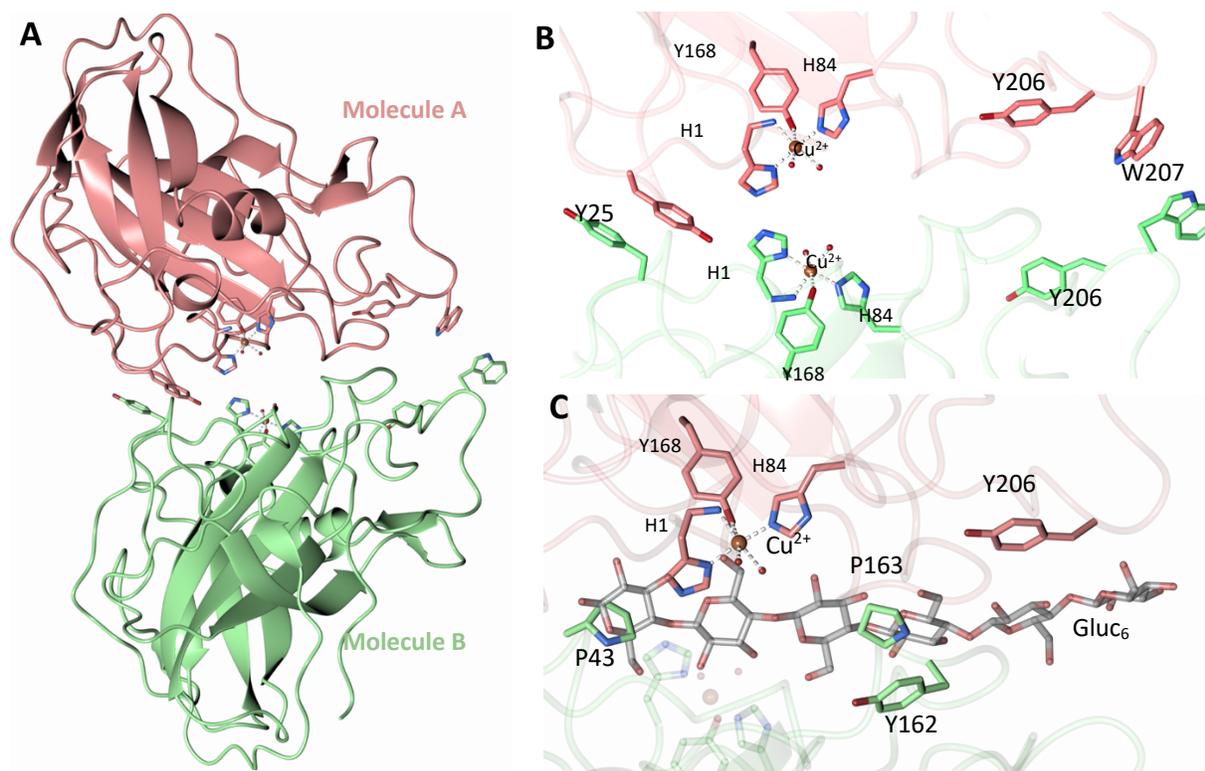


Figure 2.12 Crystal packing of the Molecule A and B in the LPMO9D crystals. **(A)** Cartoon of Molecule A (coral) and Molecule B (light green) of the planar active site surfaces packed facing each other. **(B)** Crystal symmetry in Molecule A and B is broken by alternate conformations of Tyr 25, Tyr206 and Trp207. **(C)** Alignment of Molecule B of *Nc*LPMO9D with the Gluc₆ bound *Ls*LPMO9A (PDB 5ACI) in grey to illustrate superposition of proline residues with glucose pyranose rings.

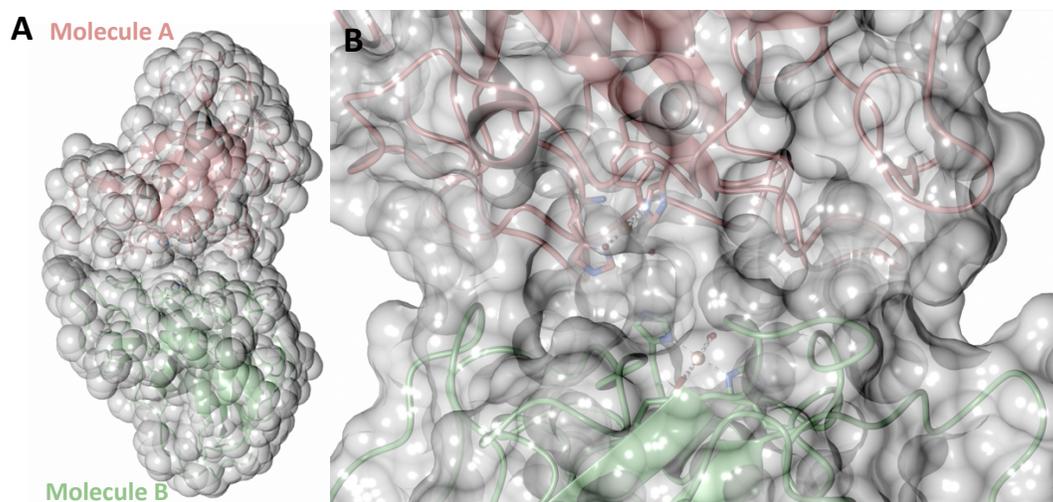


Figure 2.13 Solvent accessible surface in the crystal packing of Molecule A and Molecule B of the asymmetric unit of LPMO9D.

2.3.3.2 Ascorbate-reduced copper active site

To probe the protonation states of key catalytic residues and intermediates in the *Nc*LPMO9D active site, a neutron diffraction dataset to 2.4 Å resolution was collected at 100 K on a crystal of *Nc*LPMO9D reduced with ascorbate at pD 6.0 (pH 5.6) in the presence of atmospheric oxygen. Structure refinement was performed solely against neutron data to circumvent potential artefacts induced by X-ray radiation damage. The axial position is vacant, supporting displacement of both the equatorial and axial water molecules upon Cu^{2+} reduction. The neutron scattering length density maps at the active site of Molecule A indicate the presence a fully occupied equatorially coordinated dioxygen species in a mixed state two-atom superoxo species and a three-atom hydroperoxo species with occupancies of 0.30 and 0.70, respectively (Figure 2.14 A and 2.14 B). The intermediates have η_1 end-on geometry at a Cu–O1 distance of ~ 1.96 Å and ~ 1.98 Å for the superoxo and hydroperoxo species, respectively. The superoxo species is modelled with an O–O bond length of 1.28 Å while the hydroperoxo species is modelled with a bond length of 1.46 Å. The Cu–O1–O2 angles are $\sim 147^\circ$ and 151° for the superoxo and hydroperoxo species respectively. A Glu30 residue from Molecule B is modeled 2.06 Å and 2.08 Å away from the O2 of the superoxo and hydroperoxo species, respectively, however it displays a disordered conformation with limited NSLD, as has been observed for *Nc*LPMO9D structures with the same crystal packing.^{7,29,30} The

modeling of an activated dioxygen species is in good accordance with our previous high resolution X-ray structures which indicated the presence of a peroxo species with a 1.44 ± 0.06 Å bond length with a Cu-O1-O2 angle of 140.52° and 1.90 Å from the Cu following ascorbate reduction (Figure 2.14 C).²⁹ Activated dioxygen species have been observed in further LPMO structures including *NcLPMO9D* and *NcPMO-3* from Li *et al.*, *JdLPMO10A* from Bacik *et al.* as well as in an artificial copper protein from Mann *et al.* (Table 2.3).^{7,12,25,29,35} In our previous high-resolution X-ray structure, discrimination between a putative superoxo and peroxo species was based on bond length determination due to the lack of visibility of hydrogen atoms. In addition, the occupancy of the activated dioxygen species only refined to 0.59 and an axial water was present with a 0.48 occupancy. We reason that photoreduction due to X-ray beam exposure resulted in presence of intermediates with low occupancy. The *NcLPMO9D* active site presented here is absent of radiation damage since neutron protein crystallography represents a non-destructive technique permitting capture of a mixed occupancy of a superoxo and hydroperoxo species.²⁸ The observed hydroperoxo species may represent an intermediate in the mechanism of LPMO H₂O₂ formation, since LPMOs are known to produce H₂O₂ in the absence of substrate.³⁶ Caldararu *et al.* have further shown it is more energetically favorable for H₂O₂ formation to proceed by dissociation of H₂O₂ from the copper active site than HO₂.²⁵ This higher dissociation energy of HO₂ is supported by the observed binding stability of the hydroperoxo species in our structure. The Molecule B copper active site contains a water molecule coordinated in the axial position with a partial occupancy of 0.65 and a Cu-O distance of ~ 2.2 Å while the equatorial position is vacant (Figure 2.14 D). This geometry, while not chemically relevant, has been observed in partially photo-reduced LPMO structures,³⁷ suggesting that the active site of Molecule B underwent only partial reduction during the ascorbate soak.

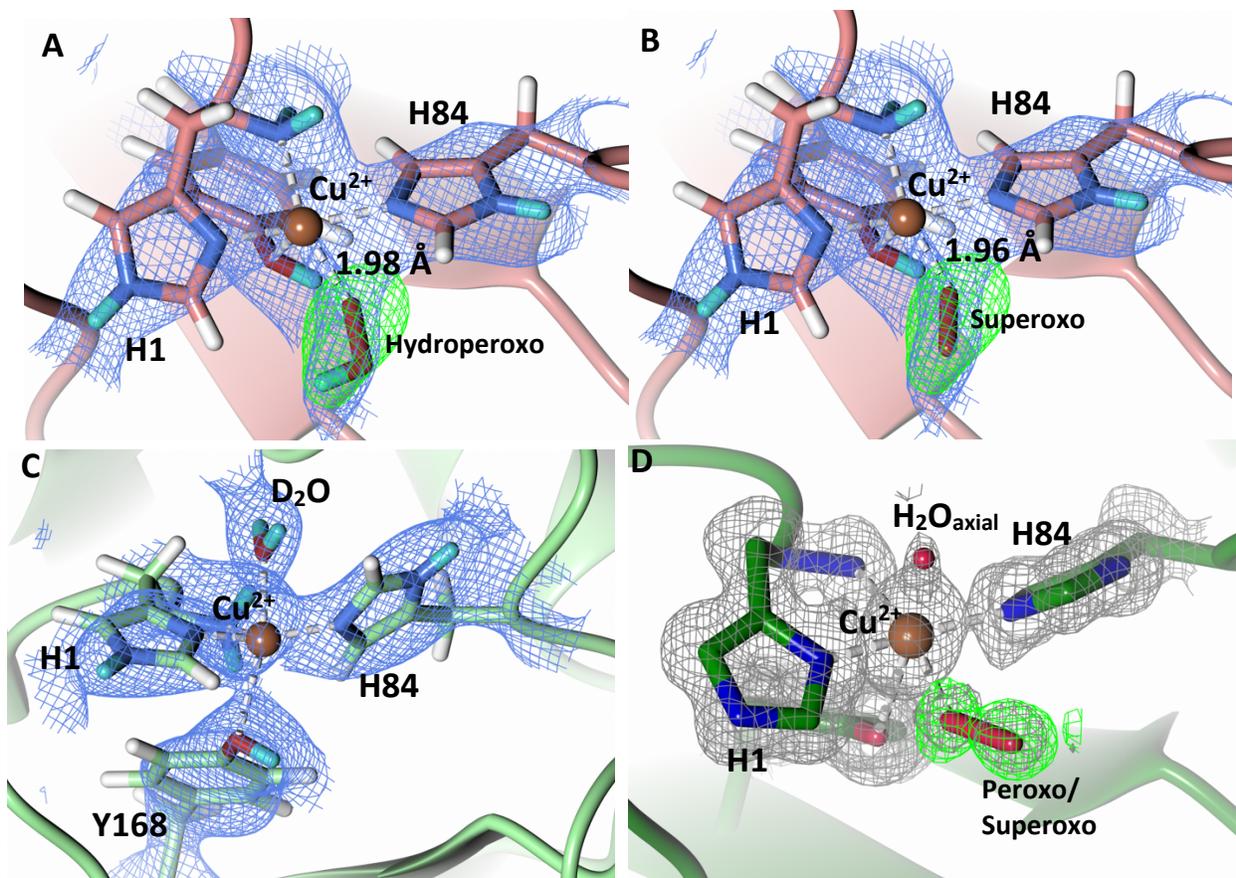


Figure 2.14 Neutron diffraction structure of the ascorbate reduced LPMO9D crystal. **(A)** Active site of Molecule A (coral) with a hydroperoxo species bound at the equatorial position. **(B)** Active site of Molecule A (coral) with a superoxo species bound at the equatorial position. **(C)** Active site of Molecule B (light green) with a water bound at the axial position. **(D)** Equatorially coordinated activated dioxygen superoxo/peroxo in X-ray diffraction structure of ascorbate reduced LPMO9D (green) by O'Dell *et al.*²⁹ (PDB 5TKH). Neutron scattering length density $2F_o-F_c$ maps displayed in blue at a 1.0σ cut-off. Neutron scattering length density F_o-F_c maps displayed in green at a 2.5σ cut-off. H atoms displayed in white and D atoms displayed in turquoise. Electron density $2F_o-F_c$ maps displayed in grey at a 1.0σ cut-off. Electron density F_o-F_c maps displayed in green at a 3.0σ cut-off.

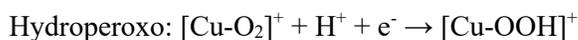
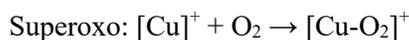
Table 2.3 Structurally characterized dioxygen species in LPMOs.

LPMO	PDB	Dioxygen species	Position	Cu-O (Å)	O1-O2 (Å)	Cu-O-O (°)
<i>Nc</i> LPMO9D	4EIR - Chain A	Superoxide – end-on	Axial	2.96	1.16	147.75
	4EIR- Chain B	Superoxide – end-on	Axial	2.92	1.15	142.11
<i>Nc</i> PMO-3	4EIS	Peroxo – end-on	Axial	3.44	1.49	117.16
<i>Jd</i> LPMO10A	5VG0 – Chain A	Peroxo – side-on	Equatorial	O1 – 2.14 O2 – 1.84	1.48	O1 – 57.67 O2 – 79.39
<i>Jd</i> LPMO10A	5VG0 – Chain A	Peroxo – end-on	Equatorial	1.83	1.46	109.52
<i>Jd</i> LPMO10A	5VG1 – Chain B	Peroxo – end-on	Equatorial	2.49	1.45	115.32
	5VG0/5VG1 joint quantum refinement (ComQum-X) – Chain A	Superoxide – end-on	Equatorial	2.31	1.25	108.6
	5VG0/5VG1 joint quantum refinement (ComQum-X) – Chain B	Superoxide – end-on	Equatorial	2.27	1.25	113.6
<i>Nc</i> LPMO9D	5TKH – Chain A	Peroxo – end-on	Equatorial	1.90	1.44	140.52
Artificial copper protein	6ANX	Hydroperoxo	Intermediate	1.94	1.52	137.60
<i>Nc</i> LPMO9D (current neutron structure)	Chain A	Superoxo	Equatorial	1.96	1.28	147
	Chain A	Hydroperoxo	Equatorial	1.98	1.46	151

2.3.3.3 Protonation state of His157

Examination of the $2F_o-F_c$ and F_o-F_c omit neutron scattering length density maps of the Molecule A secondary coordination shell His157 suggest that this residue is best modelled by the neutral N_ϵ -protonated form and the imidazolate form with occupancies of 0.30 and 0.70, respectively (Figure 2.15 A and 2.15 B). Analysis of the Molecule B neutron scattering length density maps indicate that His157 is

singly N_ε-protonated at full occupancy (Figure 2.15 C). The N_ε nitrogen of His157 in Molecule A is observed at a distance of ~3.25 Å from the distal oxygen O2 of the equatorially bound superoxo species and ~2.86 Å from the hydroperoxo species. These findings indicate that His157 plays a role in protonation of the activated dioxygen species observed in the barricaded active site presented here. Activity and structural studies of *Mt*PMO3* as well as QM/MM calculations of *Ls*LPMO9A have implicated this conserved histidine residue in proton transfer to the copper-bound superoxide during LPMO catalysis (His161 and His147 in *Mt*PMO3* and *Ls*LPMO9A, respectively).^{22,24} QM/MM studies have also indicated that the Glu201 residue near the copper active site in *Jd*LPMO10A is capable of donating a proton to a bound superoxo to form a hydroperoxo species.²⁵ Computational studies have suggested that ascorbate may function as a proton donor to the superoxo species;²⁶ however, the active site occlusion due to crystal packing in our structure does not permit ascorbate to access the copper active site, and there was no evidence in the NSLD maps for an ascorbate molecule near the active site. Following formation of the superoxide in *Nc*LPMO9D, protonation by His157 may form the copper(II)-hydroperoxyl species [CuOOH]²⁺, the copper(III)-hydroperoxo, or upon further reduction a copper(II)-hydroperoxo [CuOOH]⁺, all of which have been posited as viable reaction intermediates.^{12,23,38,39}



While the neutron data supports modeling of a two and three atom dioxygen species at the active site, it does however not permit definitive discrimination between the hydroperoxyl or hydroperoxo species nor does it resolve the redox state of the copper ion. We performed preliminary DFT studies on active site models to further assess the intermediates modelled in the copper active site of Molecule A (Table 2.4). The calculations indicate that the N_ε-protonated His157 plus superoxo species and the imidazolate plus hydroperoxo species are both feasible states in that they both remain bound when geometry optimized to a

stationary point on the potential energy surface; however, they each have a different combination of charge and copper spin state precluding a direct comparison of their DFT calculated energies. The second reduction to form the hydroperoxo species can potentially be derived from ascorbate by long-range intramolecular electron transfer to the active site as has been proposed to occur via LPMO aromatic residues localized near the copper active site.^{5,7,40,41} An accurate calculated energy difference between the two states will require a multi-determinant QM treatment due to propensity of copper centers to exhibit multiple near degenerate electronic states. While beyond the scope of the current paper, this will be addressed in planned future calculations using QM/MM with a multi-determinant wavefunction QM method. The energy levels of different models used for the DFT calculations cannot be directly compared. These models, which have different charges and spin states, represents points on different energetic surfaces and a direct comparison of energy is not possible as they have different number of electrons. A direct comparison of the energetic differences would require QM/MM calculations with same number of electrons, which can be pursued in future studies.

Table 2.4 Stability and electronic state of DFT models.

Model	DFT active site residues	Starting proton position	System net charge	Multiplicity	Stability	Final state
Superoxo, His157 NE2-protonated	His1, His84, His157, Gln166, Tyr168	His157 NE2	+1	2	Stable	His157 singly NE2-protonated
Hydroperoxyl radical, His157 doubly deprotonated	His1, His84, His157, Gln166, Tyr168	O-O-H	+1	2	Stable	Double deprotonated His157
Hydroperoxo, His157 doubly deprotonated	His1, His84, His157, Gln166, Tyr168	O-O-H	0	2	Stable	Double deprotonated His157
Hydroperoxo, His157 doubly deprotonated	His1, His84, His157, Gln166, Tyr168	O-O-H	-1	1	Stable	Double deprotonated His157

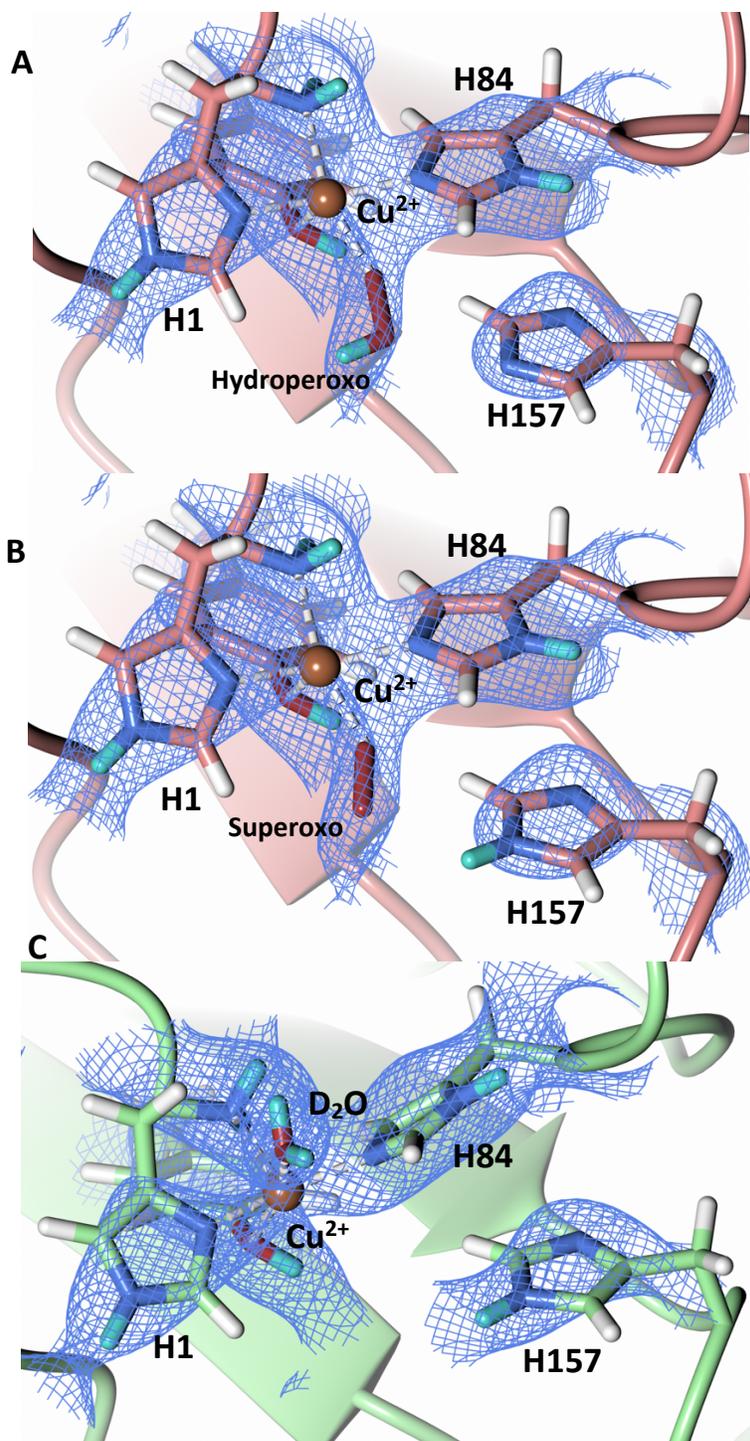


Figure 2.15 Mixed protonation state His157 in ascorbate reduced LPMO9D. **(A)** The Molecule A (coral) second shell His157 is deprotonated with an equatorial hydroperoxo species coordinated to Cu^{2+} . **(B)** The Molecule A second shell His157 is singly N_ϵ -protonated with an equatorial superoxo species coordinated to Cu^{2+} . **(C)** The Molecule B (light green) second shell His157 is singly N_ϵ -protonated with an axial water coordinated to Cu^{2+} . Neutron scattering length density $2F_o - F_c$ maps displayed in blue at a 1.0σ cut-off. H atoms displayed in white and D atoms displayed in turquoise.

2.3.3.4 N-terminal protonation following initial dioxygen activation

The N-terminal amino group of the histidine brace has been postulated to play a role in intermediate formation and stabilization during LPMO catalysis.⁴² A previous neutron protein crystallography study of unreduced *JdLPMO10A*, a bacterial LPMO, suggested that the amino terminal is present as a mixture of the -ND^\cdot and -ND_2 protonation states.³⁵ The *JdLPMO10A* neutron structure was, however, later revisited and a joint X-ray–neutron quantum refinement concluded that the amino terminal is present solely in the -ND_2 state.²⁵ We therefore sought to determine the protonation state of the N-terminal amino group upon *NcLPMO9D* reduction by ascorbate in the presence of atmospheric O_2 . The neutron scattering length density map indicates that both Molecule A and Molecule B contain the amino terminal in the -ND_2 protonation state. F_o-F_c omit difference maps calculated confirm the -ND_2 state in both Molecule A and Molecule B (Figure 2.16 A – D). The role of possible deprotonation of the N-terminal histidine amino group came to the forefront when it was proposed to promote reactive intermediate formation and stabilization based on findings from a small molecule copper complex.⁴³ Structural and spectroscopic studies of a substrate-bound *LsLPMO9A* suggested that the two protons of the N-terminus are exposed to different chemical environments, with one being involved in a hydrogen bond network, which potentially promotes N-terminal deprotonation during LPMO catalysis.³⁴ Our neutron protein crystallography structures show that the N-terminal amine is not involved in protonation upon initial Cu^{2+} reduction and molecular dioxygen activation. However, a mechanistic role for the N-terminal amine in other contexts cannot be conclusively ruled out by our experiments.

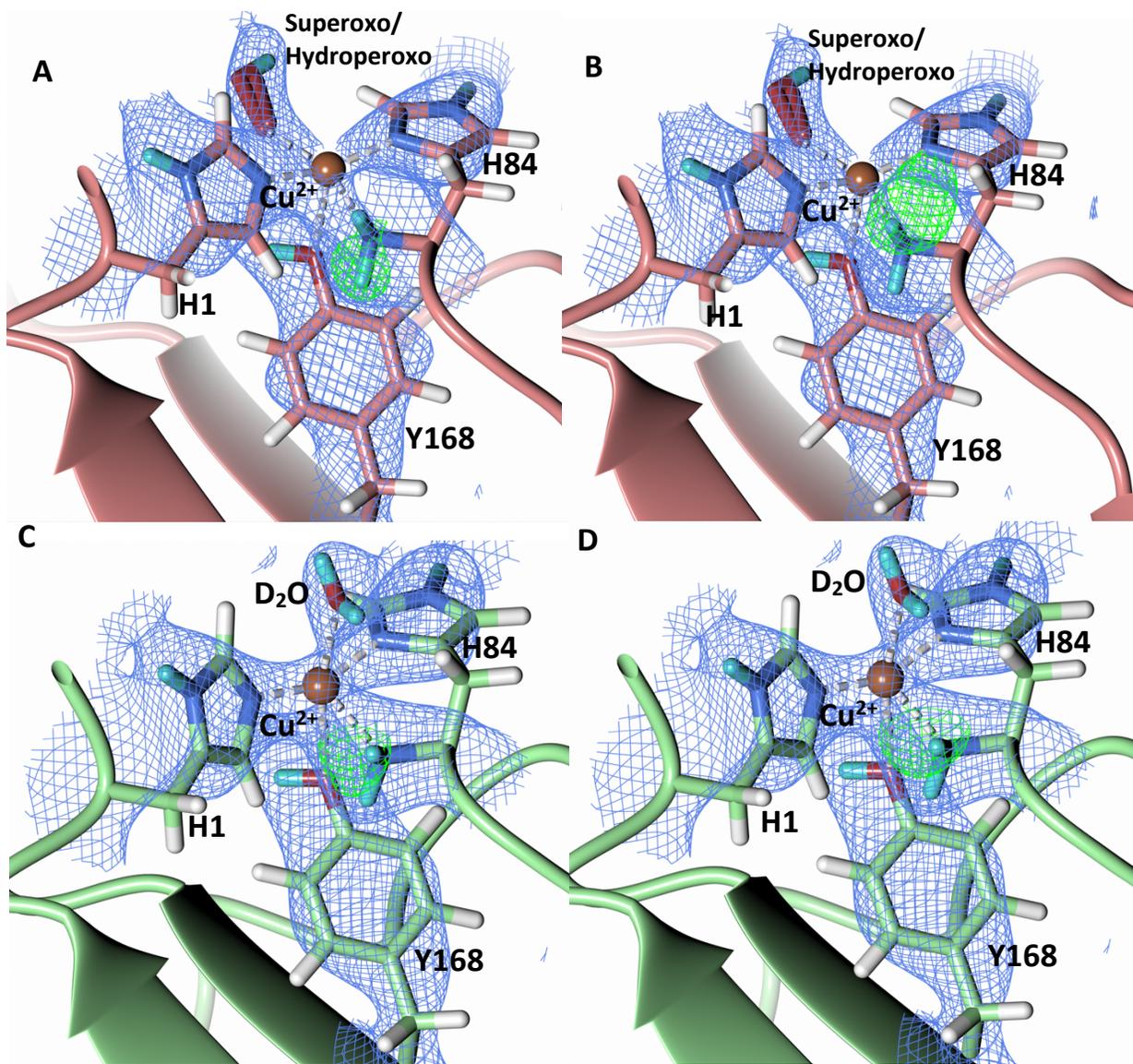


Figure 2.16 Protonation state of the N-terminal His157 amino group in the ascorbate reduced LPMO9D crystal. **(A)** Molecule A (coral) omit map displaying D2-protonation and **(B)**, D3-protonation of the N-terminal nitrogen. **(C)** Molecule B (light green) omit map displaying D2-protonation and **(D)** D3-protonation of the N-terminal nitrogen. Neutron scattering length density $2F_o-F_c$ maps displayed in blue at a 1.0σ cut-off. Neutron scattering length density F_o-F_c maps displayed in green at a 3.0σ cut-off. H atoms displayed in white and D atoms displayed in turquoise.

2.3.3.5 Histidine protonation at acidic conditions

During LPMO catalysis, a second shell histidine has been proposed to be involved in protonation of active site intermediates with QM/MM studies by Hedegård and Ryde on *LsAA9A* suggesting that His147 in a doubly protonated state can act as a proton donor.²² Additionally, Span *et al.* have proposed

that His 161 in *MtPMO3* plays a role as a proton donor based on structural, mutagenesis and spectroscopic studies as well as activity assays.²⁴ LPMOs are expected to catalyze oxygen insertion under acidic conditions since they function synergistically alongside cellulases which show optimum activity at pH 3.5 - 5.5.⁴⁴⁻⁴⁶ Given the surprising role of a neutral His157 as a proton donor to the activated dioxygen species in our crystal packing environment, we sought to probe the protonation state of His157 when exposed to acidic buffer conditions. We collected a 2.14 Å room-temperature neutron diffraction dataset under low pH conditions (pH 4.4/pD 4.8) achieved by vapor exchange of the crystal with acidified buffer. Analysis of the nuclear scattering length density maps indicate that the His157 residue remains singly protonated at the N_ε-position at this pH, which is supported by F_o-F_c omit difference maps in both Molecule A and B (Figure 2.17). To further support our acidic pH structure obtained by vapor exchange, a 1.5 Å X-ray dataset was collected on a crystal directly soaked in acidic buffer. The position of His157 at pH 4.4 superimposes to its position at pH 6.0. In this position, N_δ-protonation would sterically clash with the backbone amide proton, further supporting the single N_ε-protonation state observed in the neutron structure. The barricaded active site observed *in crystallo* and electropositive environment of the copper potentially play a role in maintaining the observed single histidine protonation state and prevent formation of protonation pathways to the active site. In both the neutron and X-ray low pH structures, the copper remains coordinated in the histidine brace, contrary to the disordered His78 observed in *LsLPMO9A* at pH 3.5 (PDB 5N04).⁴⁷ The low pH structure of *LsLPMO9A* was obtained using high NaCl concentrations, however, resulting in His 78 displacement by a chloride anion. The absence of NaCl in our crystallization conditions precludes coordinating histidine residue displacement by a chloride anion.

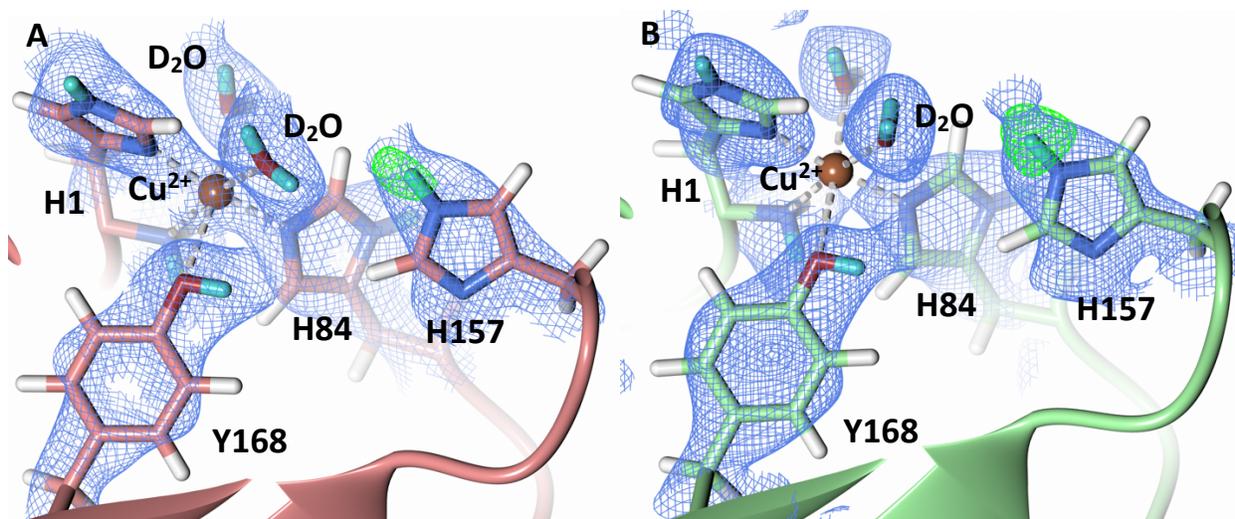


Figure 2.17 Protonation state of second-shell His157 at acidic conditions. **(A)** Molecule A (coral) His157 N ϵ -protonation. **(B)** Molecule B (light green) His157 N ϵ -protonation. Neutron scattering length density $2F_o - F_c$ maps displayed in blue at a 1.0 σ cut-off. Neutron scattering length density $F_o - F_c$ maps displayed in green at a 3.0 σ cut-off. H atoms displayed in white and D atoms displayed in turquoise.

2.3.3.6 Conformational flexibility of His157

Analysis of the crystal structure of *MtPMO3** deposited by Span *et al.* indicate that the second shell His161 corresponding to *NcLPMO9D* His157 is present in two conformations: an “inward” copper-facing conformation with a with a Cu-N ϵ distance of ~ 5 Å, and an “outward” copper-distant conformation with a Cu-N ϵ distance of ~ 10 Å (Figure 2.18).²⁴ The experimentally observed existence of both “inward” and “outward” conformations indicates mobility of the conserved second shell histidine, at least when the copper active site is solvent exposed. To further probe the extent of such local flexibility, if any, in our own structure, we carried out 2nd-generation mining minima (VM2) free energy calculations,⁴⁸⁻⁵¹ and compared the predicted thermodynamically accessible *NcLPMO9D* His157 conformations found when the active site is solvent exposed to those found when the active site is in the crystal structure with binding-face to binding-face packing.

When applied to *NcLPMO9D* with its active site exposed to solvent (i.e. the monomeric structure), the VM2 method, which performs extensive molecular conformational searching and provides a total free

energy via a Boltzmann weighted ensemble of the resultant low energy conformers, predicts that both “inward” and “outward” His157 conformations are populated. Focusing only on the two lowest energy conformers found by VM2 (Figure 2.19), the first conformer is “inward” and the second conformer is “outward” with a free energy difference of only 0.11 kcal mol⁻¹ between them. Furthermore, summing the probabilities of each type of conformation populated, so that 99.0% of the total population is included, His157 has a predicted 3:1 “inward” to “outward” probability distribution. In contrast, when applied to the NcLPMO9D dimer, as per the crystal structure (Figure 2.12 A), VM2 finds only conformations where His157 remains fully “inward”. This is despite VM2’s ability to drive molecular conformations over large energy barriers, as well as a free energy window set at 10 kcal mol⁻¹ for retaining minima found by the conformational search. This result strongly supports preclusion of movement of His157 to an “outward” conformation due to the barricaded active site present in the crystal structure.

We conclude, then, that the steric obstruction resulting from crystal packing of the dimer has significantly limited the conformational flexibility of His157 and suggest that the likely mobility of this residue in the monomer, which has a solvent exposed active site, may facilitate its protonation and serve a functional role. Taken together with the crystal structure active site occlusion, these factors may explain the single N_ε-protonation of His157 observed even at acidic conditions. However, the single protonation of His157 solely at the N_ε position in our crystal structures does not rule out formation of the doubly protonated second shell histidine residue under similar acidic *operando* conditions where the active site would not be barricaded by crystal packing.

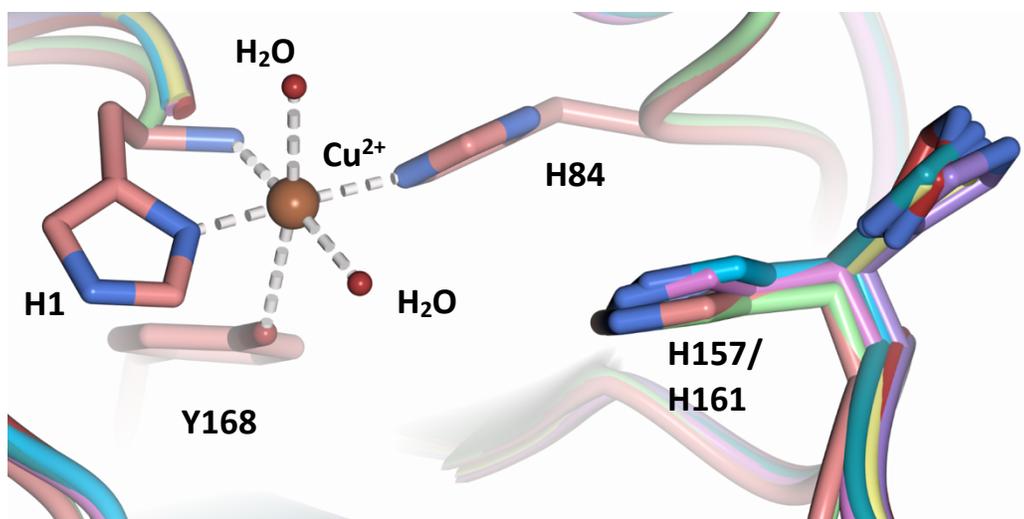


Figure 2.18 Alignment of a second shell residue His157 in the “inward” and “outward” conformation. Alignment of Molecule A *MtPMO3** His161 (PDB 5UFV) Molecule A (teal), Molecule B (yellow), Molecule C (pink), Molecule D (purple), Molecule E (turquoise) and Molecule F (red) with *NcLPMO9D* His157 from Molecule A (coral) and Molecule B (green).

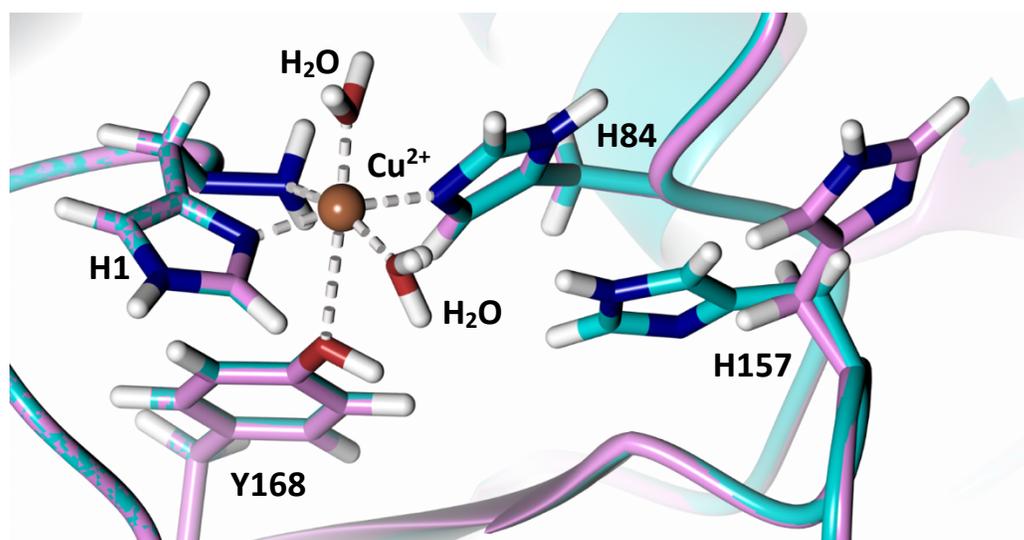


Figure 2.19 Second shell His157 conformations from VM2 free energy minima calculations. For *NcLPMO9D* with its active site exposed to solvent His157 adopts an “inward” (cyan) and “outward” (pink) conformation. The inward and outward conformations are separated by a $0.11 \text{ kcal mol}^{-1}$. H atoms displayed in white.

2.3.4 Conclusion

The neutron structure presented here provides a full-atom snapshot of the early catalytic intermediates in the LPMO reaction pathway. Our results provide the first direct observation of a cryo-trapped protonated copper–dioxygen species at the LPMO active site. The unique occluded active site environment found in the crystal structure presented here has resulted in stabilization of an activated dioxygen species at the *NcLPMO9D* active site with joint occupancy as a superoxo intermediate and a hydroperoxo intermediate, potentially to ultimately form H₂O₂ in the absence of substrate. A second-shell histidine was further observed in a single N_ε-protonation state and in the imidazolate form, leading us to postulate that it plays a role in the early stages of dioxygen intermediate protonation in the unique barricaded environment provided by the binding-face to binding-face crystal packing. Although unusual, a neutral histidine functioning as a general acid has been mechanistically proposed for the His95 residue in triose phosphate isomerase and the His98 residue in methylglyoxal synthase.^{52–54} To quote Martin Karplus: “Although it is generally thought that the imidazole of histidine must be doubly protonated to act as an acid, the simulation suggests that neutral imidazole can transfer a proton to O₂ of the enediolate.”⁵⁵ It remains to be determined whether the role of the second shell His157 in protonation of the activated dioxygen species prevails *operando* when *NcLPMO9D* is solvent-exposed and intermittently binds substrate. The flexibility of His157 and potentially that of neighboring active site residues may serve a role in catalysis that must be further explored.

2.3.5 Experimental section

A detailed account of the experimental procedure, data collection and refinement strategy has been described.⁵⁶ Provided here are the experimental details in brief.

2.3.5.1 Protein expression and crystallization

Neurospora crassa LPMO9D (*Nc*PMO-2) was heterologously expressed in the *Pichia pastoris* SuperMan₅ strain and purified and crystallized as described by O'Dell *et al.* previously^{29,30} Crystal quality was assessed on beamline CG4-D, the IMAGINE instrument at the High Flux Isotope Reactor at Oak Ridge National Laboratory.⁵⁷

2.3.5.2 Crystal ascorbate soak and data collection

To reduce the copper active site, a large crystal was harvested and placed in a deuterated buffer containing 100 mM HEPES pD 6.0 (pH 5.6), 25 % PEG 3350 and 100mM ascorbic acid pD 6.0 (pH 5.6). The crystal was incubated for two hours after which it was flash-frozen in liquid nitrogen using the deuterated buffer supplemented with 25 % glycerol as cryoprotectant. Time-of-flight neutron diffraction data were collected on beamline 11B, the Macromolecular Neutron Diffractometer (MaNDi) at the Spallation Neutron Source (SNS) at Oak Ridge National Laboratory.⁵⁸

2.3.5.3 Crystal low pH vapor exchange data collection

To obtain low pH conditions in the LPMO crystal, a large crystal was harvested and mounted in a quartz capillary to which a deuterated low pH buffer was added to facilitate vapor exchange in three stages.^{59,60} The first vapor exchange was performed in 22% PEG 3350 and 100 mM sodium acetate pD 5.6 (pH 5.2) for three days, followed by 22% PEG 3350 and 100 mM sodium acetate pD 5.2 (pH 4.8) for six days and 22% PEG 3350 and 100 mM sodium acetate pD 4.8 (pH 4.4) for 16 days. Time-of-flight neutron diffraction data were collected on beamline 11B, the Macromolecular Neutron Diffractometer (MaNDi) at the Spallation Neutron Source (SNS) at Oak Ridge National Laboratory.⁵⁸ Following this, an X-ray dataset was collected on the same crystal at 298 K using a copper rotating-anode home source.

2.3.5.4 Crystal low pH direct soak and data collection

LPMO crystals were directly exposed to acidic conditions by soaking crystals in low pH buffer. A crystal was transferred to a buffer composed of 22% PEG 3350, 100 mM and 100 mM sodium acetate with sequential incubation at pH 5.6, pH 5.2 and finally pH 4.8 for ten minutes each. Crystals were subsequently flash-frozen in liquid nitrogen with the pH 4.8 acidic buffer supplemented with 25 % glycerol used as a cryoprotectant. Data were collected at 100 K using a copper rotating-anode home source.

2.3.5.5 Model refinement

X-ray data were indexed, integrated and scaled with *CrysAlisPRO* and the *CCP4* suite.^{61,62} Neutron diffraction data were reduced using the *Mantid* data analysis and visualization package.⁶³ The time-of-flight diffraction data were integrated using three-dimensional profile fitting of the Bragg peaks.⁶⁴ Structure solution and refinement of the X-ray diffraction data of the low pH structures as well the neutron diffraction data of the ascorbate-soaked structure were performed by utilizing the *PHENIX* software suite with manual model building to fit density maps performed in *Coot*.⁶⁵⁻⁶⁷ Dioxygen species were modelled using restrained bond length refinements.

2.3.5.6 Electronic structure (density functional theory) calculations

The active site models used for DFT calculation were derived by extracting coordinates from the resting state X-ray structure (PDB 5TKG, O'Dell *et al.*²⁹) including residues His1, His84, His157, Gln166 and Tyr168 along with the active site copper(II) ion, pre-bound molecular O₂ and the axial and equatorial water molecules. To reduce the total number of atoms in these models, the main chain of residues His84, His157, Gln166 and Tyr168 were truncated at C_β which was modeled as a methyl group. DFT calculations were performed with Gaussian 09 using the B3LYP functional with the 6-31G** basis set applied to all atoms. The model was implicitly solvated using the polarizable continuum model as implemented in Gaussian 09 with a dielectric constant of 4.24 (scrf keyword solvent=DiethylEther). For geometry optimization, the coordinates of one heavy atom per residue (as performed previously in DFT calculation

of LPMO9D in O'Dell *et al.*²⁹) were constrained to their starting values to maintain the relative conformation imposed by the protein backbone.

2.3.5.7 VM2 calculations

VM2 is an implementation of the second-generation Mining Minima (M2) method originally developed by Gilson *et al.*⁶⁸⁻⁷³ It is a free energy calculation method, which combines extensive molecular conformational searching^{48,74} with a rigorous statistical mechanics approach.⁶⁸ Molecular configuration integrals over all space are closely approximated in a tractable manner by summation of local configuration integrals associated with the low energy minima of the system.^{75,76} VM2's search algorithm can drive molecular conformations over high-energy barriers, via a mode-distort-minimize procedure, to find relevant minima even when they are structurally diverse. As such, VM2 is well suited for exploration of thermally accessible conformations of His157 in monomeric and dimeric NcLPMO9D, respectively, as carried out in this study.

We applied molecular mechanics (MM) based VM2, taking the AMBER ff14SB parameter files (prmtop etc.) generated by the setup for MD simulations (see above) and converting them to formats readable by the VM2 software. It should be noted that for computational efficiency, VM2 includes solvation effects through continuum models (see below for details), so only the water molecules coordinating directly with active site copper centers were explicitly included. To further manage computational expense, and to focus on the region of interest, not all protein atoms were included in the VM2 calculation. Rather, a "cutout" approach was taken, where all residues containing an atom within 14 Å of the His157 alpha carbon (CA) were present in the energy model, and of these all atoms within 12 Å of His157 CA were set as mobile. For the monomer, this resulted in 904 atoms present in the calculation, with 476 of these mobile; for the dimer, this resulted in 1256 atoms present, with 600 of these mobile. To maintain the basic arrangement of the active site copper first-shell coordination residues (histidine brace, tyrosine, and two water molecules)

during aggressive conformational searching, a flat-bottomed energy well restraint potential¹¹ was applied to their coordinating atoms, as well as the central copper atom itself.

VM2 is an iterative method,⁴⁸ with each iteration comprising a conformational search, with removal of any repeat conformers produced, calculation of the configuration integrals at 300K for the resulting set of new conformers,^{75,76} addition of this set of conformers to the pool already established, and then determination of the total free energy by summation of the local configuration integrals of all conformers (minima) currently in the pool. Once the conformational search no longer finds new low energy conformers the total free energy stops changing between iterations. The calculations in this study were deemed converged when the absolute value of the free energy change between iterations was less than 0.01 kcal mol⁻¹. The conformational search carried out during each VM2 iteration included 1200 mode-distort-minimize procedures. Molecular distortions were from atom displacements along single-modes produced by diagonalization of the Hessian matrix in torsional space and, in addition, atom displacements based on random combinations of pairs of these modes.^{48,74} For energy-derivative involved steps, e.g. the mode-distort-minimize procedure and the Hessian (energy 2nd-derivative) calculation, required for configurational entropy terms, solvation energy was included using the generalized Born (GB) continuum model.^{77,78} The Poisson–Boltzmann Surface Area (PBSA) method⁷⁹ was applied to provide a more accurate single-point solvation energy correction at GB determined geometries. On completion, VM2 calculations output the total free energy, and in addition the relative free energy of each individual conformer along with its Boltzmann distribution probability. Also outputted are PDB and mol2 formatted files containing all protein conformers produced within 10 kcal mol⁻¹ of the lowest energy conformer, up to a maximum of 1,000, allowing for structural visualization and analysis.

2.3.6 Conflicts of interest

SPW declares an equity interest in VeraChem LLC, which develops the VM2 free energy software package. The remaining authors declare no conflict of interest.

2.3.7 Acknowledgements

Protein expression, purification and crystallization experiments were conducted at the Center for Structural Molecular Biology (CSMB), a U.S. Department of Energy Biological and Environmental Research User Facility at Oak Ridge National Laboratory. Neutron diffraction data was collected at BL-11B MaNDi at the Spallation Neutron Source (SNS) at ORNL which is sponsored by the Scientific User Facilities Division, Office of Basic Energy Sciences, U.S. Department of Energy. The authors thank Brendan Sullivan for assistance with data reduction. X-ray diffraction data was collected at the Molecular Education, Technology, and Research Innovation Center (METRIC) facilities at North Carolina State University, which is supported by the State of North Carolina. GCS acknowledges support in part from the National Research Foundation (NRF), South Africa and the Oak Ridge National Laboratory Graduate Opportunities (GO!) program. FM acknowledges support from USDA NIFA Hatch 211001.

2.3.8 References

1. Meier, K. K. *et al.* Oxygen Activation by Cu LPMOs in Recalcitrant Carbohydrate Polysaccharide Conversion to Monomer Sugars. *Chem. Rev.* **118**, 2593–2635 (2018).
2. Quinlan, R. J. *et al.* 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 (2011).
3. Phillips, C. M., Beeson, W. T., Cate, J. H. & Marletta, M. A. Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem. Biol.* **6**, 1399–1406 (2011).
4. Beeson, W. T., Phillips, C. M., Cate, J. H. D. & Marletta, M. A. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J. Am. Chem. Soc.* **134**, 890–892 (2012).
5. Beeson, W. T., Vu, V. V., Span, E. A., Phillips, C. M. & Marletta, M. A. Cellulose Degradation by Polysaccharide Monooxygenases. *Annu. Rev. Biochem.* **84**, 923–946 (2015).
6. Bissaro, B. *et al.* Oxidative cleavage of polysaccharides by monocopper enzymes depends on H₂O₂. *Nat. Chem. Biol.* **13**, 1123–1128 (2017).
7. Li, X., Beeson, W. T., Phillips, C. M., Marletta, M. A. & Cate, J. H. D. Structural Basis for Substrate Targeting and Catalysis by Fungal Polysaccharide Monooxygenases. *Structure* **20**, 1051–1061 (2012).
8. Kjaergaard, C. H. *et al.* Spectroscopic and computational insight into the activation of O₂ by the mononuclear Cu center in polysaccharide monooxygenases. *Proc. Natl. Acad. Sci.* **111**, 8797–8802 (2014).
9. Hangasky, J. A., Detomasi, T. C. & Marletta, M. A. Glycosidic Bond Hydroxylation by Polysaccharide Monooxygenases. *Trends Chem.* **1**, 198–209 (2019).
10. Chylenski, P. *et al.* Lytic Polysaccharide Monooxygenases in Enzymatic Processing of Lignocellulosic Biomass. *ACS Catal.* **9**, 4970–4991 (2019).
11. Peterson, R. L. *et al.* Cupric superoxo-mediated intermolecular C-H activation chemistry. *J. Am. Chem. Soc.* **133**, 1702–1705 (2011).
12. Mann, S. I., Heinisch, T., Ward, T. R. & Borovik, A. S. Peroxide Activation Regulated by Hydrogen Bonds within Artificial Cu Proteins. *J. Am. Chem. Soc.* **139**, 17289–17292 (2017).
13. Trammell, R. *et al.* Decoding the Mechanism of Intramolecular Cu-Directed Hydroxylation of sp³ C-H Bonds. *J. Org. Chem.* **82**, 7887–7904 (2017).
14. Garcia-Bosch, I. & Sieglar, M. A. Copper-Catalyzed Oxidation of Alkanes with H₂O₂ under a Fenton-like Regime. *Angew. Chemie Int. Ed.* **55**, 12873–12876 (2016).
15. Donoghue, P. J. *et al.* Rapid C–H Bond Activation by a Monocopper(III)–Hydroxide Complex. *J. Am. Chem. Soc.* **133**, 17602–17605 (2011).
16. Dhar, D. *et al.* Perturbing the Copper(III)-Hydroxide Unit through Ligand Structural Variation. *J. Am. Chem. Soc.* **138**, 356–368 (2016).
17. Schröder, D., Holthausen, M. C. & Schwarz, H. Radical-like activation of alkanes by the ligated copper oxide cation (phenanthroline)CuO⁺. *J. Phys. Chem. B* **108**, 14407–14416 (2004).
18. Yoshizawa, K., Kihara, N., Kamachi, T. & Shiota, Y. Catalytic mechanism of dopamine β-monooxygenase mediated by Cu(III)-oxo. *Inorg. Chem.* **45**, 3034–3041 (2006).
19. Dietl, N., Van Der Linde, C., Schlangen, M., Beyer, M. K. & Schwarz, H. Diatomic [CuO]⁺ and its role in the spin-selective hydrogen- and oxygen-atom transfers in the thermal activation of methane. *Angew. Chemie - Int. Ed.* **50**, 4966–4969 (2011).
20. Kim, S., Ståhlberg, J., Sandgren, M., Paton, R. S. & Beckham, G. T. Quantum mechanical calculations suggest that lytic polysaccharide monooxygenases use a copper-oxyl, oxygen-rebound mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 149–154 (2014).
21. Hedegård, E. D. & Ryde, U. Targeting the reactive intermediate in polysaccharide monooxygenases. *J. Biol. Inorg. Chem.* **22**, 1029–1037 (2017).
22. Hedegård, E. D. & Ryde, U. Molecular mechanism of lytic polysaccharide monooxygenases. *Chem. Sci.* **9**, 3866–3880 (2018).

23. Bertini, L. *et al.* Catalytic Mechanism of Fungal Lytic Polysaccharide Monooxygenases Investigated by First-Principles Calculations. *Inorg. Chem.* **57**, 86–97 (2018).
24. Span, E. A., Suess, D. L. M., Deller, M. C., Britt, R. D. & Marletta, M. A. The Role of the Secondary Coordination Sphere in a Fungal Polysaccharide Monooxygenase. *ACS Chem. Biol.* **12**, 1095–1103 (2017).
25. Caldararu, O., Oksanen, E., Ryde, U. & Hedegård, E. D. Mechanism of hydrogen peroxide formation by lytic polysaccharide monooxygenase. *Chem. Sci.* **10**, 576–586 (2019).
26. Wang, B., Walton, P. H. & Rovira, C. Molecular Mechanisms of Oxygen Activation and Hydrogen Peroxide Formation in Lytic Polysaccharide Monooxygenases. *ACS Catal.* **9**, 4958–4969 (2019).
27. Schröder, G. C., O'Dell, W. B., Myles, D. A. A., Kovalevsky, A. & Meilleur, F. IMAGINE: Neutrons reveal enzyme chemistry. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 778–786 (2018).
28. O'Dell, W. B., Bodenheimer, A. M. & Meilleur, F. Neutron protein crystallography: A complementary tool for locating hydrogens in proteins. *Arch. Biochem. Biophys.* **602**, 48–60 (2016).
29. O'Dell, W. B., Agarwal, P. K. & Meilleur, F. Oxygen Activation at the Active Site of a Fungal Lytic Polysaccharide Monooxygenase. *Angew. Chemie - Int. Ed.* **56**, 767–770 (2017).
30. O'Dell, W. B., Swartz, P. D., Weiss, K. L. & Meilleur, F. Crystallization of a fungal lytic polysaccharide monooxygenase expressed from glycoengineered *Pichia pastoris* for X-ray and neutron diffraction. *Acta Crystallogr. Sect. Struct. Biol. Commun.* **73**, 70–78 (2017).
31. Karkehabadi, S. *et al.* The First Structure of a Glycoside Hydrolase Family 61 Member, Cel61B from *Hypocrea jecorina*, at 1.6 Å Resolution. *J. Mol. Biol.* **383**, 144–154 (2008).
32. Harris, P. V. *et al.* Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: Structure and function of a large, enigmatic family. *Biochemistry* **49**, 3305–3316 (2010).
33. Eibinger, M., Sattelkow, J., Ganner, T., Plank, H. & Nidetzky, B. Single-molecule study of oxidative enzymatic deconstruction of cellulose. *Nat. Commun.* **8**, 894 (2017).
34. Frandsen, K. E. H. *et al.* The molecular basis of polysaccharide cleavage by lytic polysaccharide monooxygenases. *Nat. Chem. Biol.* **12**, 298–303 (2016).
35. Bacik, J. P. *et al.* Neutron and Atomic Resolution X-ray Structures of a Lytic Polysaccharide Monooxygenase Reveal Copper-Mediated Dioxxygen Binding and Evidence for N-Terminal Deprotonation. *Biochemistry* **56**, 2529–2532 (2017).
36. Isaksen, T. *et al.* A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *J. Biol. Chem.* **289**, 2632–2642 (2014).
37. Gudmundsson, M. *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).
38. Neisen, B. D., Gagnon, N. L., Dhar, D., Spaeth, A. D. & Tolman, W. B. Formally Copper(III)-Alkylperoxo Complexes as Models of Possible Intermediates in Monooxygenase Enzymes. *J. Am. Chem. Soc.* **139**, 10220–10223 (2017).
39. Kim, B., Jeong, D., Ohta, T. & Cho, J. Nucleophilic reactivity of a copper(II)-hydroperoxo complex. *Commun. Chem.* **2**, 1–6 (2019).
40. Westereng, B. *et al.* Enzymatic cellulose oxidation is linked to lignin by long-range electron transfer. *Sci. Rep.* **5**, 1–9 (2015).
41. Courtade, G. *et al.* Interactions of a fungal lytic polysaccharide monooxygenase with β -glucan substrates and cellobiose dehydrogenase. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 5922–5927 (2016).
42. Ciano, L., Davies, G. J., Tolman, W. B. & Walton, P. H. Bracing copper for the catalytic oxidation of C–H bonds. *Nat. Catal.* **1**, 571–577 (2018).
43. Dhar, D. & Tolman, W. B. Hydrogen atom abstraction from hydrocarbons by a copper(III)-hydroxide complex. *J. Am. Chem. Soc.* **137**, 1322–1329 (2015).
44. Schüle, M. Enzymatic properties of cellulases from *Humicola insolens*. *J. Biotechnol.* **57**, 71–81 (1997).
45. Boer, H. & Koivula, A. The relationship between thermal stability and pH optimum studied with

- wild-type and mutant *Trichoderma reesei* cellobiohydrolase Cel7A. *Eur. J. Biochem.* **270**, 841–848 (2003).
46. Cragg, S. M. *et al.* Lignocellulose degradation mechanisms across the Tree of Life. *Curr. Opin. Chem. Biol.* **29**, 108–119 (2015).
 47. Frandsen, K. E. H., Poulsen, J. C. N., Tandrup, T. & Lo Leggio, L. Unliganded and substrate bound structures of the cellobiosaccharide active lytic polysaccharide monooxygenase LsAA9A at low pH. *Carbohydr. Res.* **448**, 187–190 (2017).
 48. Chen, W., Gilson, M. K., Webb, S. P. & Potter, M. J. Modeling protein-ligand binding by mining minima. *J. Chem. Theory Comput.* **6**, 3540–3557 (2010).
 49. Huang, Y. M. M., Chen, W., Potter, M. J. & Chang, C. E. A. Insights from free-energy calculations: Protein conformational equilibrium, driving forces, and ligand-binding modes. *Biophys. J.* **103**, 342–351 (2012).
 50. You, W., Huang, Y. ming M., Kizhake, S., Natarajan, A. & Chang, C. en A. Characterization of Promiscuous Binding of Phosphor Ligands to Breast-Cancer-Gene 1 (BRCA1) C-Terminal (BRCT): Molecular Dynamics, Free Energy, Entropy and Inhibitor Design. *PLoS Comput. Biol.* **12**, 1–25 (2016).
 51. Xu, P. *et al.* Computation of host-guest binding free energies with a new quantum mechanics based mining minima algorithm. *J. Chem. Phys.* **154**, (2021).
 52. Lodi, P. J. & Knowles, J. R. Neutral imidazole is the electrophile in the reaction catalyzed by triosephosphate isomerase: structural origins and catalytic implications. *Biochemistry* **30**, 6948–6956 (1991).
 53. Lennartz, C., Schäfer, A., Terstegen, F. & Thiel, W. Enzymatic reactions of triosephosphate isomerase: A theoretical calibration study. *J. Phys. Chem. B* **106**, 1758–1767 (2002).
 54. Marks, G. T., Susler, M. & Harrison, D. H. T. Mutagenic Studies on Histidine 98 of Methylglyoxal Synthase: Effects on Mechanism and Conformational Change. *Biochemistry* **43**, 3802–3813 (2004).
 55. Bash, P. A. *et al.* Computer simulation and analysis of the reaction pathway of triosephosphate isomerase. *Biochemistry* **30**, 5826–5832 (1991).
 56. Schröder, G. C., O’Dell, W. B., Swartz, P. D. & Meilleur, F. Preliminary results of neutron and X-ray diffraction data collection on a lytic polysaccharide monooxygenase under reduced and acidic conditions. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **77**, 128–133 (2021).
 57. Meilleur, F., Kovalevsky, A. & Myles, D. A. A. IMAGINE: The neutron protein crystallography beamline at the high flux isotope reactor. in *Methods in enzymology* **634**, 69–85 (Elsevier Inc., 2020).
 58. Coates, L., Stoica, A. D., Hoffmann, C., Richards, J. & Cooper, R. The macromolecular neutron diffractometer (MaNDi) at the Spallation Neutron Source, Oak Ridge: enhanced optics design, high-resolution neutron detectors and simulated diffraction. *J. Appl. Crystallogr.* **43**, 570–577 (2010).
 59. Wan, Q. *et al.* Direct determination of protonation states and visualization of hydrogen bonding in a glycoside hydrolase with neutron crystallography. *Proc. Natl. Acad. Sci.* **112**, 12384–12389 (2015).
 60. Gerlits, O. *et al.* Long-Range Electrostatics-Induced Two-Proton Transfer Captured by Neutron Crystallography in an Enzyme Catalytic Site. *Angew. Chemie - Int. Ed.* **55**, 4924–4927 (2016).
 61. Kubicek, C. P. *Fungi and Lignocellulosic Biomass.* (Wiley-Blackwell, 2012). doi:10.1002/9781118414514
 62. Winn, M. D. *et al.* Overview of the CCP 4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 235–242 (2011).
 63. Arnold, O. *et al.* Mantid - Data analysis and visualization package for neutron scattering and μ SR experiments. *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* **764**, 156–166 (2014).
 64. Sullivan, B. *et al.* Improving the accuracy and resolution of neutron crystallographic data by three-dimensional profile fitting of Bragg peaks in reciprocal space. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 1085–1095 (2018).

65. Adams, P. D. *et al.* PHENIX : a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010).
66. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
67. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 352–367 (2012).
68. Gilson, M. K., Given, J. A., Bush, B. L. & McCammon, J. A. The statistical-thermodynamic basis for computation of binding affinities: A critical review. *Biophys. J.* **72**, 1047–1069 (1997).
69. Head, M. S., Given, J. A. & Gilson, M. K. ‘Mining minima’: Direct computation of conformational free energy. *J. Phys. Chem. A* **101**, 1609–1618 (1997).
70. David, L., Luo, R. & Gilson, M. K. Ligand-receptor docking with the Mining Minima optimizer. *J. Comput. Aided. Mol. Des.* **15**, 157–71 (2001).
71. Kairys, V. & Gilson, M. K. Enhanced docking with the mining minima optimizer: Acceleration and side-chain flexibility. *J. Comput. Chem.* **23**, 1656–1670 (2002).
72. Chang, C. E. & Gilson, M. K. Free energy, entropy, and induced fit in host-guest recognition: Calculations with the second-generation mining minima algorithm. *J. Am. Chem. Soc.* **126**, 13156–13164 (2004).
73. Chen, W., Chang, C. E. & Gilson, M. K. Calculation of cyclodextrin binding affinities: Energy, entropy, and implications for drug design. *Biophys. J.* **87**, 3035–3049 (2004).
74. Chang, C. E. & Gilson, M. K. Tork: Conformational Analysis Method for Molecules and Complexes. *J. Comput. Chem.* **24**, 1987–1998 (2003).
75. Chang, C. E., Potter, M. J. & Gilson, M. K. Calculation of molecular configuration integrals. *J. Phys. Chem. B* **107**, 1048–1055 (2003).
76. Potter, M. J. & Gilson, M. K. Coordinate systems and the calculation of molecular properties. *J. Phys. Chem. A* **106**, 563–566 (2002).
77. Qiu, D., Shenkin, P. S., Hollinger, F. P. & Still, W. C. The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate Born radii. *J. Phys. Chem. A* **101**, 3005–3014 (1997).
78. Still, W. C., Tempczyk, A., Hawley, R. C. & Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **112**, 6127–6129 (1990).
79. Luo, R., David, L. & Gilson, M. K. Accelerated Poisson-Boltzmann calculations for static and dynamic systems. *J. Comput. Chem.* **23**, 1244–1253 (2002).

Chapter 3 : Optimization of expression and crystallization of CYP450cam

3.1 Introduction to the structure and mechanism of cytochrome P450s.

Cytochrome P450 (CYP450) enzymes comprise a superfamily of heme-containing proteins ubiquitous in most life kingdoms that catalyze the important process of oxygen insertion into a wide range of substrates.¹ The origin of the name CYP450 (Pigment-450) derives from the Soret peak observed in pigments of rat liver microsomes at 450nm in the presence of carbon monoxide for these enzymes in 1958 when they were still functionally uncharacterized.^{2,3} CYP450s most commonly catalyze the highly regio- and stereoselective oxyfunctionalization of a wide variety of substrates.⁴ While substrate hydroxylation is most commonly catalyzed by CYP450s, they also catalyze epoxidation, sulfoxidation and N- or O-dealkylations.⁵ Their biological functions include drug metabolism, biosynthesis of lipids, steroids and degradation of xenobiotics and currently approximately 300 000 sequences that code for CYP450s have been identified.⁶ While their ubiquity in nature is well characterized, CYP450s are absent in certain *Archaea* and *Eubacteria* as well as the model prokaryotic *Escherichia coli*.^{7,8} CYP450s are present in prokaryotes, viruses, fungi, plants and mammals with the number and diversity of functions closely related to environment and physiological function.⁹

3.1.1 Cytochrome P450s in the different kingdoms of life

3.1.1.1 Prokaryotic CYP450s

P450s are plentiful in prokaryotes spanning more than 600 families and display a broad range of activities.⁶ Their roles include fatty acid hydroxylation and metabolism of xenobiotics.¹⁰ In contrast to eukaryotic CYP450s that are membrane-bound, prokaryotic CYP450s are soluble, which facilitated their characterization.¹¹ Indeed, the first X-ray crystal structure of a CYP450 was of CYP450cam (CYP101A1) from the prokaryotic Gram-negative *Pseudomonas putida* by Poulos *et al.* in 1985.¹²

3.1.1.2 Viral CYP450s

Although exceptionally rare in viruses, the virus *Acanthamoeba polyphaga mimivirus* (APMV) was first found to encode for two CYP450 genes in 2009.¹³ Two genes were identified: MIMI_L532 (YP_142886) and MIMI_L808(YP_143162). MIMI_L532 was found to have homology to a number of bacterial and nematode CYP450s proteins, however expression was not possible by Lamb *et al.*¹³ MIMI_L808 possessed homology to CYP51, a sterol 14-demethylase, and was successfully expressed. The native function of MIMI_L808 remains unknown, however, with sterol substrates showing no binding. Several giant viruses have now been proposed to contain CYP450 genes as well as a herpesvirus and a mycobacteriophage.¹⁴ Lamb *et al.* were successfully able to crystallize a phage CYP450 classified as CYP102L1, representing an exciting new insight into viral CYP450s.¹⁴

3.1.1.3 Fungal CYP450s

The fungal kingdom contains approximately 800 CYP450 families and close to 8000 annotated members.⁶ The families CYP51 and CYP61 are found in most species of fungi since they play a role in the synthesis of ergosterol, a membrane component.¹⁵ Fungal CYP450s are associated with many roles in metabolism, with various species displaying a range of CYP450s dependent on their ecological niche.¹⁶

3.1.1.4 Plant CYP450s

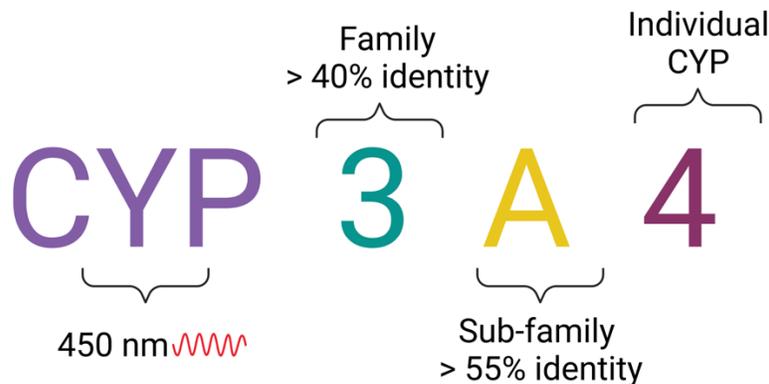
The sheer diversity of plant metabolites, particularly oxygenated low molecular weight natural products involved in signaling, defense and adaptation have necessitated large number of specialized CYP450s.¹⁷ There are about 277 plant CYP450s and more than 16 000 annotated members.⁶ Plant CYP450s can be categorized into A-type and non-A-type; the A-type CYP450s are now part of the CYP71 family and play a role in plant specific metabolism including lignin alkaloid and flavonoid biosynthesis while the non-A-type CYP450s are involved in a more general role of oxygenation of lipids and sterols.¹⁸

3.1.1.5 Mammalian CYP450s

Due to their central role in medical research, mammalian CYP450s have been extensively studied.¹⁹ CYP450s play an essential role in drug, steroid and fat-soluble vitamin metabolism as well as xenobiotic oxygenation, but they are also involved in the formation of reactive intermediates that affect signal transduction pathways, causing tumorigenesis and inflammation.²⁰ Mammalian CYP450s involved in detoxification are located in liver microsomes where they are membrane-bound through an N-terminal anchor.²¹ There are 18 CYP450 families in the human genome with 17 annotated members.²² CYP3A4 is the best characterized and most targeted of human CYP450s since it is involved in half of all drugs metabolized by CYP450s.²³

3.1.2 Nomenclature of CYP450s

The vast number of CYP450s has necessitated the development of a CYP450 nomenclature system.²⁴ Within this hierarchical system based on sequence identity, CYP450s are divided into families if they have more than 40% sequence identity which is further divided into subfamilies if they share more than 55% sequence identity (Figure 3.1).^{25,26} This is designated as the name “CYP” followed by a family number, then the subfamily is designated with a letter and finally the individual CYP is given by the last number. It should be noted that while this system groups CYP450s according to their sequence identity, often the CYP450s within these families have dissimilar substrates and physiological functions.²⁷ As of 2018, there were 791 crystal structures of CYP450s with close to 55 000 named CYP450 genes.²⁸



Created in BioRender.com bio

Figure 3.1 Nomenclature of cytochrome P450s (CYP450s). CYP450s are given the name “CYP” to indicate the superfamily of “pigments” that have an absorbance maximum at 450 nm when reduced and bound to CO. This is further divided into families and sub-families depending on sequence identity with individual CYP450s designated by the final number. Figure prepared with BioRender.

3.1.3 Classes of CYP450s

In order to perform catalysis CYP450s require an electron source or reductase partner protein.²⁹ CYP450s are classified into ten different classes depending on their reductase type and electron transport chain architecture and cellular localization.³⁰ While the classes are diverse and may expand when new CYP450s are discovered, Class I and Class II make up about 90% of all CYP450s.²⁶ Class I CYP450s are subdivided into Class Ia and Class Ib and encompass bacterial and mitochondrial CYP450s which are cytosolic and mitochondrial-membrane-bound, respectively (Figure 3.2 and Figure 3.3).³⁰ In Class I, the electron transport chain is composed of a three component system which is expressed from separated genes: the FAD-dependent ferredoxin reductase (FdxR), an iron-sulfur cluster ferredoxin (Fdx) and the heme-center CYP450. For both Class Ia and Ib, NAD(P)H functions as the electron donor with Fdx acting as a soluble, mobile electron shuttle. While all the components in Class Ia are soluble, both FdxR and CYP450 of Class Ib are embedded in the mitochondrial membrane (Figure 3.2). The prokaryotic Class Ia plays a role in the metabolism of fatty acids and xenobiotics,³¹ while the eukaryotic Class Ib is involved in synthesis of vitamin D and steroid hormones derived from cholesterol.³²

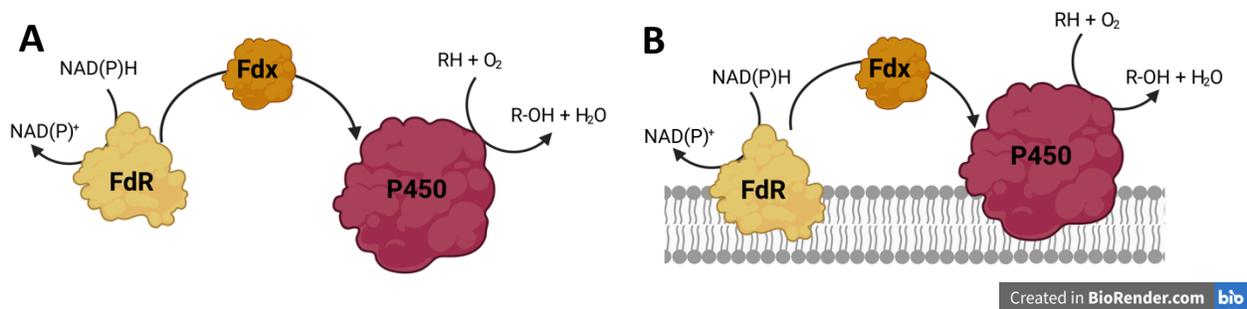


Figure 3.2 CYP450 Class I electron transport system. **(A)** The bacterial Class Ia system is a soluble system composed of ferredoxin reductase (FdR), an iron-sulfur cluster ferredoxin (Fdx) and the heme-center CYP450. **(B)** The mitochondrial Class Ib is membrane bound and composed of a ferredoxin reductase (FdR), a soluble iron-sulfur cluster ferredoxin (Fdx) to act as an intermediary and the heme-center CYP450. Figure prepared with BioRender.

Class II represents the most common class amongst eukaryotic CYP450s and have a range of functions including metabolizing endogenous compounds in mammals to playing a role in structural polymer synthesis in plants.³⁰ Class II is a two component integral membrane system composed of CYP450 reductase (CPR) domain which has a FAD and FMN cofactor and the CYP450 domain. Like Class I, the electron source for Class II is reduced NAD(P)H (Figure 3.3).

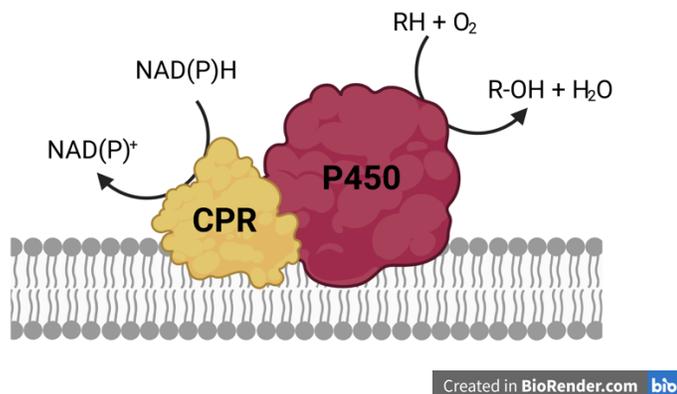


Figure 3.3 CYP450 Class II electron transport system. The eukaryotic integral membrane protein Class II system is composed of a CYP450 reductase (CPR) and a CYP450. Figure prepared with BioRender.

The remaining eight classes of CYP450s vary according to their domain architecture and electron sources and cofactors.³⁰ Of note are Class VII and VIII which are designated as fusion systems derived from a single gene. Class VII systems are composed of an N-terminal CYP450 domain fused to a C-terminal phthalate family oxygenase (PFOR)-like reductase domain that has a FMN and iron-sulfur cluster.³³ Class

VIII is composed of an N-terminal CYP450 domain fused to a C-terminal FMN and FAD cofactor dependent CYP450 reductase (CPR) domain. Self-sufficient CYP450s have garnered much scientific interest for biotechnological applications because of their broad substrate specificity, chemical versatility and self-sufficient nature that makes them catalytically efficient and amenable to enzyme engineering.³⁴

3.1.4 Structure of CYP450s

The overall structure of CYP450s is very similar, even with sequence identity across the CYP450 superfamily of less than 20%.³⁵ CYP450 forms an overall triangular prism shape composed primarily of α -helices and some β -sheets, with a catalytic heme center (Figure 3.4).³⁶ First structural insights were gained from the crystal structure of CYP101A1 (CYP450cam).¹² This structure contained 12 α -helices designated A – L and 5 β -sheets designated β 1 - β 5. This structural knowledge has been expanded to approximately 800 total structures with the scope extending from prokaryotic and eukaryotic to even include viral structures.^{14,28} As the number of CYP450 structures has increased, so have the labels to describe the secondary structure components which currently include α -helices A, A', B, B', B'', C, D, D', E, F, F', G, G', H', J, J', K, K', K'', L, L' and β -sheets β 1- β 6.³⁷

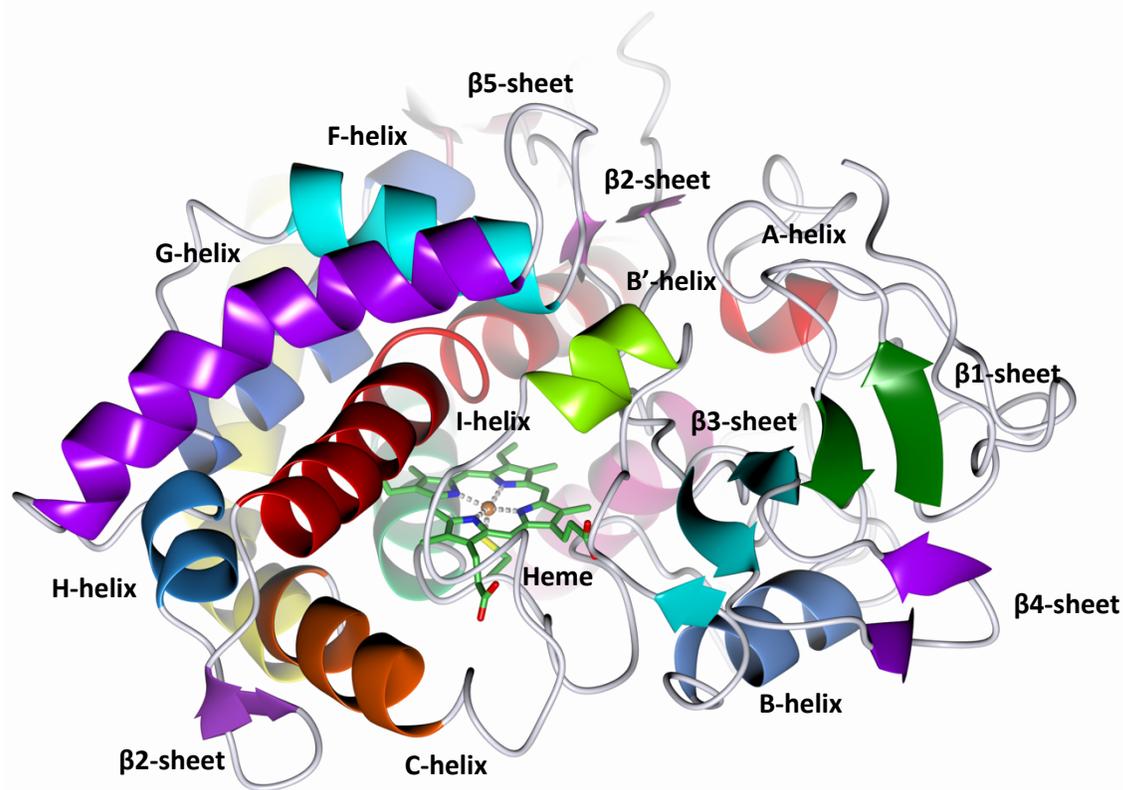


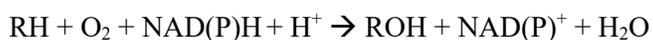
Figure 3.4 Representation of the structural elements of CYP450s (PDB 1DZ4). CYP450 forms a triangular prism shape composed of α -helices and β -sheets, with a catalytic heme center next to the I-helix.

The core of CYP450s is composed of a four-helix bundle in which three helices (D, L and I) are parallel and the E-helix is antiparallel³⁸ The active site of a CYP450 contains a heme group coordinated to the sulfur of a conserved cysteine residue within the CYP450 “signature sequence” FXXGXbXXCXG, where X represents any residue and Xb represents a basic residue⁴ The sulfur atom of the cysteine is liganded to the heme iron, forming the proximal ligand.³⁹ The region near the heme active site is the most conserved structurally with the active site heme neighbored by the I-helix and L-helix on the distal face, and the proximal cysteine coordinated to the heme iron. The I-helix is the longest helix on CYP450 and contains the conserved (A/G)GX(E/D)T sequence, resulting in a characteristic “kink” believed to form the oxygen binding site^{40,41} Schlichting et al. subsequently showed that a water molecule binds in this groove and stabilizes the bound oxygen.⁴² Regions proposed to be involved in substrate recognition and binding are called substrate recognition sites (SRS) and include the B'-helix (SRS1), F- and G-helices (SRS2 and

SRS3), the I-helix (SRS4), the β 4-sheet hairpin (SRS5) and the K-helix β 2-sheet loop (SRS6).⁴³ The regions involved in substrate specificity show the most structural divergence permitting the broad range of CYP450 substrates and varied regio- and stereo-selectivity.⁴⁴

3.1.5 The Cytochrome P450 Catalytic cycle

Central to the chemistry of CYP450s is the cysteine coordinated heme in the active site.⁴⁵ The overall reaction catalyzed by CYP450s requires molecular oxygen and two electrons derived from the reduced cofactors NAD(P)H and can be summarized as follows:⁴⁶



In its resting state, the heme iron is in a low-spin ferric Fe^{3+} state with a water molecule coordinated in the distal site.⁴⁷ Upon substrate binding, the resting state water molecule is displaced and the Fe^{3+} shifts from a low-spin to a high-spin state (Figure 3.5).⁴⁸ This change in spin-state results in a lowered redox potential and Fe^{3+} is subsequently reduced to the Fe^{2+} ferrous heme. It has, however, also been shown that one-electron reduction can occur before the substrate binds.⁴⁹ The molecular oxygen co-substrate subsequently binds and following a second reduction a ferric-peroxo intermediate is formed.⁵⁰ This ferric-peroxo intermediate is protonated to the ferric hydroperoxo species, termed Compound 0.^{51,52} Further protonation results in scission of the O-O bond and the release of water and a ferryl-oxo porphyrin radical. This ferryl-oxo porphyrin radical is highly reactive and has been designated “Compound I” since it reacts directly with the substrate.^{4,53} Compound I performs hydrogen atom abstraction from the substrate by homolytic bond cleavage forming a substrate radical and a ferryl hydroxide termed Compound II.^{54,55} Since the substrate radical is highly reactive, this recombines with the ferryl hydroxide to form the hydroxylated product.^{56,57} Following product release, water binds to the ferric heme, returning the enzyme to its resting state. In an alternative pathway, H_2O_2 can also serve as the CYP450 co-substrate, resulting in direct formation of Compound I via a “peroxide shunt”.⁵⁸

There are several routes for the “uncoupling” of the CYP450 reaction cycle. This can occur after the formation of the ferric-peroxo intermediate, resulting in formation of a superoxo species, or by protonation of the ferric hydroperoxo in absence of substrate resulting in formation of H₂O₂. The Compound I can also be reduced protonated and reduced to form water and the resting state ferric heme.

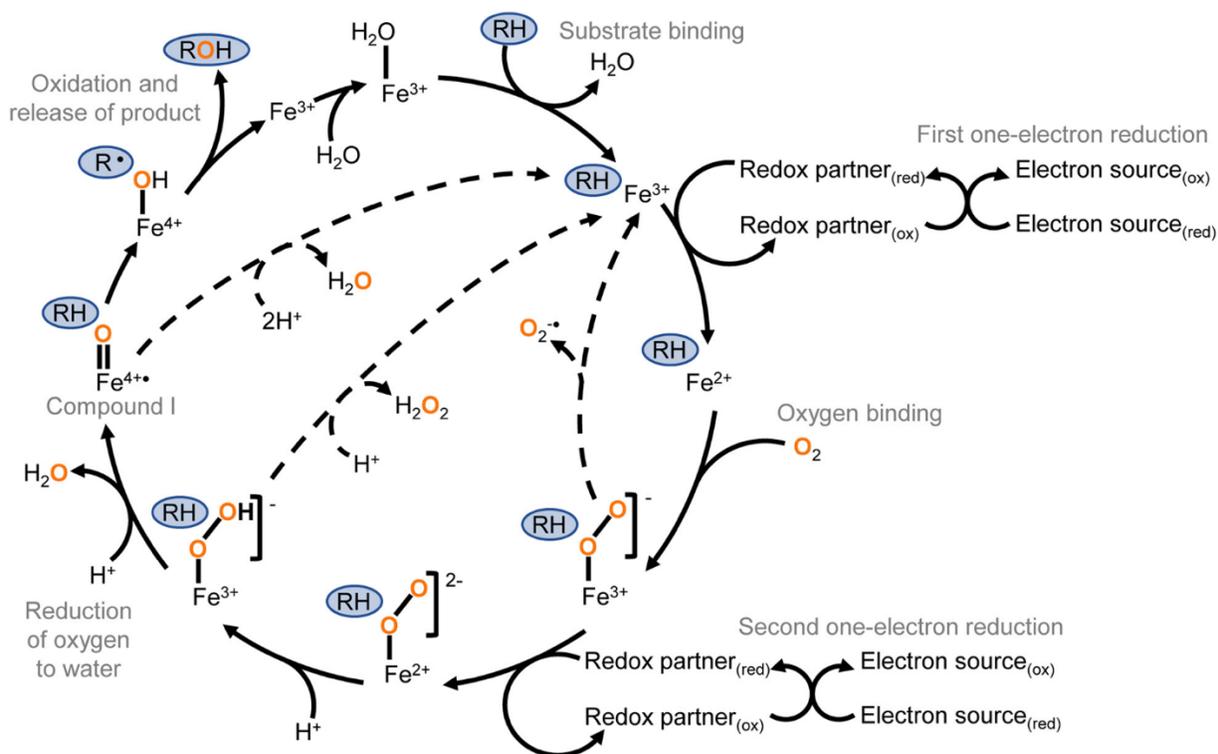


Figure 3.5 The overall CYP450 catalytic cycle. The CYP450 cycle is initiated by substrate binding followed by reduction and oxygen binding. Following a second reduction and two protonation steps, the ferryl-oxo porphyrin radical Compound I forms. Compound I abstracts a hydrogen from the substrate and forms the hydroxylated product following recombination with the substrate radical. Solid arrows represent the productive cycle and the dashed lines represent the uncoupling reaction. Adapted from Behrendorff *et al.*⁵⁹ and reprinted with permission of Frontiers in Microbiology.

In order to perform catalysis, CYP450s require the addition of two electrons. The transfer of electrons to the CYP450 active site is often the rate limiting step in catalysis.⁶⁰ The primary source of these electrons in two consecutive single reduction steps are reduced NAD(P)H which is oxidized by a diflavin reductase partner protein, the CYP450 reductase (CPR), which transfers the electrons to the CYP450 heme either directly or *via* additional diflavin, flavodoxin or ferredoxin redox partner proteins.^{30,61,62} In addition

to partner protein-dependent reduction, a number of chemical reductants can serve as electron sources such as sodium dithionite and titanium(II)citrate, as well as electrodes light-derived electron sources.⁶³⁻⁶⁵

3.1.6 Protonation pathway

Oxygen insertion is catalyzed by the high-valent iron-oxo species termed Compound I.⁴ This requires cleavage of the O-O bond in molecular oxygen by protonation of the leaving oxygen.⁴⁴ A complicated proton delivery pathway involving active site residues and water molecules is likely to be involved, but the exact nature of the pathway is still debated since there is no residue in the vicinity that can function as an acid-base catalyst such as the histidine residue as observed in peroxidases.⁶⁶ Currently, most of the insight into the protonation pathway has been obtained from studies on CYP450cam from *Pseudomonas putida* which catalyzes the hydroxylation of camphor in the 5-*exo* position.^{12,67,68} Neighboring the kink of the I-helix of CYP450, there is a highly conserved “acid-alcohol pair” composed of Asp/Glu and Ser/Thr which are designated Asp251 and Thr252 in CYP450cam (Figure 3.6).^{40,41,45} Early crystallographic and mutagenesis studies by Gerber *et al.* suggested that Thr252 protonates the dioxygen species with Asp251 acting as a proton donor and that there may be a bridging water molecule involved.^{69,70} Asp251 shuttles protons from the surface residues Lys178, Asp182 and Arg186 to the active site, thereby functioning as a “carboxylate switch”.

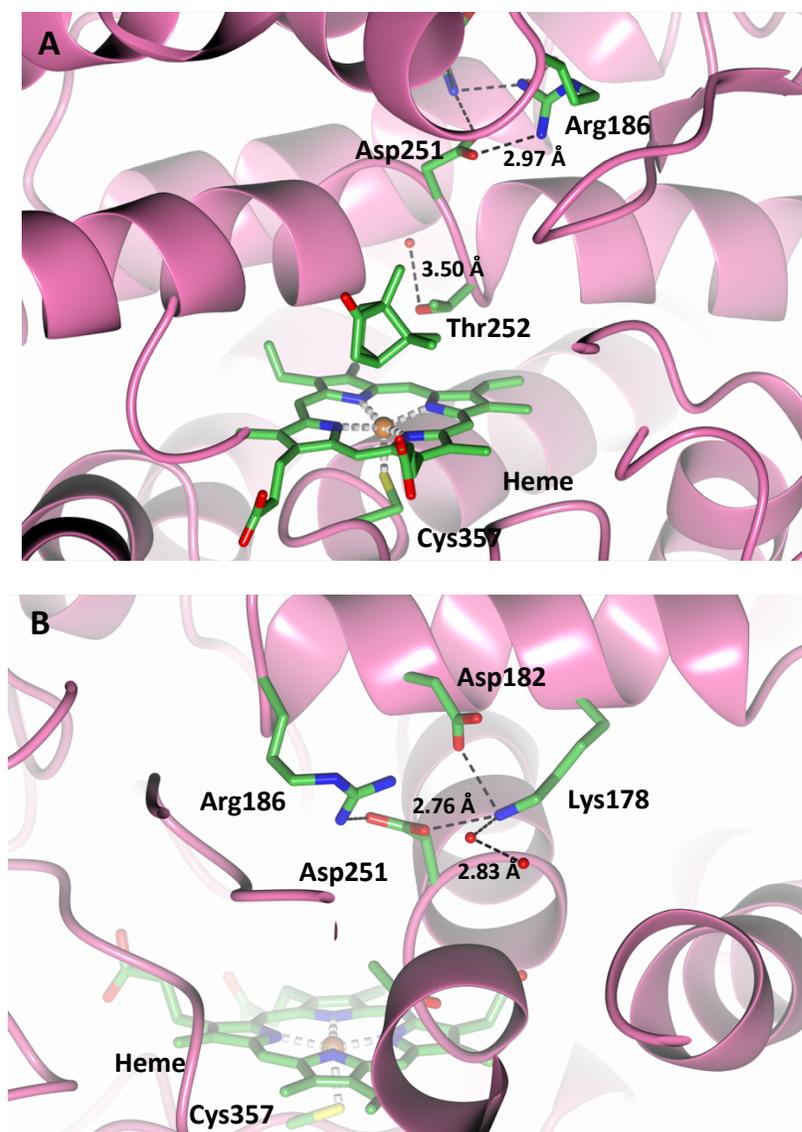


Figure 3.6 Residues involved in the CYP450 protonation pathway (PDB 1DZ4). **(A)** Thr252 is located in the I-helix kink and in the closed state there is only one water molecule in the I-helix and Asp251 is restrained in a salt bridge interaction. **(B)** Asp251 is held in a salt bridge interaction in the closed state with Lys178 and Arg186 (camphor omitted for clarity).

Mutagenesis of the Thr252 residue indicate that while a Ala252 mutant does not show reduced NADH consumption, there is significant uncoupling with H₂O₂ formation observed.^{71,72} Further mutagenesis studies by Vidakovic *et al.* supported that Asp251 functions as a “carboxylate switch” mediating protons from the surface of the protein.⁷³ Their activity studies indicated that mutation of Asp251 to asparagine significantly reduced the rate of camphor hydroxylation to 1% of that observed in the wild type. In the model proposed by Vidakovic *et al.*, an additional water molecule is present in the active site

with Asp251 shuttling protons to the dioxygen species *via* a two water molecule network stabilized by Thr252. Driving the reaction with in-beam chemistry Schlichting *et al.* found that Asp251 may act as a “carbonyl switch”, stabilizing active site waters instead of functioning as a proton donor, a role assigned to Thr252 and two water molecules which form a part of a proton delivery mechanism.⁴² These findings implied that following oxygen binding, there is a conformational change in the I-helix in which the carbonyl group of the Asp251 residue flips. This causes rearrangement of the active site in which Thr252 forms a hydrogen bond with the bound dioxygen and two water molecules enter the active site.^{41,42} Early studies by Kimata *et al.* put the role of Thr252 as a direct source of protons into question since a CYP450cam with chemical modification of the Thr252 hydroxyl group to a methoxy group still displayed activity.⁷⁴ An additional structural and mutagenesis study by Hishiki *et al.* on a Thr252Ile mutant proposed that a proton delivery network is formed by Lys178/Arg186, Asp251 and two waters molecules. In this scenario, a water molecule serves as the proton donor and Thr252 stabilizes this water *via* a hydrogen bond.⁷⁵ Structural studies of the ferrous dioxygen mutants Thr252Ala and Asp251Asn by Nagano *et al.* also suggested that Thr252 does not play a role as a hydrogen bond donor or acceptor during dioxygen activation.⁴¹ It was concluded that Thr252 may acts as a hydrogen bond acceptor from the hydroperoxy intermediate, thereby increasing affinity of the distal oxygen in the hydroperoxy species for second protonation. QM/MM studies by Zheng *et al.* and Wang *et al.*, however, suggested that protonation of the dioxygen species occurs through a concerted mechanism involving transfer from a protonated Asp251 *via* a water molecule and Thr252.^{76,77} An alternative to an acid-alcohol pair being involved in the first protonation step has been recently proposed by Kalita *et al.* used QM/MM studies to show that a heme propionate/Arg299 pair may be involved in intermediate protonation.⁷⁸ The discussed findings illustrate that the first protonation in the CYP450cam first protonation step still remains a debated topic. Thus far it can be concluded that the Asp/Glu and Ser/Thr acid-alcohol motif is important for intermediate protonation during catalysis, however, the exact proton delivery pathway may differ amongst different CYP450s.⁷⁹

3.1.7 CYP101A1 (CYP450cam)

CYP450cam (CYP101A1) from *Pseudomonas putida* was the first CYP450 structure to be determined and most mechanistic insights into CYP450s have been gained from studies of this “archetypal” system.^{67,68} CYP450cam is a 46.7 kDa protein that catalyzes the regio- and stereoselective hydroxylation of camphor in the 5-*exo* position. This structure contains 12 α -helices designated A – L and 5 β -sheets designated β 1 - β 5. The active site heme is positioned between helices I and L and the heme iron is coordinated to the cysteine residue Cys357 (Figure 3.7). The substrate access channel is composed of the antiparallel F and G-helices, the F-G loop and a part of the B'-helix. These structural elements are positioned around the heme to ensure correct camphor positioning for catalysis.^{80,81}

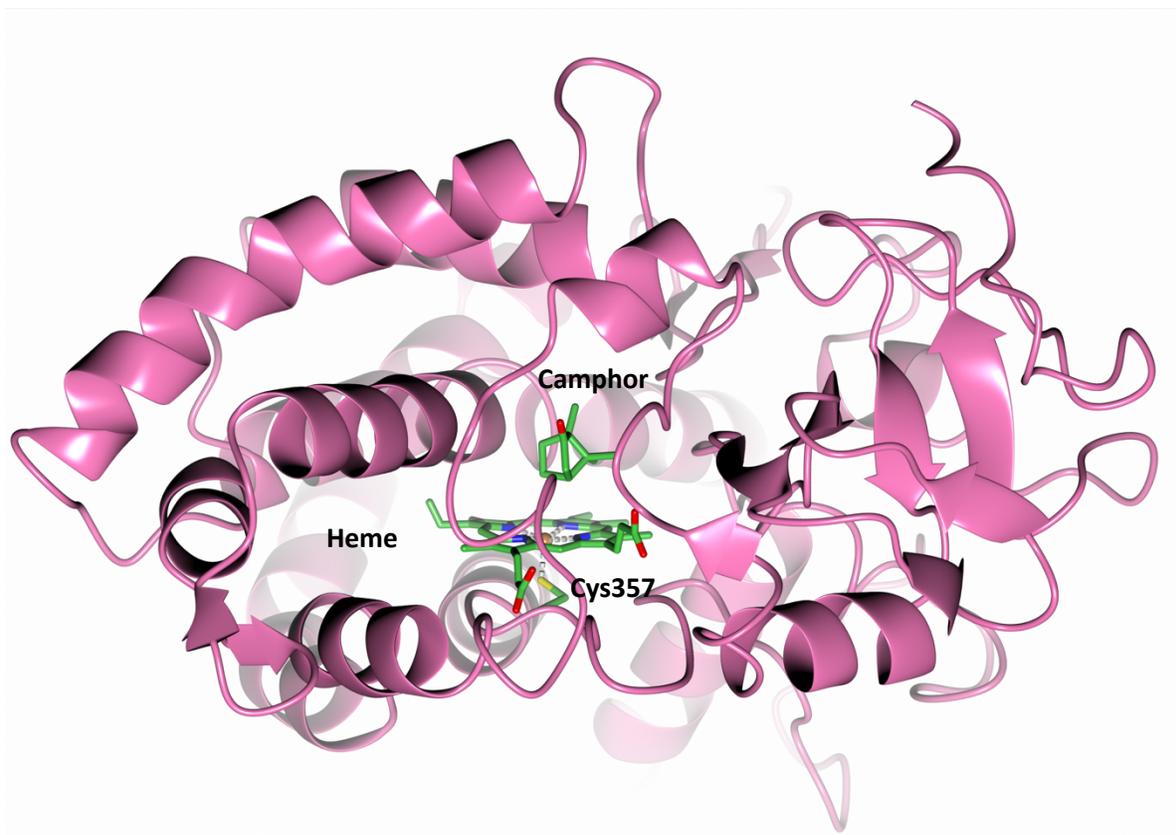


Figure 3.7 Overview of the CYP450cam structure (PDB 1DZ4). The heme iron is coordinated to the sulfur of a cysteine on the proximal side and the substrate camphor binds in the vicinity of the I-helix on the distal face. The substrate entry channel is bounded by the F-, G- and B-helices and FG-loop.

In the catalytic mechanism of CYP450cam, camphor binding results in a shift of the low spin ferric heme to a high spin state which also results in a change of the reduction potential from -303 mV to -173 mV (Figure 3.8).⁸² Following this, the heme can be reduced by a 11.6 kDa Fe₂S₂ ferredoxin redox partner protein Putidaredoxin which has a reduction potential of -240 mV.⁸³ Following this molecular dioxygen binds and is activated for substrate hydroxylation.

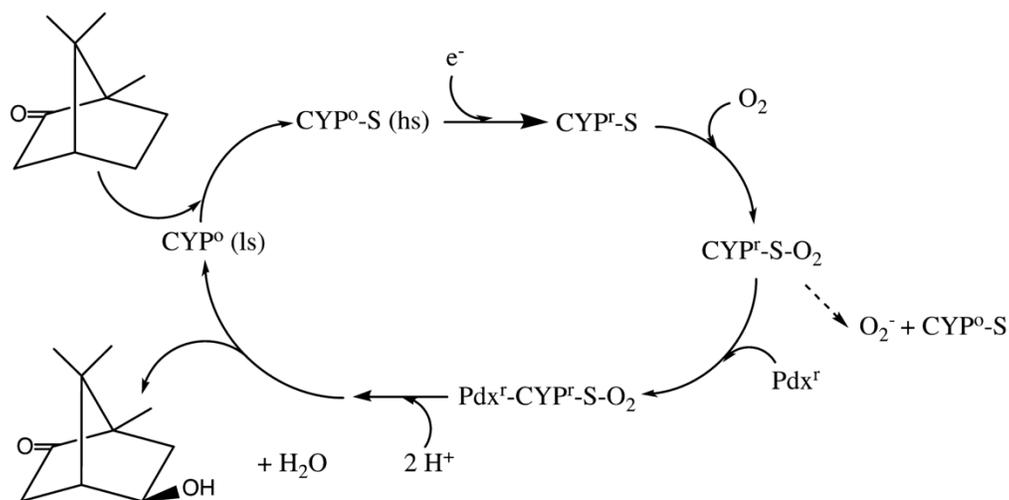


Figure 3.8 The catalytic cycle of CYP450cam. The oxidized CYP450cam (CYP⁰) in the low spin state (ls) binds the substrate, D-camphor and transitions to a high spin state (hs). The reduced CYP^r binds oxygen and reduced Pdx^r which can form the active complex to form the product 5-*exo*-camphor. If Pdx^r does not bind, the oxygen is reduced by the heme iron to superoxide and dissociates. Adapted from Pochapsky *et al.*⁸⁴ and reprinted with permission of the American Chemical Society (ACS). Please contact publisher for reuse.

In addition to its role in electron delivery to the CYP450cam heme, Pdx plays an important effector role, inducing conformational changes essential for substrate hydroxylation.⁸⁵⁻⁸⁹ During catalysis, the first reduction of CYP450cam can be achieved by non-native redox partners including small molecules, however Pdx is essential for the second reduction.^{90,91} Crystallographic studies by Tripathi *et al.* showed that substrate bound CYP450cam adopts a closed conformation which shifts toward an open conformation upon Pdx binding.⁹² This conformational change involves movement of the F and G helices and FG loop of CYP450cam that form the substrate binding channel. Conformational changes by Pdx binding also caused

Asp251 to be freed from a salt bridge interaction with Arg186 and Lys178, which triggers formation of the proton relay network required for camphor hydroxylation.⁹²

3.1.8 Neutron Studies of CYP450cam

The protonation pathway responsible for delivery of a proton to the peroxo intermediate during CYP450cam catalysis remains a field of active investigation as discussed in **Section 3.1.6**. Since detailed knowledge of this complex hydrogen bonding network is necessary to determine the exact architecture of the acid-alcohol pair and surrounding water molecules, neutron protein crystallography is the ideal probe. As discussed in **Chapter 1**, neutron protein crystallography allows the position of hydrogen (H) and deuterium (²H/D) atoms to be determined. In contrast to X-ray crystallography, where the magnitude of scattering is proportional to the number of electrons, the magnitude of neutron scattering is dependent on complex interactions with the nucleus.^{93,94} This makes light H/D atoms similarly visible to protein C, N, O and S atoms in neutron scattering length density maps.^{95,96} Thus protonation states, hydrogen bonding interactions and water molecule orientations are visible at moderate resolutions in contrast to the sub-atomic resolutions necessary for X-ray diffraction.⁹⁷ While there have been numerous studies on metalloproteins using neutron protein crystallography to elucidate their reaction mechanism as discussed in **Chapter 1.1**,⁹⁸ currently the only heme-center proteins characterized by neutron protein crystallography are the heme peroxidases.⁹⁹ Studies on the Class I heme peroxidases cytochrome *c* peroxidase (CcP) permitted characterization of compound I as a deprotonated Fe(IV)=O and with a positively charged, doubly protonated distal histidine.¹⁰⁰ Further neutron protein crystallography studies of ascorbate peroxidase (APX) permitted characterization of Compound II as a protonated Fe(IV)-OH.¹⁰¹ The structure further indicated that the distal histidine was doubly protonated and therefore positively charged. A recent neutron crystal structure investigating a distal arginine as a dynamic proton donor, found this residue to be deprotonated and neutral, supporting its capacity as a proton donor.¹⁰²

The studies of CcP and APX represent an excellent precedent of the information that can be gleaned from neutron protein crystallography of heme-center proteins. In particular, the capture of catalytic

intermediates permits further insight into the mechanism. Although there is currently no complete structure of CYP450cam, preliminary studies to determine its suitability to neutron studies have yielded promising results. Due to the negative coherent neutron scattering length and a high incoherent cross section of hydrogen, it is advantageous perdeuterate the protein and to exchange all the H atoms within the protein to D to improve data quality.^{103,104} Perdeuteration of a protein may however affect its physiochemical properties.¹⁰⁵ Studies by Meilleur *et al.* on CYP450cam have shown that the hydrogenated and perdeuterated forms are structurally and physiochemically similar and that a perdeuterated protein would therefore serve as a suitable system to further study CYP450cam catalysis.¹⁰⁶ To ascertain this, CYP450cam was expressed in hydrogenated and perdeuterated form and characterized by Fourier transform infrared (FTIR) spectroscopy. The thermal stability and unfolding was investigated by using the amide band absorption peaks. Additionally, step-scan time-resolved FTIR was used to study the binding of the CO ligand by monitoring the CO stretch vibration band and the amide band absorption peaks. It was found that perdeuteration does not affect CO binding, however the perdeuterated CYP450cam was less thermally stable. Meilleur *et al.* further demonstrated perdeuterated CYP450cam crystals were very suitable for diffraction experiments.¹⁰⁷ X-ray diffraction of crystals of perdeuterated CYP450cam that were >99% deuterium enriched diffracted to a resolution of 1.7 Å. Hydrogenated crystals of CYP450cam grown using the same conditions diffracted up to 1.4 Å. Structural analysis revealed no significant differences between the hydrogenated and perdeuterated forms. Taken together, these successful structural studies support the use of perdeuterated CYP450cam as a system for characterization by neutron protein crystallography.

3.1.9 Large crystal growth for neutron protein crystallography

While neutron protein crystallography provides mechanistically important structural information about the protonation states, hydrogen bonding interactions and water molecule orientations, it presents challenges due to neutron source flux limitations and weak sample scattering.¹⁰⁸ During neutron diffraction, the intensity of diffraction can be given by the following equation¹⁰⁹:

$$I = \frac{I_0 F^2 V A}{v_0^2}$$

where I is the diffraction intensity, I_0 is the incident neutron intensity, F is the structure factor, V is the volume of the crystal sample, A is the detector area subtended by the sample and v_0 is the unit cell volume. This equation illustrates that $I \propto V$ indicating that the larger the crystal, the better the signal-to-noise ratio, I/σ . Neutron protein crystallography, therefore, requires large crystals for successful data collection with a minimum volume of 0.1 mm^3 .^{110,111} It should be noted that in addition to crystal volume, the unit cell volume also contributes to the diffraction intensity.¹¹² To prevent close reflections that are difficult to discern due to overlap, large crystals with small unit cells are optimal.¹¹³

The number of neutron protein diffraction structures is significantly lower than the number of X-ray diffraction structures with neutrons constituting close to 190 deposited structures in the Protein Data Bank (PDB) in contrast to more than 160 352 X-ray structures as of October 2021.¹¹³ The majority of solved neutron structures utilized hydrogenated protein that had undergone H/D vapor exchange since perdeuteration can hamper crystal growth and has lower protein expression yields.¹¹⁴ Analysis of the submitted neutron crystal structures indicates that large crystal growth is achieved by vapor diffusion and batch crystallization, while dialysis and capillary counter-diffusion have also yielded a few successful results.^{103,113} Additionally, crystal seeding techniques have contributed to the success of the growth of larger crystals with improved diffraction.^{115,116}

3.1.10 General principles of protein crystallization

In biology, the central principle of crystallization is to achieve a state of supersaturation of a macromolecule that facilitates their aggregation subsequent growth into ordered assemblies.¹¹⁷ Crystallization proceeds in two phases termed nucleation and growth.¹¹⁸ Nucleation is the first step during which self-assembly into an organized structure for crystal growth occurs.¹¹⁹ These ordered nuclei continue to grow until a critical size is exceeded and crystal growth commences. In order to achieve a state of supersaturation where the protein concentration is greater than its solubility limit, a number of precipitants are used to reduce the proteins solubility.¹²⁰ These include salts, organic solvents, polymers or alcohols,

however, every protein is unique and the exact composition of precipitant that will yield successful crystallization can only be determined by trial and error.

Crystal growth can be described using kinetics and it is governed by equilibrium thermodynamics.¹²¹ Formation of nuclei and subsequent crystal growth can be illustrated with a phase diagram in which the states of the protein are described as a function of precipitant.¹²² This provides insight into the protein solubility at increasing precipitant concentrations and provides a set of conditions at which there is an equilibrium between solution and crystal growth.¹²³ When a protein crystallizes, a phase transition occurs with the solubility curve demarcating the a change in phase from protein in solution in an undersaturated state to four saturation zones: metastable, nucleation, precipitation and phase separation (Figure 3.9).¹²⁴ In the undersaturation zone, there is no nucleation or crystal growth. Increase in protein and precipitant concentration results in a supersaturation environment in which the concentration of the protein is greater than its solubility limit. The lowest level of supersaturation i.e., the metastable zone, does not have sufficient levels of supersaturation for nucleation to take place, however it presents the ideal environment for crystallization from formed nuclei or seeded crystals. The nucleation zone presents moderate supersaturation conditions in which proteins can aggregate to form nuclei from which crystals will grow. Further increasing protein and precipitant concentrations results in the precipitation zone where concentrations are too high for nucleation. An extreme case occurs when protein and precipitation concentrations are increased beyond the precipitation and phase separation, characterized by formation of droplets with an oily appearance.¹²⁵

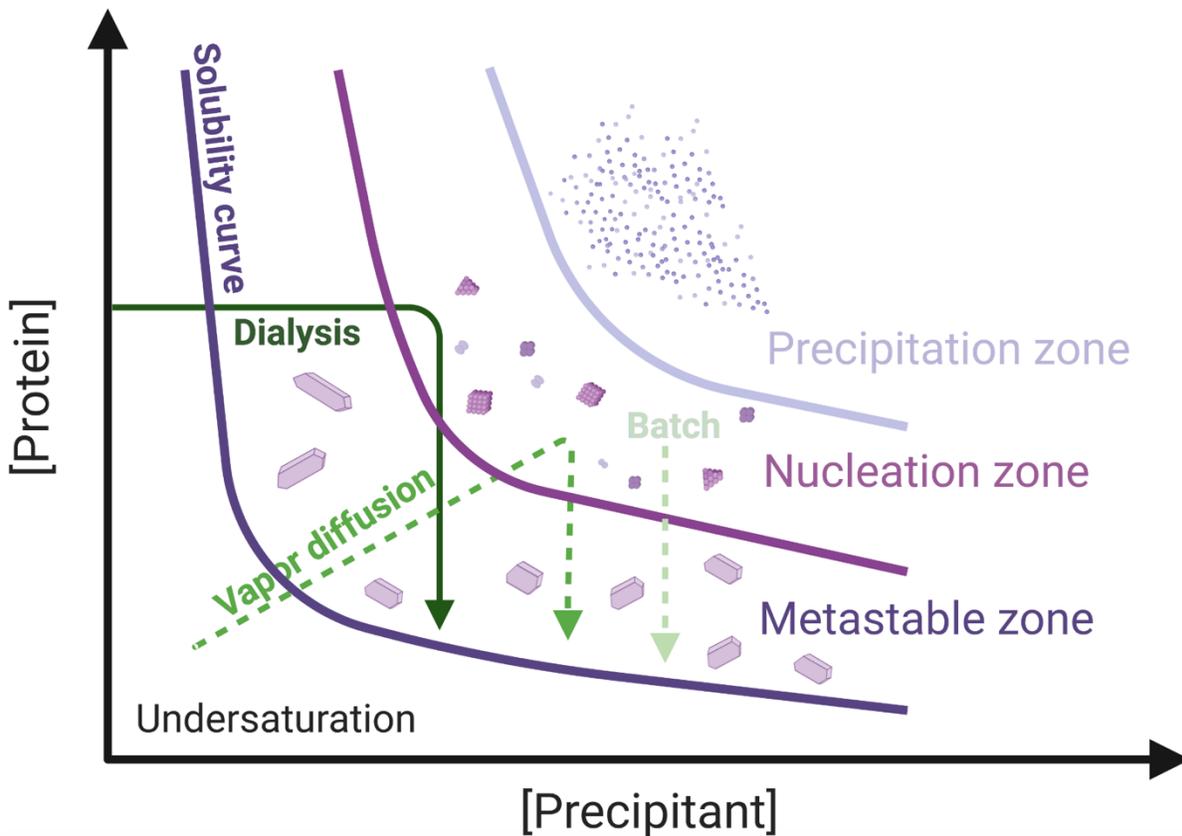


Figure 3.9 Crystallization phase diagram for proteins. The kinetic trajectories for dialysis, vapor diffusion and batch are shown in arrows. The undersaturation and supersaturation zones are separated by the solubility curve and the supersaturation zone is composed of the metastable, precipitation and nucleation zones. Figure prepared with BioRender.

Crystallization techniques represent different kinetic trajectories within this phase diagram to achieve nucleation and subsequent crystal growth and will be discussed below. Crystallization is a trial-and-error process that leverages on variations of precipitant type, precipitant concentration, pH, temperature and protein concentration to achieve successful crystal growth.¹²⁶ A plethora of precipitants are at the protein crystallographer's disposal, often conveniently provided as screens of several hundred conditions for evaluation and subsequent optimization.^{127,128} The primary crystallization techniques include vapor diffusion using both sitting and hanging drops, batch crystallization and crystallization by dialysis (Figure 3.10). While the apparatus used for protein crystallography can vary, the primary principle of achieving a state of supersaturation to enable nucleation and subsequent crystal growth remains a unifying concept.

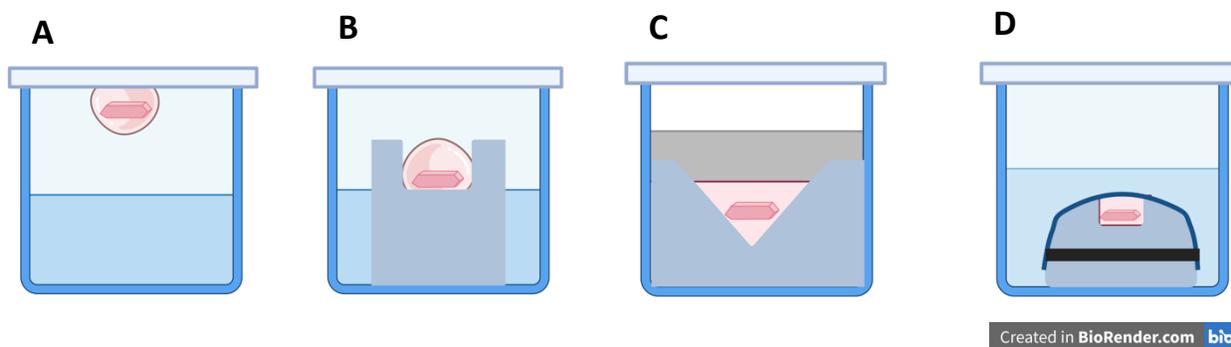


Figure 3.10 Illustration of the conventional crystallization methods. **(A)** Hanging drop vapor diffusion in which the protein droplet is suspended above the precipitant reservoir. **(B)** Sitting drop vapor diffusion in which the protein drop is sitting on a bridge above the precipitant reservoir. **(C)** Batch protein crystallization in which the protein and precipitant are combined and covered in a low-density non evaporative oil. **(D)** Dialysis crystallization in which the protein equilibrates with the precipitant through a semi-permeable dialysis membrane. Figure prepared using BioRender.

3.1.10.1 Vapor diffusion

Vapor diffusion is the most commonly applied protein crystallization technique.¹²⁹ In this technique, a droplet composed of protein in buffer is mixed with a precipitant solution is equilibrated against a reservoir with a higher concentration of precipitant.¹³⁰ The precipitant concentration within the protein droplet gradually increases as vapor diffusion to achieve equilibration of the drop and the reservoir occurs. During this equilibration with the reservoir, the drop will reach supersaturation conditions that encourage nucleation and crystal growth. In the practical application of this technique, the wells are sealed with a cover slide and vacuum grease or tape to create a closed environment. The protein droplet can be placed on a support above the precipitant reservoir in a “sitting drop” conformation, on the underside of the siliconized glass cover slide used to seal the well in a “hanging drop” or sandwiched between two siliconized glass cover slides.¹²⁶

3.1.10.2 Dialysis

Like the vapor diffusion method, dialysis also makes use of a gradual equilibration with a precipitant solution, but the equilibration occurs through a semi-permeable membrane.¹³¹ The membrane used should permit the movement of small molecules, such as salts, but prevent the diffusion of the protein. Generally, the protein in buffer is placed in a “button” which is covered in the semi-permeable membrane

of a certain molecular weight cutoff.¹³² This dialysis button is immersed in the precipitant and gradually the precipitant diffuses into the button creating conditions of supersaturation.

3.1.10.3 Batch

Batch crystallization can be thought of as an “end-point” technique because there is no gradual exploration of the phase diagram through evaporation or diffusion to achieve supersaturation.¹³³ In a batch set-up, the protein and buffer are mixed with precipitant to and nucleation must occur from these proportions of protein and precipitant. To perform a batch experiment, protein and precipitant are combined in vials to supersaturation levels and left undisturbed. An alternative to this larger scale method is microbatch in which droplets of protein mixed with precipitant are covered with an inert low density, non-evaporative oil such as paraffin, creating a closed environment.¹³⁴

3.1.10.4 Capillary counter-diffusion

This technique is also known as “free interface diffusion” since it permits free diffusion of the protein solution with the precipitant solution where they meet at an interface.¹³⁵ To achieve this a protein solution is placed in a capillary and the precipitant is carefully layered on top of this. The interface of the protein and precipitant undergoes diffuses mixing and a supersaturation environment is generated with crystals forming as a “nucleation front” progresses.¹³⁶ It is also possible to suspend the precipitant in a matrix or gel to regulate diffusion and prevent rapid mixing.

3.1.10.5 Strategies for neutron protein crystallography

The techniques applied to obtain crystal for X-ray diffraction are the same as those applied for neutron diffraction, with the primary difference being scale.^{104,110,137} While crystallization drops for X-ray diffraction usually range from 1 μL to 10 μL , neutron diffraction requires larger volumes of several hundred microliters to even the milliliter range.¹⁰³ Optimization of conditions which provide crystal “hits” is also highly advantageous. Optimization is performed by creating a grid with variations of a condition that has been found to promote crystal growth.¹³⁸ The parameters that can be adjusted include the precipitant concentration, the protein concentration and the pH; all within a range of the original condition that yielded

crystals.¹³⁹ Optimization permits improvement of the crystal morphology and allows the conditions that favor growth of large, single crystals to be established. Since larger drops result in a different thermodynamic behavior during diffusion, optimization of conditions that worked for small-scale crystallization may be necessary when bigger drops are used. Finally, seeding remains an invaluable technique to obtain large crystals.¹⁴⁰ Large crystals, by necessity, require large amounts of protein. It may occur that a high-quality crystal may have depleted all the protein material within a drop to continue growing in which case transference of this crystal to another drop with lower supersaturation levels i.e. the metastable zone to continue growth typically via vapor diffusion. Macroseeding is usually preceded by a step termed “etching” in which the crystal is placed in an unsaturated solution to remove the ultimate crystal layer that may contain malpositioned protein molecules which would adversely affect regular growth along the planes of the crystal surface.¹⁴¹ Microseeding is particularly useful if a condition that provides crystals has been established, however the morphology or quality of these crystals is undesirable.¹⁴² A nucleant solution is generated by crushing a crystal and feeding this into a protein precipitant solution that is less supersaturated and commencing with vapor diffusion crystallization. Seeding can also be performed by streak-seeding in which a fine hair/whisker is used to transfer crystal fragments from a nucleant solution or simply by streaking the hair gently across a drop containing crystal(s) to dislodge seeds.¹⁴³ The “seeded” hair/whisker is then streaked across the protein drop that will be exposed to vapor diffusion to deposit donor seeds for crystallization.

3.2 Aim of research

The wealth of structural, spectroscopic and computational research of CYP450cam has provided information on the reaction mechanism as well as proposed numerous putative pathways for intermediate protonation. Controversy of the source of these protons and their precise mechanism of delivery to the heme active site remain. Recent studies have proposed that, contrary to a protonation pathway involving the acid-alcohol pair Asp251-Thr252, a heme propionate/Arg299 pair may be involved in intermediate protonation.⁷⁸ Neutron protein crystallography represents the ideal technique to uncover the protonation pathway since it permits the visualization of H/D atoms. This will provide information on the protonation states of the active site residues, hydrogen bond interactions and the orientation of water molecules. Meilleur *et al.* have successfully perdeuterated and crystallized CYP450cam, establishing that this system is suited to neutron diffraction experiments by performing preliminary neutron diffraction data collection.^{106,107} The aim of this research is to replicate the successful crystallization of CYP450cam and optimize this for growth of large crystals. Once these large crystals are obtained, it will be possible to perform neutron diffraction studies of the CYP450cam resting state, as well as oxygen-activated reaction cycles intermediates to elucidate the protonation pathway and reaction mechanism.

3.3 Materials and methods

3.3.1 Reagents

All reagents, unless otherwise stated were obtained from Fischer Scientific (Hampton, New Hampshire, USA) and Sigma-Aldrich (St. Louis, Missouri, USA) and purifications columns were from GE Healthcare Life Sciences (Chicago, Illinois, USA).

3.3.2 Protein Expression

The CYP101A1 (CYP450cam) from *Pseudomonas putida* with the C334A mutation was cloned into the pET-M11 vector (EMBL).¹⁴⁴ The C334A mutation prevents dimerization and has been shown to not influence activity.¹⁴⁵ This CYP450cam contains a N-terminal His₆-tag with a TEV protease site and will be termed CYP450cam(C334A)-His. The CYP450cam C334A mutant was additionally synthesized (Genscript) into the pET-15b vector (Novagen). This CYP450cam contains no His₆-tag and will be termed CYP450cam(C334A). CYP450cam(C334A)-His₆ and CYP450cam(C334A) were overexpressed in *E. coli* BL21(DE3) grown in Luria-Bertani broth for to an OD₆₀₀ of 0.7 after which expression was induced with 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) supplemented with 1 mM 5-aminolevulinic acid (heme precursor) and 50 μM FeCl₃ for 24 h at 25°C. Cell pellets were resuspended in either 50 mM KP_i pH 7.6, 50 mM KP_i pH 6 or 50 mM Tris pH 7.8 which each also contained 1 mM D-camphor and 0.5 mg/mL lysozyme. Cells were lysed by sonication and centrifuged at 30 000g for 45 min.

3.3.3 Protein purification

3.3.3.1 Ni-affinity purification

P450cam(C334A)-His₆ lysate was loaded onto a HisTrap HP column equilibrated with 50 mM KP_i pH 7.6, 250 mM KCl, 1 mM D-camphor and 20 mM imidazole pH 7.6. Elution was performed using a linear gradient of 50 mM KP_i pH 7.6, 250 mM KCl, 1 mM D-camphor and 500 mM imidazole pH 7.6. The eluate was pooled from a peak obtained monitoring the absorbance peak at 391 nm and fractions were concentrated using a 30 kDa cutoff centrifugal filter (Amicon).

3.3.3.2 Anion exchange purification

P450cam(C334A) lysate loaded onto either a DEAE HiTrap DEAE FF equilibrated with 50 mM KP_i pH 6 and 1 mM camphor or a HiTrap Q XL column equilibrated with 50 mM Tris pH 7.8 and 1 mM camphor. Elution was performed using a linear gradient of 50 mM KP_i pH 6 or 50 mM Tris pH 7.8 for the DEAE and Q-columns, respectively, as well as 1 mM D-camphor and 1 M KCl. The eluate was pooled from a peak obtained monitoring the absorbance peak at 391 nm and fractions were concentrated using a 30 kDa cutoff centrifugal filter (Amicon).

3.3.3.3 Size exclusion chromatography

P450cam(C334A)-His₆ and CYP450cam(C334A) from the affinity anion exchange purification was loaded onto a Superdex 75 prep grade, HiLoad 16/60, 120 mL column equilibrated with a wash buffer composed of 50 mM KP_i pH 7.6, 250 mM KCl and 1 mM D-camphor. The eluate was pooled from a peak obtained monitoring the absorbance peak at 391 nm and fractions were concentrated using a 30 kDa cutoff centrifugal filter (Amicon).

3.3.4 Hemin reconstitution

To ensure that any apoprotein without heme coordinated in the active site would be reconstituted with heme, hemin chloride was added to purified CYP450cam according to the protocol by Auclair *et al.*¹⁴⁶ A hemin solution was prepared in which 24 mg of hemin chloride was dissolved in 330 μ L 0.1 M NaOH. This was diluted to 30 mL in 50 mM KP_i pH 7.6 and 4 mM D-camphor. 10 mL of this hemin solution was added to 57 mg of CYP450cam(C334A) in 190 mL of 50 mM Tris pH7.4, 250 mM KCl and 1 mM camphor. This reaction was stirred overnight at room temperature. A buffer exchange was performed using a 30 kDa cutoff centrifugal filter (Amicon) to 50 mM KP_i pH 7.6, 250 mM KCl and 1 mM camphor. This gave a final concentration of \sim 60 μ M hemin chloride added to 6 μ M CYP450cam.

Alternatively, the Lee *et al.* protocol was followed in which 0.1 mM hemin chloride added to the *E. coli* BL21(DE3) that had expressed CYP450cam lysis buffer.¹⁴⁷

3.3.5 Characterization

3.3.5.1 SDS-PAGE

Gel electrophoresis was performed on whole cells following expression as well as purification fractions using sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE).¹⁴⁸ Samples were run on Mini-PROTEAN TGX Stain-Free Precast Gels (BioRad Laboratories, Hercules, California, USA) and performed according to the manufacturer's protocol using 5 μ L of the Precision Plus Protein Unstained Standards (BioRad Laboratories, Hercules, California, USA).

3.3.5.2 Native SDS-PAGE

Blue native PAGE (BN-PAGE) was performed using precast Bis-Tris 4 – 16% gels from the NativePAGE™ Novex® System (Life Technologies, Carlsbad, California, USA) in the XCell SureLock minicell with 5 μ L of NativeMark unstained protein standard and stained with Coomassie G-250 for analysis.¹⁴⁹

3.3.5.3 Absorption spectrum

UV-Vis spectroscopy was performed using a Shimadzu UV-VIS spectrophotometer (SHIMADZU SCHWEIZ GmbH, Switzerland) using 1 cm pathlength quartz cuvettes.

3.3.5.4 CO-difference spectrum

Whole cells or purified CYP450cam(C334A)/P450camC334A-His₆ was used for CO difference spectroscopy. CYP450cam was reduced with a few grains of sodium dithionite in duplicate and one of the samples was exposed to CO in a sealed environment for 10 min. Absorbance of the CO-bound and free CYP450cam was measured from 360 – 500 nm in a microtiter plate using a SpectraMax i3 spectrophotometer (Molecular Devices, San Jose, California, USA).

3.3.6 Crystallization

3.3.6.1 Vapor diffusion

Crystallization was performed using purified CYP450cam(C334A) or CYP450camC334A-His₆ exchanged into a buffer composed of Tris pH7.4 and 1 mM camphor using a 30 kDa cutoff centrifugal filter (Amicon). Crystallization was performed with CYP450cam in a concentration range of 5 – 40 mg/mL. Crystallization was performed in both sitting and hanging drops in 24-well Cryschem and VDX plates (Hampton Research, Aliso Viejo, California, USA) sealed with tape or 22 mm siliconized cover slides. A range of conditions were used for crystallization with polyethylene glycol (PEG) PEG 3350 - PEG 12 000 in a 2% - 22% range, 100 – 250 mM KCl, 1 – 2 mM D-camphor and 50 mM Tris in a 6.5 – 8.5 pH range and temperatures ranging from 4°C - 25°C. Protein drops were added to precipitant in multiple ratios (1:1, 1:2, 1:3, 2:1, 3:1) and drop sizes ranged from 2 µL – 60 µL. Additionally, large sitting drops ranging from 100 µL– 800 µL were set up in the 9 well glass plate and sandwich box setup (Hampton Research, Aliso Viejo, California, USA).

3.3.6.2 Microbatch

Crystallization was performed using purified CYP450cam(C334A) or CYP450camC334A-His₆ exchanged into a buffer composed of in Tris pH7.4 and 1 mM camphor using a 30 kDa cutoff centrifugal filter (Amicon). Microbatch crystallization was performed using 72 well Greiner microbatch plates (Hampton Research, Aliso Viejo, California, USA) and protein drops were submerged in paraffin or Al's oil. Crystallization was performed with CYP450cam in a concentration range of 5 – 40 mg/mL. A range of conditions were used for crystallization with PEG 3350 - PEG 12 000 in a 2% - 22% range, 100 – 250 mM KCl, 1 – 2 mM D-camphor and 50 mM Tris in a 6.5 – 8.5 pH range and temperatures ranging from 4°C - 25°C. Protein drops were added to precipitant in multiple ratios (1:1, 1:2, 1:3, 2:1, 3:1) and drop sizes ranged from 2 µL – 20 µL.

3.3.6.3 Microdialysis

Crystallization was performed using purified CYP450cam(C334A) or CYP450camC334A-His₆ exchanged into a buffer composed of in Tris pH7.4 and 1 mM camphor using a 30 kDa cutoff centrifugal filter (Amicon). Dialysis was performed using the dialysis button setup from Hampton (Hampton Research, Aliso Viejo, California, USA) with buttons ranging from 5 μ L – 350 μ L with a Spectra/Por6 25 kDa cutoff dialysis membrane (Repligen, Waltham, Massachusetts, USA). Crystallization was performed with CYP450cam in a concentration range of 5 – 40 mg/mL. A range of conditions were used for crystallization with PEG 3350 - PEG 12 000 in a 2% - 22% range, 100 – 250 mM KCl, 1 – 2 mM D-camphor and 50 mM Tris in a 6.5 – 8.5 pH range and temperatures ranging from 4°C - 25°C.

3.3.6.4 Screens

Crystallization was performed using purified CYP450cam(C334A) or CYP450camC334A-His₆ exchanged into a buffer composed of in Tris pH7.4 and 1 mM camphor using a 30 kDa cutoff centrifugal filter (Amicon). Crystallization was performed with CYP450cam in a concentration range of 5 – 40 mg/mL.

A number of screens from Qiagen (Hilden, Düsseldorf, Germany) including JCSG Core I-IV. PEGs Suite I-III and NexTal MB Class Suite I – II were used. Proteins were crystallized in a 1:1 ratio with precipitant in drops ranging from 1 μ L – 2 μ L at temperatures from 4°C - 25°C.

3.3.6.5 Microseeding

Formed crystals were harvested and microseeds were generated using Seed Beads (Hampton Research, Aliso Viejo, California, USA). Stabilizing solutions for microseed solutions were generated using the well conditions for subsequent crystallization in a dilution series of 10^0 , 10^{-1} , 10^{-2} , 10^{-3} , 10^{-4} and 10^{-5} . Drops composed of 1 μ L protein and 1 μ L precipitant pre-equilibrated overnight at 6°C and 1 μ L seed was added to this. Crystallization was performed with CYP450cam at a concentration range of 20mg/mL. Crystallization conditions contained PEG 3350 - PEG 12 000 in a 2% - 22% range, 250 mM KCl, 1 mM D-camphor and 50 mM Tris pH 7.6.

3.3.7 X-ray diffraction

X-ray diffraction data at 295 K were collected on a Rigaku MicroMax-007 HF X-ray generator coupled with a Dectris EIGER R 4M detector. Collected data was reduced with CrysAlis^{Pro} software suite (Rigaku, The Woodlands, Texas, USA) and scaled and merged with *AIMLESS* in the *CCP4* suite^{150,151}. Structure solution and refinement of the X-ray diffraction data was performed using the *PHENIX* software suite with manual model building to fit electron density maps in *Coot*.^{152–154}

3.4 Results and discussion

3.4.1 Protein expression and purification

In order to successfully obtain crystals of sufficient size for neutron protein crystallography, sufficient quantities of pure, high-quality protein are necessary. With respect to CYP450cam, this necessitates expression of a correctly folded protein with the heme prosthetic group incorporated. To ensure this, expression was performed with the addition of 5-aminolevulinic acid, a heme precursor,¹⁵⁵ and FeCl₃. Analysis of expression indicated that both CYP450cam(C334A)-His₆ and the CYP450cam(C334A) without the tag showed expression. This was seen on an SDS-PAGE gel as a band at approximately 47 kDa (Figure 3.11).

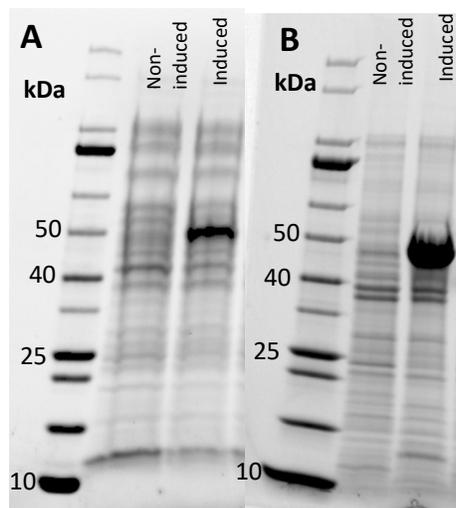


Figure 3.11 SDS-PAGE of protein expression. (A) CYP450cam(C334A)-His₆. (B) CYP450cam(C334A) without a tag.

To evaluate expression of CYP450cam with heme incorporated into the active site, a CO-difference spectrum analysis was performed. The active site of CYP450cam consists of an iron protoporphyrin IX complex with the iron coordinated to the sulfur of a deprotonated cysteine on the proximal face.⁶⁸ Reduction of CYP450cam with dithionite and exposure to CO results in formation of a peak at 446 nm termed the Soret band.^{156,157} The Soret band is an indicator that a CYP450 is in its native state and its attributed to mixing of the cysteine thiolate sulfur and heme iron orbital charge transfer transition with the porphyrin π -

π^* transitions.¹⁵⁸ Displacement of the thiolate or protonation represent loss of the native active state and are characterized by a Soret absorbance maximum at 420 nm. This P420 species is considered catalytically inactive.¹⁵⁹ The CO-difference spectrum is obtained by subtracting the oxidized spectrum from the reduced form. Appearance of a peak at 450 nm following whole cell CO difference spectroscopic analysis indicate that expression of CYP450cam in a native form had been achieved (Figure 3.12).

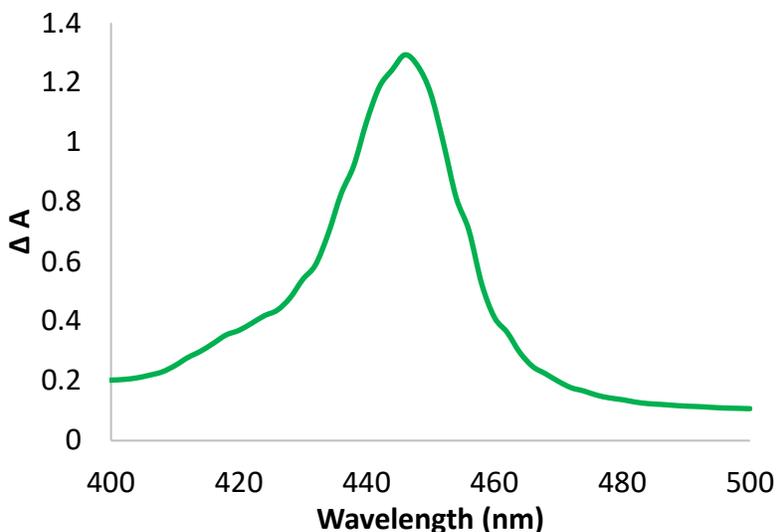


Figure 3.12 Representative whole cell CO-difference spectrum of CYP450cam after expression.

Following confirmation that the native form of CYP450cam had been expressed for both CYP450cam(C334A)-His₆ and CYP450cam(C334A), purification was performed by Ni-affinity and anion exchange chromatography.

3.4.1.1 His-tag dependent purification

His-tag dependent purification was performed using immobilized metal ion affinity (IMAC) with a Ni²⁺-NTA column. During this purification, the N-terminal hexahistidine (His₆) tag of the CYP450cam binds to the Ni²⁺ charged resin.¹⁶⁰ Imidazole, which competes with the His₆-tag for binding, is used for elution of the protein. As a first step, the cell lysate of CYP450cam(C334A)-His₆ was purified on a HisTrap HP column which resulted in enrichment of the CYP450cam(C334A)-His₆ (Figure 3.13 A). Size exclusion chromatography (SEC) in which molecules are separated according to their size because of their different

migration rates through a porous gel or polymer.¹⁶¹ One round of SEC was performed after affinity purification (Figure 3.13 B).

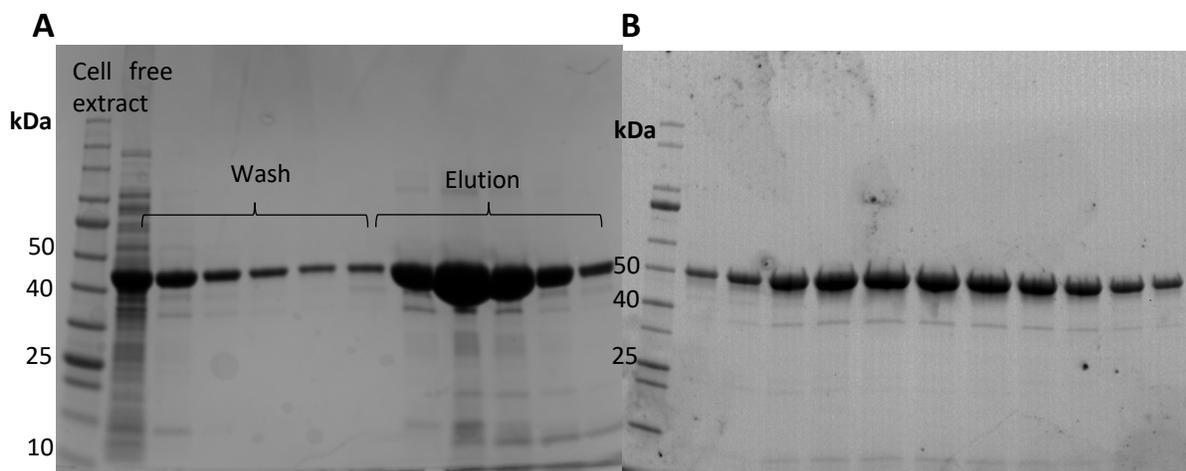


Figure 3.13 SDS-PAGE CYP450cam(C334A)-His₆ after affinity purification. **(A)** After Ni-affinity column purification. **(B)** Peak fractions after size exclusion column purification.

3.4.1.2 Anion exchange chromatography

In addition to the purification of the His₆-tagged form of CYP450cam, untagged CYP450cam was also purified using anion exchange chromatography. Ion exchange chromatography separates proteins according to their charge. Proteins are eluted by use of a salt which has a high ionic strength, competing with the bound protein and displacing it. In this case, anion exchange chromatography was used in which a positively charged column binds to negatively charged proteins. The strength of an anion exchanger depends on its ability to maintain its positive charge over a range of pH conditions. To ensure strong binding to the anion exchange column, it is necessary to ensure that the net charge of the protein is negative, therefore the pH of the buffer should be greater than the isoelectric point (pI) of the protein.

In order to determine the best protocol, two columns were tested: the weak anion exchanger diethyl aminoethyl (DEAE-column) and the strong anion exchanger (Q-column). The pI of CYP450cam(C334A) is 5.24, therefore buffers with a pH > 5.24 were used to ensure a net negative charge on CYP450cam.

3.4.1.3 DEAE column

The DEAE column was used with a equilibration and wash buffer composed of 50 mM KP_i pH 6 and 1 mM camphor and gradient elution was performed using with KCl. While protein did bind to the column, and elution with KCl did yield some CYP450cam enrichment, it was apparent that along with CYP450cam, there were many additional cellular proteins binding to the column. This was apparent after SEC which did not yield a pure CYP450cam (Figure 3.14).

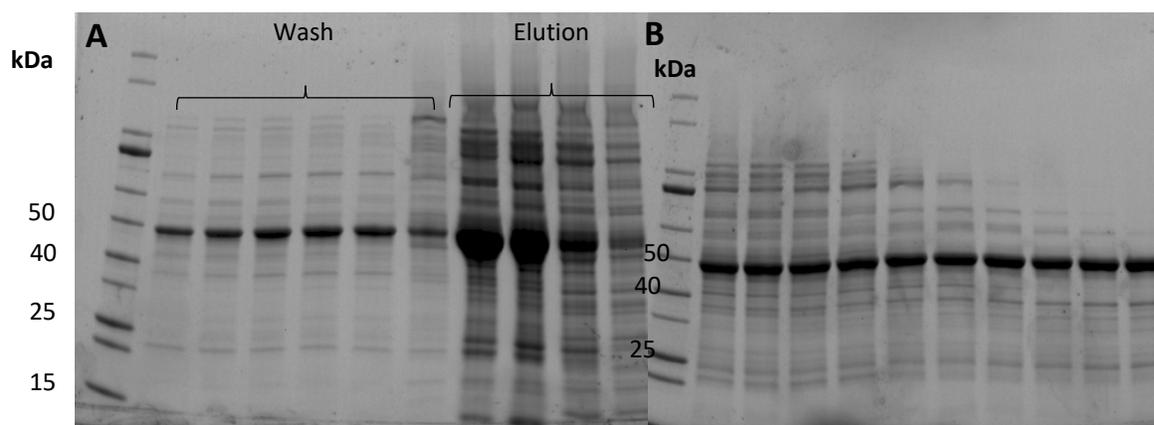


Figure 3.14 SDS-PAGE CYP450cam(C334A) after anion exchange purification. **(A)** After DEAE anion exchange column purification. **(B)** Peak fractions after size exclusion column purification.

3.4.1.4 Q-column purification

Due to the limited purity using the DEAE weak anion exchange column, a strong Q-column anion exchanger was used instead with 50 mM Tris pH 7.8 and 1 mM camphor as the equilibration and wash buffer. Subsequent purification of CYP450cam yielded a more enriched sample, indicating that the Q-column along with the Tris buffer were more effective as was also seen in a reasonable pure sample after SEC (Figure 3.15).

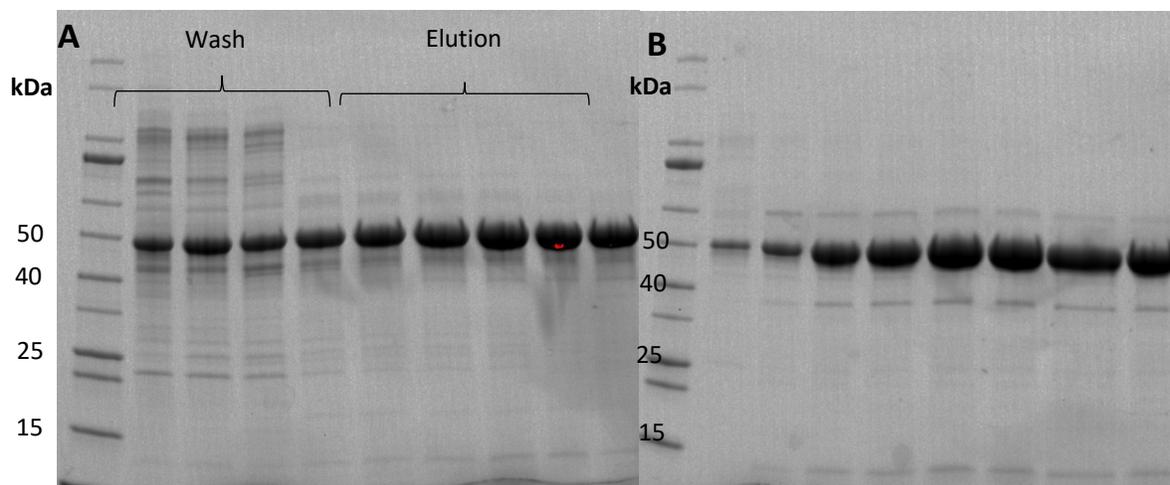


Figure 3.15 SDS-PAGE CYP450cam(C334A). **(A)** After Q-column anion exchange column purification. **(B)** Peak fractions after size exclusion column purification.

While both Ni-affinity and Q-column anion exchange purification yielded reasonably pure protein, there were a few impurities still visible in SDS-PAGE gels following one round of SEC. To circumvent this, a second round of SEC was performed which yielded pure CYP450cam (Figure 3.16).

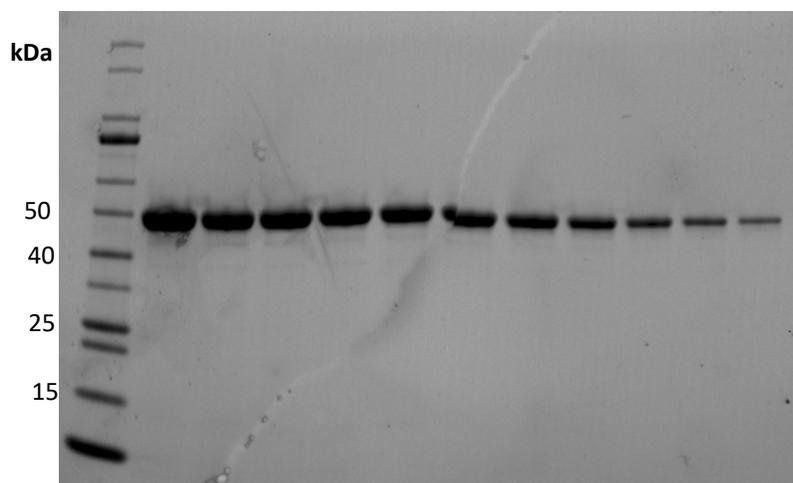


Figure 3.16 Representative SDS-PAGE of pure CYP450cam used for crystallization shown in a concentration range.

3.4.2 Heme reconstitution

While a protocol had been established to obtain pure CYP450cam, spectroscopic analysis indicated that heme incorporation levels were low. While a CO-difference spectrum confirms the presence of correctly folded native CYP450cam, it does not provide information on the apoprotein which is devoid of

its iron protoporphyrin IX prosthetic group. The high spin substrate bound CYP450cam exhibits an absorbance peak at 391 nm.¹⁶² A further indication of heme incorporation and purity is the “Rz” value in which camphor bound CYP450cam with a $A_{391}:A_{280}$ ratio of greater than 1.45 is considered a native CYP450cam with approximately 90% purity.^{162–166} Additionally, the low-spin substrate free form of CYP450cam exhibits an absorbance peak at 417 nm.^{162,167} Initial fractions of CYP450cam showed very low $A_{391}:A_{280}$ values of ~ 0.46 (Figure 3.17).

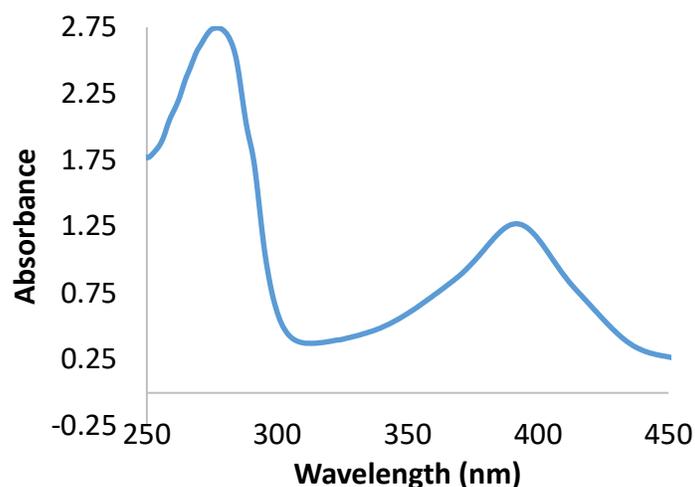


Figure 3.17 Absorbance spectrum of CYP450cam to determine sample purity.

In order to improve heme incorporation, a heme reconstitution protocol may be followed in which hemin chloride is added to the purified protein. In order to reconstitute the heme, the protocol by Auclair *et al.* was followed in which hemin is added in a $\sim 10:1$ ratio to CYP450cam.¹⁴⁶ Addition of the hemin chloride and subsequent spectroscopic analysis after buffer exchange indicated an improvement in the incorporation levels to ~ 1.44 in comparison to 0.55 before the addition of hemin (Figure 3.18). However, a CO difference spectrum indicated a prominent P420 peak, suggesting incorrect heme coordination (Figure 3.19). Indeed, analysis of the absorbance spectrum indicates significant peak broadening around 350 – 400 nm. This is a hallmark of unbound hemin in the hydroxide form at 386 nm as well as a μ -oxo bridged hemin. Therefore, the hemin had not been correctly incorporated into the apoprotein and excess unbound hemin was potentially still present.

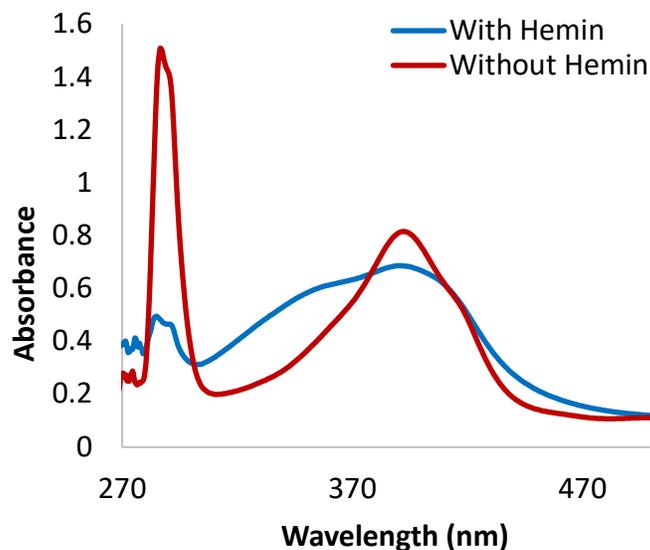


Figure 3.18 Absorbance spectrum of CYP450cam to determine sample purity and heme incorporation after hemin addition.

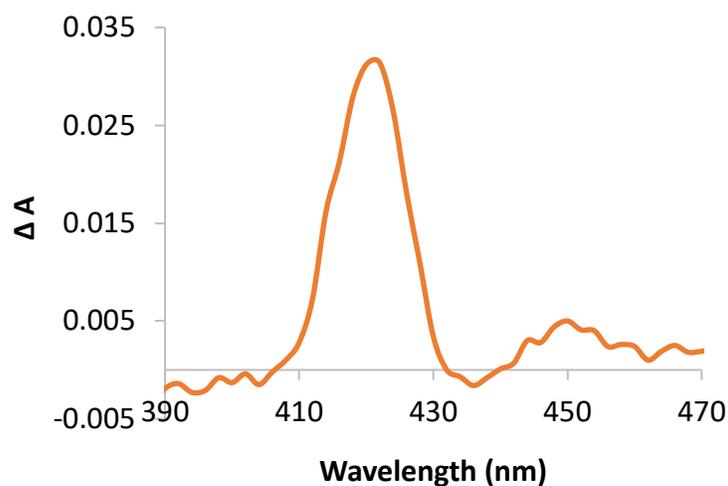


Figure 3.19 Representative whole cell CO-difference spectrum of CYP450cam after addition of hemin.

To avoid the presence of excess hemin and minimize incorrectly incorporated heme, the protocol applied by Lee *et al.* was utilized.¹⁶⁸ In this heme reconstitution method, hemin is added to the *E.coli* BL21(DE3) cells following CYP450cam expression in the lysis buffer. This ensures that hemin is present for reconstitution, but excess is removed during subsequent purification. Following purification, heme

incorporation was improved with a $A_{391}:A_{280}$ ratio RZ ratio above the required 1.46 (Figure 3.20). Additionally, the CO difference spectrum no longer displayed a P420 peak (Figure 3.21).

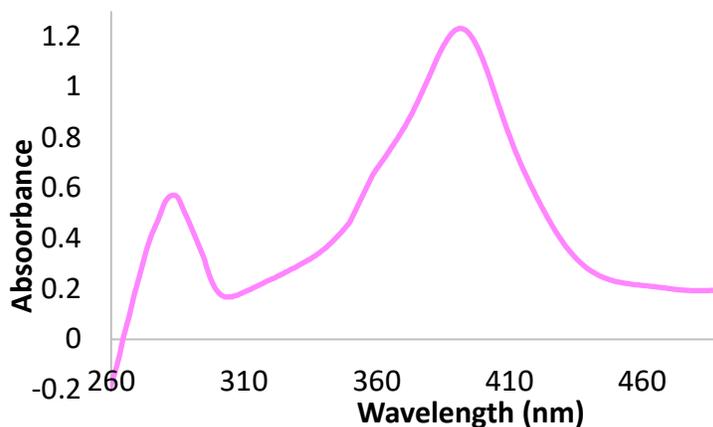


Figure 3.20 Absorbance spectrum of CYP450cam to determine sample purity and heme incorporation after the Lee *et al.* protocol was used for heme incorporation.

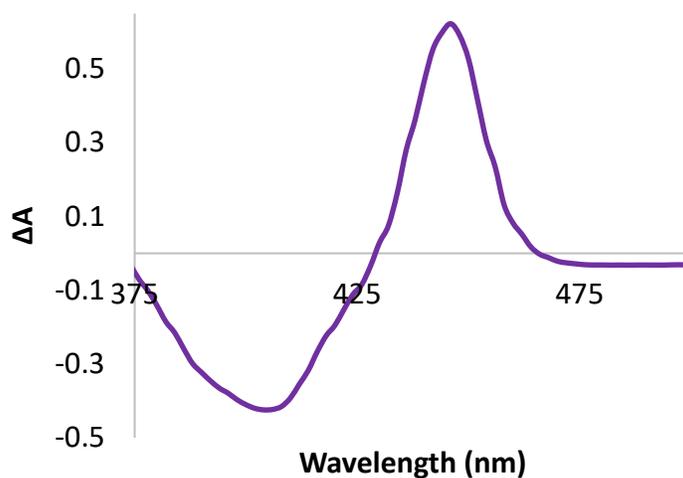


Figure 3.21 Representative CO-difference spectrum of CYP450cam after the Lee *et al.* protocol was used for heme incorporation.

3.4.3 Native SDS-PAGE

In order to verify that CYP450cam CYP450cam(C334A)-His₆ and CYP450cam(C334A) were not forming dimeric assemblies a blue native polyacrylamide gel electrophoresis (BN-PAGE) was run. Like SDS-PAGE, BN-PAGE enables protein separation in a polyacrylamide gel.^{148,149} However, BN-PAGE

allows the resolution of complexes since it does not use a strong ionic detergent and reductants such as β -mercaptoethanol (β ME) which cause protein unfolding.¹⁶⁹ SDS-PAGE relies on protein separation according to their mass since one molecule of SDS binds to every two amino acids ensuring a uniform charge-to-mass ratio in the denatured protein.¹⁷⁰⁻¹⁷² BN-PAGE does not use denaturants, instead using Coomassie Blue-G250 dye to impart proteins with a negative charge and facilitate their migration through the electric field applied to the polyacrylamide matrix.^{173,174} During BN-PAGE separation is dependent on the mass, size and structure/shape of the protein which migrates through the polyacrylamide gel pores until it arrives at its pore size limit.¹⁷⁵ Analysis of CYP450cam using BN-PAGE indicated that no dimeric complexes were formed (Figure 3.22). This further verifies that introduction of the C334A mutation prevents dimeric assemblies.¹⁴⁵

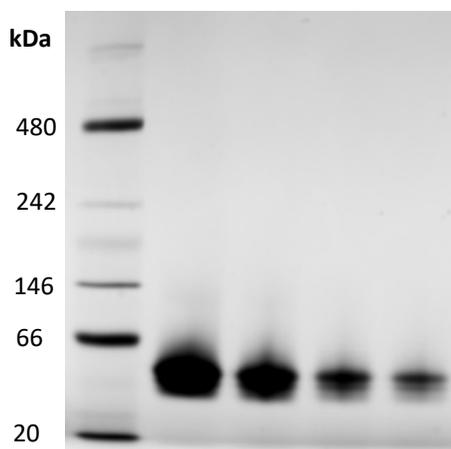


Figure 3.22 Blue-native PAGE of CYP450cam(C334A) shown in a concentration range.

3.4.4 Crystallization

P450cam crystallizes well in PEG, however the crystal morphology is often that of needles. While needles are sufficient to collect X-ray diffraction data, neutron protein crystallography requires crystals that have a minimum size of 0.1 mm^3 . Attempts were therefore made to grow CYP450cam crystals with a less elongated shape. Attempts at crystallization were successful, however it was not possible to obtain crystals other than needles or thin plates for either CYP450cam(C334A)-His₆ or CYP450cam(C334A).

3.4.4.1 Vapor diffusion

Sitting drop vapor diffusion experiments were conducted in a range of volumes 2 μL – 60 μL . Precedence by Meilleur *et al.* for successful crystallization using polyethylene glycol (PEG) was used as a starting condition for optimization.¹⁰⁷ A range of conditions were used for crystallization with polyethylene glycol (PEG) PEG 3350 - PEG 12 000 in a 2% - 22% range, 100 – 250 mM KCl, 1 – 2 mM D-camphor and 50 mM Tris in a 6.5 – 8.5 pH range and temperatures ranging from 4°C - 25°C. Protein drops were added to precipitant in multiple ratios (1:1, 1:2, 1:3, 2:1, 3:1). Shown below are representative results of these crystallization conditions (Figure 3.23). The morphology remained needle and plate-like even when temperatures as well as protein concentrations were lowered to slow the nucleation and crystal growth.

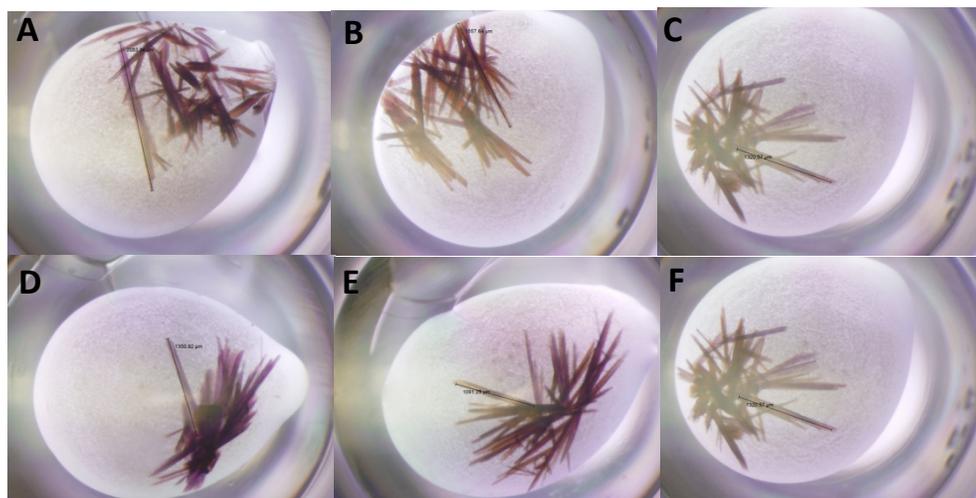


Figure 3.23 Crystallization trials using sitting drop vapor diffusion. 22 mg/mL CYP450cam in a 6 μL drop with 1:1 protein to reservoir ratio. (A) 50 mM Tris pH = 7.4, 15.6% PEG4000, 250 mM KCl, 1 mM camphor (B) 50 mM Tris pH = 7.4, 16% PEG8000, 250 mM KCl, 1 mM camphor (C) 50 mM Tris pH = 7.4, 16% PEG6000, 250 mM KCl, 1 mM camphor (D) 50 mM Tris pH = 7.4, 14% PEG6000, 250 mM KCl, 1 mM camphor.

Large sitting drop were additionally set up to determine whether this could encourage large crystal growth or whether the altered vapor diffusion rates on a larger drop may contribute to crystal growth with a different morphology. Needle crystals were still obtained however some drops remained clear without any crystal growth (Figure 3.24).

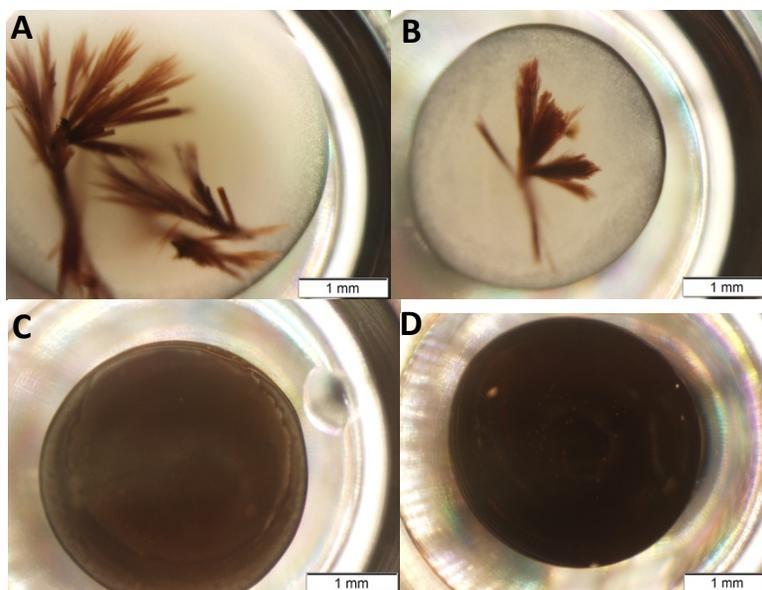


Figure 3.24 Crystallization trials using large scale sitting drop vapor diffusion. 35 mg/mL CYP450cam in a 60 μ L drop with 1:1 protein to reservoir ratio. (A) 50 mM Tris pH = 7.4, 16% PEG8000, 250 mM KCl, 2 mM camphor (B) 50 mM Tris pH = 7.4, 17% PEG6000, 250 mM KCl, 2 mM camphor (C) 50 mM Tris pH = 7.4, 15.2% PEG6000, 250 mM KCl, 2 mM camphor (D) 50 mM Tris pH = 7.4, 10% PEG4000, 250 mM KCl, 2 mM camphor.

3.4.4.2 Microbatch

Microbatch crystallization was performed under oil (Paraffin or Al's Oil). A range of conditions were used for crystallization with polyethylene glycol (PEG) PEG 3350 - PEG 12 000 in a 2% - 22% range, 100 – 250 mM KCl, 1 – 2 mM D-camphor and 50 mM Tris in a 6.5 – 8.5 pH range and temperatures ranging from 4°C – 25°C. Since these microbatch drops were run as trials, smaller volumes of 2 μ L – 20 μ L were used. Protein drops were added to precipitant in multiple ratios (1:1, 1:2, 1:3, 2:1, 3:1). The morphology continued to be needle-like, however the dynamics of growth appeared different with the needles appearing to grow radially from a point of origin (Figure 3.25).

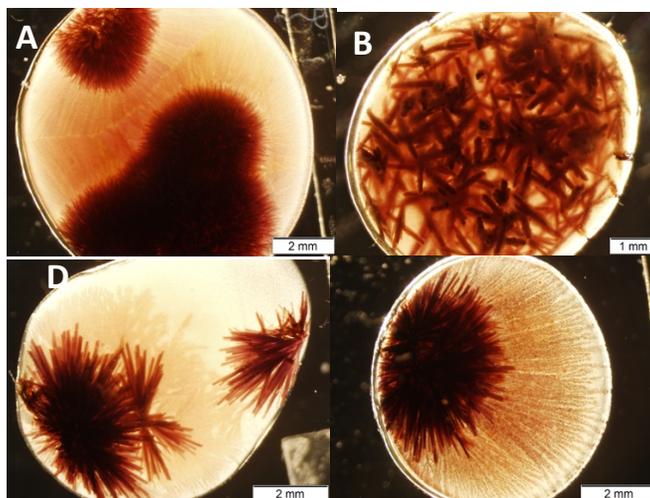


Figure 3.25 Crystallization trials using microbatch under paraffin oil. 18 mg/mL CYP450cam in a 4 μ L drop. (A) 50 mM Tris pH = 7.4, 25% PEG3350, 250 mM KCl, 1 mM camphor, 3:1 protein to reservoir ratio (B) 50 mM Tris pH = 7.4, 16.8% PEG4000, 250 mM KCl, 1 mM camphor, 2:1 protein to reservoir ratio (C) 50 mM Tris pH = 7.4, 15.2% PEG6000, 250 mM KCl, 2 mM camphor, 1:2 protein to reservoir ratio (D) 50 mM Tris pH = 7.4, 10% PEG8000, 250 mM KCl, 1 mM camphor, 3:1 protein to reservoir ratio.

3.4.4.3 Microdialysis

Meilleur *et al.* had successfully used microdialysis buttons for the growth of crystals large enough for neutron diffraction tests. Therefore, microdialysis buttons with a range of sizes from 5 μ L – 350 μ L were used. Consistent with the conditions used to successfully grow voluminous crystals by Meilleur *et al.*, a range of conditions were used for crystallization with PEG 3350 - PEG 12 000 in a 2% - 22% range, 100 – 250 mM KCl, 1 – 2 mM D-camphor and 50 mM Tris in a 6.5 – 8.5 pH range and temperatures ranging from 4°C - 25°C. Crystals grown were up to 2 mm in length, but still in a needle shape, often growing as hollow tubes (Figure 3.26).



Figure 3.26 Crystallization trials using microdialysis buttons. 37 mg/mL CYP450cam in a 50 μ L volume button. **(A)** 50 mM Tris pH = 7.4, 18% PEG4000, 250 mM KCl, 1 mM camphor **(B)** 50 mM Tris pH = 7.4, 16.8% PEG8000, 250 mM KCl, 1 mM camphor **(C)** 50 mM Tris pH = 7.4, 32% PEG6000, 500 mM KCl, 2 mM camphor.

3.4.4.4 Screens

To expand upon the conditions used to obtain crystals that had a needle-like morphology, various commercially available crystal screens from Qiagen were used. The crystals obtained continued to have needle or plate morphologies or formed hollow rectangular cylinders with very jagged edges (Figure 3.27).

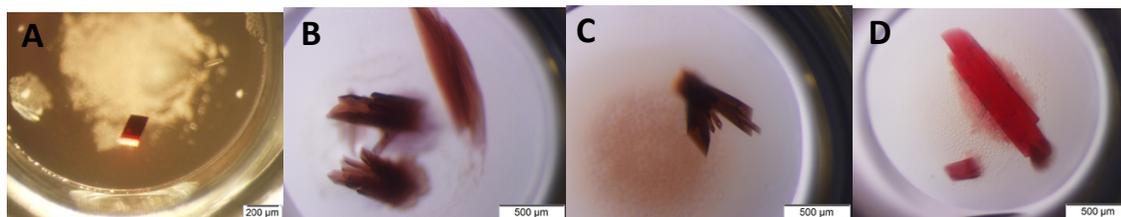


Figure 3.27 Crystallization screens in sitting drop vapor diffusion. 20 mg/mL CYP450cam in a 2 μ L drop. **(A)** 0.2 M Magnesium acetate, 20% PEG 3350, JCSG Core I condition 12 **(B)** 0.2 M Lithium sulfate, 0.1 M Tris pH = 8.5, 16% PEG4000, PEGS II Suite condition 31 **(C)** 0.1 M HEPES pH = 7.0, 30% PEG6000, JCSG Core IV condition 60 **(D)** 0.1 M Sodium acetate pH = 5.0, 2M Sodium formate, Protein Complex Suite condition 78.

One trial condition yielded what appeared to be a rhombus shape crystal as opposed to needles or plates. Larger sitting drop optimization trials of this condition were performed; however, they did not yield the same rhomboid shape.

3.4.4.5 Microseeding

Seeding pre-equilibrated drops with microseeds was also attempted. The crystals grown had a needle like morphology or formed small crystalline aggregates (Figure 3.28). Additionally, many drops remained clear and did not form any crystalline material.

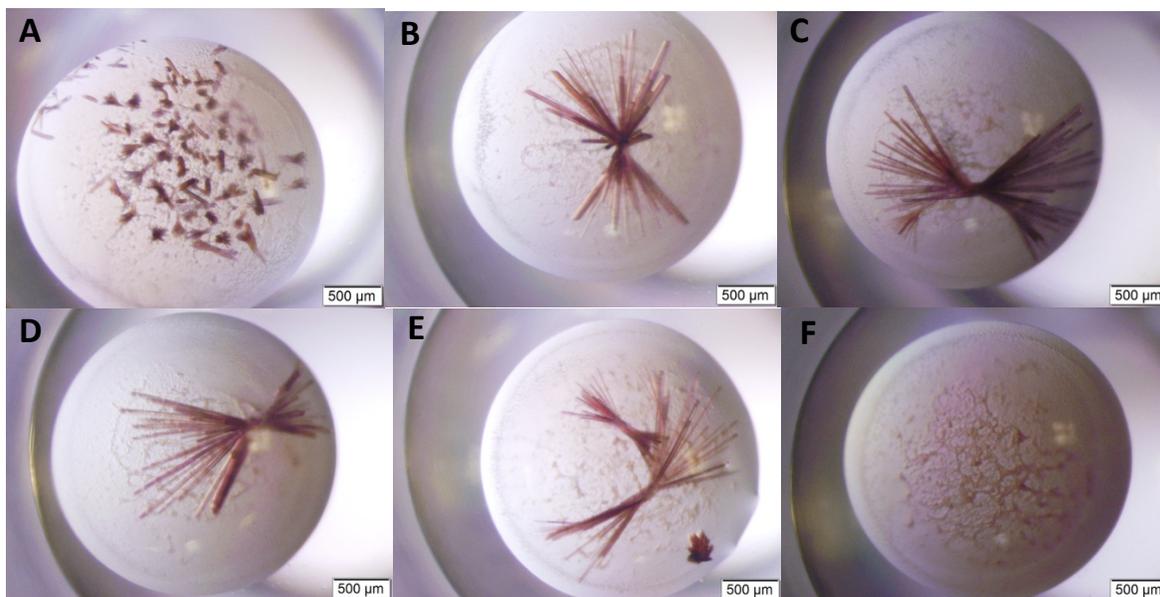


Figure 3.28 Crystallization screens in sitting drop vapor diffusion. 18 mg/mL CYP450cam using microseeding with 1 μ L seed, 1 μ L protein and 1 μ L reservoir. (A) 50 mM Tris pH = 7.4, 16% PEG6000, 250 mM KCl, 2 mM camphor, 10^{-4} dilution seed (B) 50 mM Tris pH = 7.4, 16% PEG8000, 250 mM KCl, 2 mM camphor, 10^{-4} dilution seed (C) 50 mM Tris pH = 7.4, 16.8% PEG6000, 250 mM KCl, 2 mM camphor, 10^{-3} dilution seed (D) 50 mM Tris pH = 7.4, 18% PEG6000, 250 mM KCl, 2 mM camphor, 10^{-4} dilution seed (E) 50 mM Tris pH = 7.4, 15.2% PEG8000, 250 mM KCl, 2 mM camphor, 10^{-4} dilution seed (F) 50 mM Tris pH = 7.4, 16% PEG6000, 250 mM KCl, 2 mM camphor, 10^{-2} dilution seed.

3.4.5 X-ray diffraction

While the needle and plate morphology of the crystals was not suitable for neutron protein diffraction, they were however suited for X-ray diffraction data collection. Test diffraction indicated that the needles were present in the orthorhombic $P2_12_12_1$ space group as observed by Meilleur *et al.* in initial X-ray diffraction studies. Crystals diffracted in the 2.0 – 2.2 Å resolution range.

A complete X-ray diffraction dataset was collected at room temperature. Room temperature X-ray diffraction has been shown to reveal conformational flexibility that are lost or modified by cryo-cooling.^{176–}

¹⁷⁹ This structural plasticity includes alternate conformations that may have significant functional and mechanistic implications. An example of this on the higher temperature extremes for proteins was a thermostable copper nitrite reductase that displayed monodentate binding of nitrite at cryogenic conditions, however this switched to bidentate binding in a structure collected at 320 K.¹⁸⁰ Conformational heterogeneity of a room temperature and cryogenic structure of the *Staphylococcus aureus* virulence factor staphylococcal protein A (SpA) revealed that the residues in helix1 conformational heterogeneity which facilitates in binding partner interactions.¹⁸¹

A room temperature X-ray diffraction dataset was collected of CYP450cam(C334A) at a resolution of 1.8 Å (Table 3.1). The crystal was in the P2₁ space group with the same unit cell as had been observed previously by Schlichting *et al.* in the ferric form of CYP450cam (PDB 1DZ4).⁴² Analysis of the structure confirmed correct heme incorporation in the active site with an occupancy of 1. Camphor was further visible in the in the active site with the C5 carbon 4.32 Å from the heme iron. Structural alignment of the room temperature structure with the cryogenic 1DZ4 had an RMSD of 0.20 Å for the main chain and were thus very similar. Detailed analysis of the catalytically important residues did not reveal significant conformational differences.

Table 3.1 Data collection and refinement statistics.

	P450cam
Wavelength (Å)	1.54
Resolution range (Å)	28.75 - 1.8 (1.864 - 1.8)
Space group (Å)	P 1 2 ₁ 1
Unit cell <i>a</i> , <i>b</i> , <i>c</i> (Å) α , β , γ (°)	67.78 62.77 95.76 90 90.67 90
Total reflections	490977 (42857)
Unique reflections	74684 (7382)
Multiplicity	6.6 (5.8)
Completeness (%)	99.92 (99.88)
$\langle I/\sigma(I) \rangle$	18.65 (4.00)
Wilson B-factor (Å ²)	17.75
R-merge (%)	0.1396 (0.3541)
R-meas (%)	0.1513 (0.3893)
R-pim (%)	0.05772 (0.1589)
CC _{1/2}	0.978 (0.892)
CC*	0.994 (0.971)
Reflections used in refinement	74658 (7381)
Reflections used for R-free	3707 (381)
R-work (%)	0.1399 (0.1890)
R-free (%)	0.1767 (0.2424)
CC _{work}	0.969 (0.921)
CC _{free}	0.974 (0.832)
Number of non-hydrogen atoms	7493
macromolecules	6800
ligands	111
solvent	582
Protein residues	809
RMS(bonds)	0.009
RMS(angles)	1.05
Ramachandran favored (%)	98.01
Ramachandran allowed (%)	1.99
Ramachandran outliers (%)	0
Rotamer outliers (%)	1.2
Clashscore	6.35
Average B-factor (Å ²)	22.09
macromolecules	21.14
ligands	11.21
solvent	35.3

Statistics for the highest-resolution shell are shown in parentheses.

Analysis of the active site indicated a well-ordered camphor bound to the active site with the carbonyl group forming a hydrogen bond with Tyr96 and the C5 carbon 4.32 Å from the heme iron (Figure 3.29). The CYP450cam is in the substrate-bound, closed state with the I-helix Thr252 hydroxyl group pointing away from the active site heme. Asp251 is in a conformation characteristic of the closed state, locked in a salt bridge with Lys178 and Arg186 (Figure 3.30). The I-helix is closed with Thr252 forming a distant hydrogen bond with only a single water molecule.

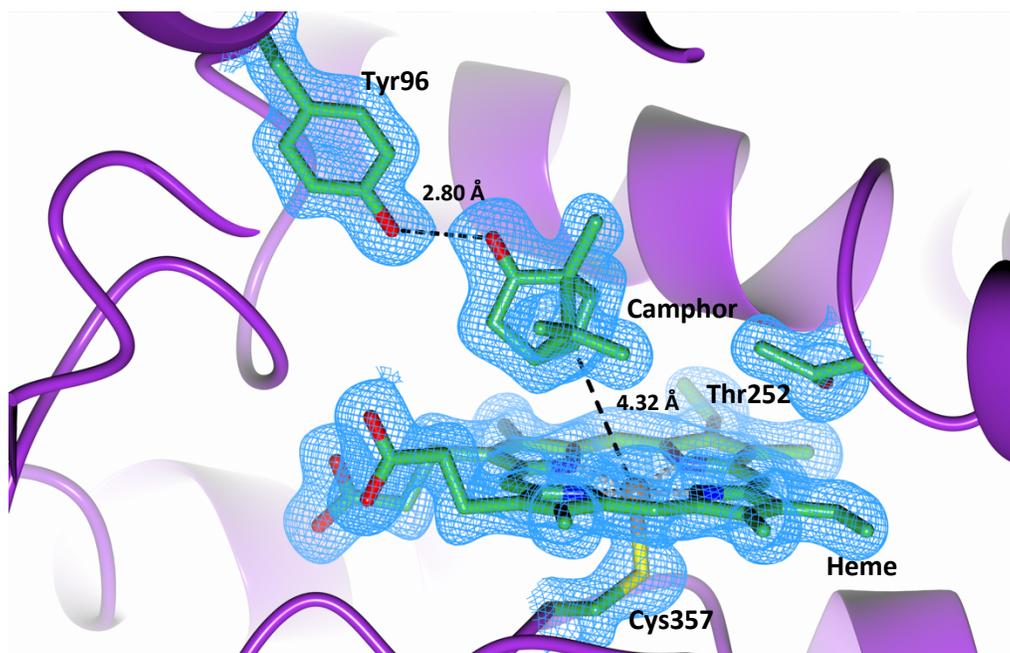


Figure 3.29 Active site of CYP450cam with camphor bound. $2F_o - F_c$ maps displayed in blue at $\sigma = 1.00$.

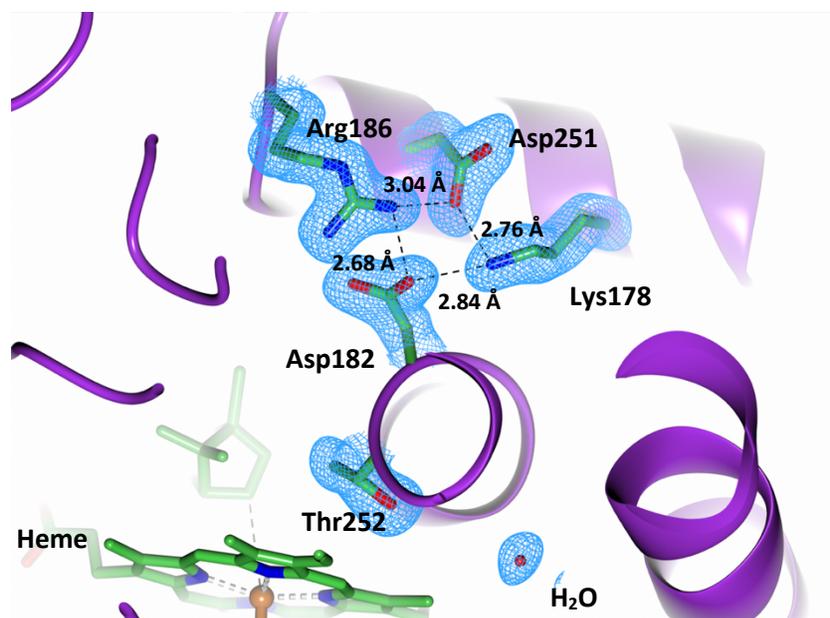


Figure 3.30 CYP450cam Asp251 salt bridge interactions. $2F_o-F_c$ maps displayed in blue at $\sigma = 1.00$.

The B'-helix is in an ordered state with a K^+ ion coordinated to the carbonyl oxygen atoms of Glu84, Gly93 and Glu94 as well as two water molecules to form a hexacoordinated system (Figure 3.31). This is characteristic of the closed state, since the B'-helix is disordered in the open state. $2F_o-F_c$ maps displayed in blue at $\sigma = 1.00$.

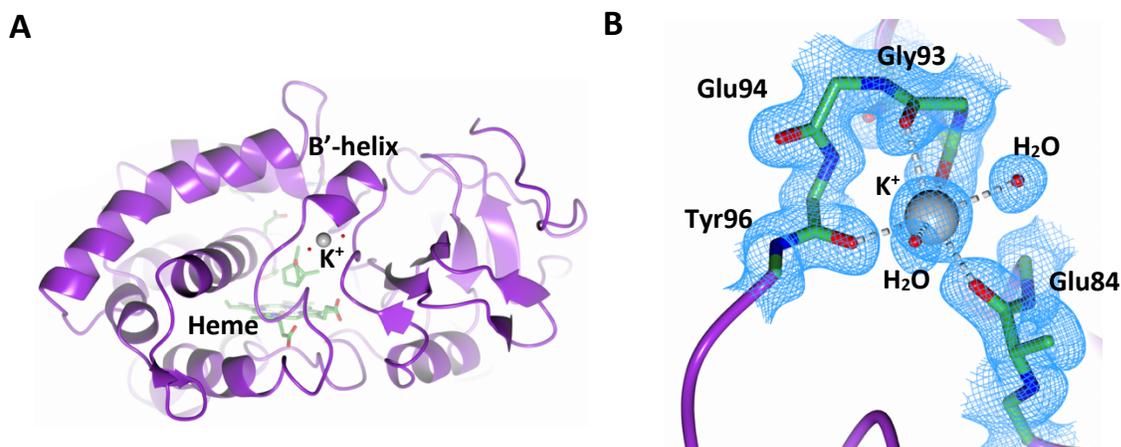


Figure 3.31 Interactions of the B'-helix in CYP450cam. **(A)** The B'-helix is ordered interacting with K^+ in the closed form. **(B)** K^+ interacts with the carbonyl oxygen atoms of Glu84, Gly93, Glu94 and Tyr96 and coordinates two water molecules. $2F_o-F_c$ maps displayed in blue at $\sigma = 1.00$.

3.5 Conclusions

The studies presented here have permitted a robust expression and purification protocol of CYP450cam(C334A)-His₆ or CYP450cam(C334A) to be established. In order to obtain a pure CYP450cam sample with sufficient heme incorporation, CYP450cam can be expressed with a His₆ tag and purified using Ni-affinity chromatography or without a tag using a strong anion exchange Q-column. This is followed by two consecutive rounds of SEC to ensure a sample with sufficient purity and heme incorporation levels (R_z of 1.45). The most effective way to ensure that any apo CYP450cam missing the heme cofactor is reconstituted is by addition of hemin chloride to the lysis buffer prior to purification. This ensures that excess unbound hemin chloride does not remain, interfering with sample purity, while also ensuring that heme is incorporated in the active site.

X-ray diffraction data collected on the crystals grown from purified protein confirmed heme incorporation at the active site and camphor substrate binding. While the majority of crystals grown were in the P₂₁2₁2₁ space group, crystals were also grown in the P₂₁ space group. The crystallization studies performed here resulted in CYP450cam crystals with a needle or plate morphology. These crystals are not suitable for neutron protein diffraction data collection. Further optimization and use of varied crystallization techniques will be necessary to obtain crystals of a sufficient volume for neutron studies.

3.6 References

1. Nelson, D. R. *et al.* CYP450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* **6**, 1–42 (1996).
2. Klingenberg, M. Pigments of rat liver microsomes. *Arch. Biochem. Biophys.* **75**, 376–386 (1958).
3. Estabrook, R. W. A passion for CYP450s (remembrances of the early history of research on cytochrome P450). *Drug Metab. Dispos.* **31**, 1461–1473 (2003).
4. Ortiz de Montellano, P. R. *Cytochrome P450: Structure, Mechanism and Biochemistry*. (Springer US, 2005). doi:10.1007/b139087
5. Robin, A. *et al.* Chimeric self-sufficient CYP450cam-RhFRed biocatalysts with broad substrate scope. *Beilstein J. Org. Chem.* **7**, 1494–1498 (2011).
6. Nelson, D. R. Cytochrome P450 diversity in the tree of life. *Biochim. Biophys. Acta - Proteins Proteomics* **1866**, 141–154 (2018).
7. Urlacher, V. B. & Girhard, M. Cytochrome P450 monooxygenases: An update on perspectives for synthetic application. *Trends Biotechnol.* **30**, 26–36 (2012).
8. Kelly, S. L. & Kelly, D. E. Microbial cytochromes CYP450: Biodiversity and biotechnology. Where do cytochromes CYP450 come from, what do they do and what can they do for US? *Philos. Trans. R. Soc. B Biol. Sci.* **368**, (2013).
9. Finnigan, J. D., Young, C., Cook, D. J., Charnock, S. J. & Black, G. W. Cytochromes CYP450 (P450s): A review of the class system with a focus on prokaryotic CYP450s. *Adv. Protein Chem. Struct. Biol.* **122**, 289–320 (2020).
10. Lamb, D. C. & Waterman, M. R. Unusual properties of the cytochrome P450 superfamily. *Philos. Trans. R. Soc. B Biol. Sci.* **368**, (2013).
11. Denisov, I. G., Shih, A. Y. & Sligar, S. G. Structural differences between soluble and membrane bound cytochrome P450s. *J. Inorg. Biochem.* **108**, 150–158 (2012).
12. Poulos, T. L., Finzel, B. C. & Gunsalus, I. C. The 2.6-Å crystal structure of *Pseudomonas putida* cytochrome P-450. *J. Biol. Chem.* **260**, 16122–16130 (1985).
13. Lamb, D. C. *et al.* The First Virally Encoded Cytochrome P450. *J. Virol.* **83**, 8266–8269 (2009).
14. Lamb, D. C. *et al.* On the occurrence of cytochrome P450 in viruses. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 12343–12352 (2019).
15. Ichinose, H. Molecular and functional diversity of fungal cytochrome P450s. *Biol. Pharm. Bull.* **35**, 833–837 (2012).
16. Chen, W. *et al.* Fungal cytochrome P450 monooxygenases: Their distribution, structure, functions, family expansion, and evolutionary origin. *Genome Biol. Evol.* **6**, 1620–1634 (2014).
17. Hamberger, B. & Bak, S. Plant CYP450s as versatile drivers for evolution of species-specific chemical diversity. *Philos. Trans. R. Soc. B Biol. Sci.* **368**, (2013).
18. Werck-Reichhart, D. & Feyereisen, R. Cytochromes CYP450 A Success Story. *Genome Biol.* **1**, 1–9 (2000).
19. Nebert, D. W. & Russell, D. W. Clinical importance of the cytochromes CYP450. *Lancet* **360**, 1155–1162 (2002).
20. Nebert, D. W. & Dalton, T. P. The role of cytochrome P450 enzymes in endogenous signalling pathways and environmental carcinogenesis. *Nat. Rev. Cancer* **6**, 947–960 (2006).
21. Williams, P. A., Cosme, J., Sridhar, V., Johnson, E. F. & McRee, D. E. Mammalian microsomal cytochrome P450 monooxygenase: Structural adaptations for membrane binding and functional diversity. *Mol. Cell* **5**, 121–131 (2000).
22. Guengerich, F. P., Waterman, M. R. & Egli, M. Recent Structural Insights into Cytochrome P450 Function. *Trends Pharmacol. Sci.* **37**, 625–640 (2016).
23. Sim, S. C. & Ingelman-Sundberg, M. The human cytochrome P450 (CYP) allele nomenclature website: A peer-reviewed database of CYP variants and their associated effects. *Hum. Genomics* **4**, 278–281 (2010).
24. Nebert, D. W. & Gonzalez, F. J. CYP450 Genes: Structure, Evolution, and Regulation. *Annu. Rev.*

- Biochem.* **56**, 945–993 (1987).
25. Nelson, D. R. Cytochrome P450 Nomenclature, 2004. in *Cytochrome P450 Protocols* **320**, 1–10 (Humana Press, 2004).
 26. Gricman, L., Vogel, C. & Pleiss, J. Conservation analysis of class-specific positions in cytochrome P450 monooxygenases: Functional and structural relevance. *Proteins Struct. Funct. Bioinforma.* **82**, 491–504 (2014).
 27. Moody, S. C. & Loveridge, E. J. CYP105-diverse structures, functions and roles in an intriguing family of enzymes in *Streptomyces*. *J. Appl. Microbiol.* **117**, 1549–1563 (2014).
 28. Nelson, D. R. & Nebert, D. W. Cytochrome P450 (CYP) Gene Superfamily. *eLS* **450**, 1–19 (2018).
 29. Ortiz de Montellano, P. R. *Cytochrome P450*. Springer (Springer International Publishing, 2015). doi:10.1007/978-3-319-12108-6
 30. Hannemann, F., Bichet, A., Ewen, K. M. & Bernhardt, R. Cytochrome P450 systems—biological variations of electron transport chains. *Biochim. Biophys. Acta - Gen. Subj.* **1770**, 330–344 (2007).
 31. Volokhan, O., Sletta, H., Ellingsen, T. E. & Zotchev, S. B. Characterization of the CYP450 monooxygenase NysL, responsible for C-10 hydroxylation during biosynthesis of the polyene macrolide antibiotic nystatin in *Streptomyces noursei*. *Appl. Environ. Microbiol.* **72**, 2514–2519 (2006).
 32. Chien, Y., Rosal, K. & Chung, B. chu. Function of CYP11A1 in the mitochondria. *Mol. Cell. Endocrinol.* **441**, 55–61 (2017).
 33. Correddu, D., Di Nardo, G. & Gilardi, G. Self-Sufficient Class VII Cytochromes CYP450: From Full-Length Structure to Synthetic Biology Applications. *Trends Biotechnol.* (2021). doi:10.1016/j.tibtech.2021.01.011
 34. Urlacher, V. B. & Girhard, M. Cytochrome P450 monooxygenases: An update on perspectives for synthetic application. *Trends Biotechnol.* **30**, 26–36 (2012).
 35. Hasemann, C. A., Kurumbail, R. G., Boddupalli, S. S., Peterson, J. A. & Deisenhofer, J. Structure and function of cytochromes CYP450:a comparative analysis of three crystal structures. *Structure* **3**, 41–62 (1995).
 36. Peterson, J. A. & Graham, S. E. A close family resemblance: The importance of structure in understanding cytochromes CYP450. *Structure* **6**, 1079–1085 (1998).
 37. Midlik, A. *et al.* Uncovering of cytochrome P450 anatomy by SecStrAnnotator. *Sci. Rep.* **11**, 1–12 (2021).
 38. Presnell, S. R. & Cohen, F. E. Topological distribution of four-alpha-helix bundles. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 6592–6596 (1989).
 39. Dawson, J. H. *et al.* Oxidized cytochrome P-450. Magnetic circular dichroism evidence for thiolate ligation in the substrate-bound form. Implications for the catalytic mechanism. *J. Am. Chem. Soc.* **98**, 3707–3709 (1976).
 40. Poulos, T. L. Cytochrome P450. *Curr. Opin. Struct. Biol.* **5**, 767–774 (1995).
 41. Nagano, S. & Poulos, T. L. Crystallographic study on the dioxygen complex of wild-type and mutant cytochrome P450cam: Implications for the dioxygen activation mechanism. *J. Biol. Chem.* **280**, 31659–31663 (2005).
 42. Schlichting, I. *et al.* The catalytic pathway of cytochrome P450cam at atomic resolution. *Science* **287**, 1615–22 (2000).
 43. Gotoh, O. Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *J. Biol. Chem.* **267**, 83–90 (1992).
 44. Poulos, T. L. & Johnson, E. F. Structures of Cytochrome P450 Enzymes. in *Cytochrome P450* 3–32 (Springer International Publishing, 2015). doi:10.1007/978-3-319-12108-6_1
 45. Denisov, I. G., Makris, T. M., Sligar, S. G. & Schlichting, I. Structure and chemistry of cytochrome P450. *Chem. Rev.* **105**, 2253–2277 (2005).
 46. Bernhardt, R. Cytochromes CYP450 as versatile biocatalysts. *J. Biotechnol.* **124**, 128–45 (2006).
 47. Harris, D. & Loew, G. Determinants of the Spin State of the Resting State of Cytochrome P450cam.

- J. Am. Chem. Soc.* **115**, 8775–8779 (1993).
48. Fisher, M. T. & Sligar, S. G. Control of heme protein redox potential and reduction rate: linear free energy relation between potential and ferric spin state equilibrium. *J. Am. Chem. Soc.* **107**, 5018–5019 (1985).
 49. Johnston, W. A. *et al.* Cytochrome P450 is present in both ferrous and ferric forms in the resting state within intact *Escherichia coli* and hepatocytes. *J. Biol. Chem.* **286**, 40750–40759 (2011).
 50. Meunier, B. & Bernadou, J. Active Iron-Oxo and Iron-Peroxo Species in Cytochromes CYP450 and Peroxidases; Oxo-Hydroxo Tautomerism with Water-Soluble Metalloporphyrins. in *Metal-Oxo and Metal-Peroxo Species in Catalytic Oxidations* **148**, 1–35 (Springer Berlin Heidelberg).
 51. Kovacs, J. A. Synthetic Analogues of Cysteinate-Ligated Non-Heme Iron and Non-Corrinoid Cobalt Enzymes. *Chem. Rev.* **104**, 825–848 (2004).
 52. Surawatanawong, P., Tye, J. W. & Hall, M. B. Density functional theory applied to a difference in pathways taken by the enzymes cytochrome P450 and superoxide reductase: Spin states of ferric hydroperoxo intermediates and hydrogen bonds from water. *Inorg. Chem.* **49**, 188–198 (2010).
 53. Rittle, J. & Green, M. T. Cytochrome P450 Compound I: Capture, Ch. *Science*. **330**, 933–937 (2010).
 54. Newcomb, M. *et al.* Cytochrome P450 compound I. *J. Am. Chem. Soc.* **128**, 4580–4581 (2006).
 55. Rittle, J. & Green, M. T. Cytochrome P450 Compound I: Capture, Characterization, and C-H Bond Activation Kinetics. *Science*. **330**, 933–937 (2010).
 56. Woggon, W.-D. Cytochrome P450: Significance, reaction mechanisms and active site analogues. in *Bioorganic Chemistry* 39–96 (Springer, 1996). doi:10.1007/3-540-61388-9_62
 57. Modi, A., Dawson, J. H., Hrycay, E. G. & Bandiera, S. M. *Peroxidase and Peroxygenase Properties and Mechanisms of Cytochrome P450. Monooxygenase, Peroxidase and Peroxygenase Properties and Mechanisms of Cytochrome P450* (Springer, 2015).
 58. Prasad, S. & Mitra, S. Substrate modulates compound I formation in peroxide shunt pathway of *Pseudomonas putida* cytochrome P450cam. *Biochem. Biophys. Res. Commun.* **314**, 610–614 (2004).
 59. Behrendorff, J. B. Y. H. Reductive Cytochrome P450 Reactions and Their Potential Role in Bioremediation. *Front. Microbiol.* **12**, 1–14 (2021).
 60. Guengerich, F. P. Rate-limiting steps in cytochrome P450 catalysis. *Biol. Chem.* **383**, 1553–64 (2002).
 61. Vermilion, J. L., Ballou, D. P., Massey, V. & Coon, M. J. Separate roles for FMN and FAD in catalysis by liver microsomal NADPH-cytochrome P-450 reductase. *J. Biol. Chem.* **256**, 266–277 (1981).
 62. Oprian, D. D. & Coon, M. J. Oxidation-reduction states of FMN and FAD in NADPH-cytochrome P-450 reductase during reduction by NADPH. *J. Biol. Chem.* **257**, 8935–8944 (1982).
 63. Fantuzzi, A., Fairhead, M. & Gilardi, G. Direct Electrochemistry of Immobilized Human Cytochrome P450 2E1. *J. Am. Chem. Soc.* **126**, 5040–5041 (2004).
 64. Li, S. & Wackett, L. P. Reductive Dehalogenation by Cytochrome P450CAM: Substrate Binding and Catalysis. *Biochemistry* **32**, 9355–9361 (1993).
 65. Park, J. H. *et al.* Cofactor-free light-driven whole-cell cytochrome P450 catalysis. *Angew. Chemie - Int. Ed.* **54**, 969–973 (2015).
 66. Erman, J. E. *et al.* Histidine 52 is a critical residue for rapid formation of cytochrome c peroxidase compound I. *Biochemistry* **32**, 9798–9806 (1993).
 67. Gunsalus, I. C. & Sligar, S. G. Equilibrium states and dynamic reactions of iron in the camphor monooxygenase system. *Adv. Exp. Med. Biol.* **74**, 254–62 (1976).
 68. Poulos, T. L., Finzel, B. C. & Howard, A. J. High-resolution crystal structure of cytochrome P450cam. *J. Mol. Biol.* **195**, 687–700 (1987).
 69. Gerber, N. C. & Sligar, S. G. Catalytic mechanism of cytochrome P-450: evidence for a distal charge relay. *J. Am. Chem. Soc.* **114**, 8742–8743 (1992).
 70. Gerber, N. C. & Sligar, S. G. A role for Asp-251 in cytochrome P-450cam oxygen activation. *J. Biol. Chem.* **269**, 4260–6 (1994).

71. Imai, M. *et al.* Uncoupling of the cytochrome P-450cam monooxygenase reaction by a single mutation, threonine-252 to alanine or valine: A possible role of the hydroxy amino acid in oxygen activation. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 7823–7827 (1989).
72. Martinis, S. A., Atkins, W. M., Stayton, P. S. & Sligar, S. G. A Conserved Residue of Cytochrome P-450 Is Involved The Dehydrophenyl Anion and the Gas Phase Ion Chemistry of Benzyne. *J. Am. Chem. Soc.* **111**, 9252–9253 (1989).
73. Vidakovic, M., Sligar, S. G., Li, H. & Poulos, T. L. Understanding the role of the essential Asp251 in cytochrome P450cam using site-directed mutagenesis, crystallography, and kinetic solvent isotope effect. *Biochemistry* **37**, 9211–9 (1998).
74. Kimata, Y., Shimada, H., Hirose, T. & Ishimura, Y. Role of THR-252 in cytochrome P450CAM: A study with unnatural amino acid mutagenesis. *Biochemical and Biophysical Research Communications* **208**, 96–102 (1995).
75. Hishiki, T. *et al.* X-ray crystal structure and catalytic properties of Thr252Ile mutant of cytochrome P450cam: roles of Thr252 and water in the active center. *J. Biochem.* **128**, 965–74 (2000).
76. Zheng, J., Wang, D., Thiel, W. & Shaik, S. QM/MM study of mechanisms for compound I formation in the catalytic cycle of cytochrome P450cam. *J. Am. Chem. Soc.* **128**, 13204–13215 (2006).
77. Wang, D., Zheng, J., Shaik, S. & Thiel, W. Quantum and molecular mechanical study of the first proton transfer in the catalytic cycle of cytochrome P450cam and its mutant D251N. *J. Phys. Chem. B* **112**, 5126–5138 (2008).
78. Kalita, S., Shaik, S., Kisan, H. K. & Dubey, K. D. A Paradigm Shift in the Catalytic Cycle of CYP450: The Preparatory Choreography during O₂ Binding and Origins of the Necessity for Two Protonation Pathways. *ACS Catal.* **10**, 11481–11492 (2020).
79. Coleman, T. *et al.* Structural insights into the role of the acid-alcohol pair of residues required for dioxygen activation in cytochrome P450 enzymes. *J. Biol. Inorg. Chem.* **25**, 583–596 (2020).
80. Lüdemann, S. K., Carugo, O. & Wade, R. C. Substrate access to cytochrome P450cam: A comparison of a thermal motion pathway analysis with molecular dynamics simulation data. *J. Mol. Model.* **3**, 369–374 (1997).
81. Cojocaru, V., Winn, P. J. & Wade, R. C. The ins and outs of cytochrome P450s. *Biochim. Biophys. Acta - Gen. Subj.* **1770**, 390–401 (2007).
82. Sligar, S. G. Coupling of Spin, Substrate, and Redox Equilibriums in Cytochrome P450. *Biochemistry* **15**, 5399–5406 (1976).
83. Reipa, V., Holden, M. J., Mayhew, M. P. & Vilker, V. L. Temperature dependence of the formal reduction potential of putidaredoxin. *Biochim. Biophys. Acta - Bioenerg.* **1459**, 1–9 (2000).
84. Pochapsky, T. C. & Pochapsky, S. S. What Your Crystal Structure Will Not Tell You about Enzyme Function. *Acc. Chem. Res.* (2019). doi:10.1021/acs.accounts.9b00066
85. Gunsalus, I. C. & Sligar, S. G. Redox regulation of cytochrome P450cam mixed function oxidation by putidaredoxin and camphor ligation. *Biochimie* **58**, 143–147 (1976).
86. Tosha, T., Yoshioka, S., Ishimori, K. & Morishima, I. L358P mutation on cytochrome P450cam simulates structural changes upon putidaredoxin binding: The structural changes trigger electron transfer to oxy-P450cam from electron donors. *J. Biol. Chem.* **279**, 42836–42843 (2004).
87. Nagano, S., Tosha, T., Ishimori, K., Morishima, I. & Poulos, T. L. Crystal structure of the cytochrome P450cam mutant that exhibits the same spectral perturbations induced by putidaredoxin binding. *J. Biol. Chem.* **279**, 42844–42849 (2004).
88. Myers, W. K., Lee, Y.-T., Britt, R. D. & Goodin, D. B. The Conformation of CYP450cam in Complex with Putidaredoxin Is Dependent on Oxidation State. *J. Am. Chem. Soc.* **135**, 11732–11735 (2013).
89. Liou, S. H., Mahomed, M., Lee, Y. T. & Goodin, D. B. Effector Roles of Putidaredoxin on Cytochrome P450cam Conformational States. *J. Am. Chem. Soc.* **138**, 10163–10172 (2016).
90. Sevrioukova, I. F., Garcia, C., Li, H., Bhaskar, B. & Poulos, T. L. Crystal structure of putidaredoxin, the [2Fe-2S] component of the CYP450cam monooxygenase system from *Pseudomonas putida*. *J. Mol. Biol.* **333**, 377–92 (2003).

91. Lipscomb, J. D., Sligar, S. G., Namtvedt, M. J. & Gunsalus, I. C. Autooxidation and hydroxylation reactions of oxygenated cytochrome P-450cam. *J. Biol. Chem.* **251**, 1116–24 (1976).
92. Tripathi, S., Li, H. & Poulos, T. L. Structural basis for effector control and redox partner recognition in cytochrome P450. *Science*. **340**, 1227–1230 (2013).
93. Sears, V. F. Neutron scattering lengths and cross sections. *Neutron News* **3**, 26–37 (1992).
94. Blakeley, M. P., Langan, P., Niimura, N. & Podjarny, A. Neutron crystallography: opportunities, challenges, and limitations. *Curr. Opin. Struct. Biol.* **18**, 593–600 (2008).
95. Oksanen, E., Chen, J. C. H. & Fisher, S. Z. Neutron crystallography for the study of hydrogen bonds in macromolecules. *Molecules* **22**, 1–26 (2017).
96. Schröder, G. C., O'Dell, W. B., Myles, D. A. A., Kovalevsky, A. & Meilleur, F. IMAGINE: Neutrons reveal enzyme chemistry. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 778–786 (2018).
97. Podjarny, A. *et al.* X-ray crystallography at subatomic resolution. *Europhys. News* **33**, 113–117 (2002).
98. Schröder, G. C. & Meilleur, F. Metalloprotein catalysis: structural and mechanistic insights into oxidoreductases from neutron protein crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **D77**, (2021).
99. Kwon, H. *et al.* *Heme peroxidase—Trapping intermediates by cryo neutron crystallography. Methods in Enzymology* **634**, (Elsevier Inc., 2020).
100. Casadei, C. M. *et al.* Neutron cryo-crystallography captures the protonation state of ferryl heme in a peroxidase. *Science*. **345**, 193–197 (2014).
101. Kwon, H. *et al.* Direct visualization of a Fe(IV)-OH intermediate in a heme enzyme. *Nat. Commun.* **7**, 1–6 (2016).
102. Kwon, H. *et al.* Visualizing the protons in a metalloenzyme electron proton transfer pathway. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 6484–6490 (2020).
103. O'Dell, W. B., Bodenheimer, A. M. & Meilleur, F. Neutron protein crystallography: A complementary tool for locating hydrogens in proteins. *Arch. Biochem. Biophys.* **602**, 48–60 (2016).
104. Schröder, G. C. & Meilleur, F. Neutron Crystallography Data Collection and Processing for Modelling Hydrogen Atoms in Protein Structures. *J. Vis. Exp.* 1–38 (2020). doi:10.3791/61903
105. Brockwell, D. Physicochemical consequences of the perdeuteration of glutathione S-transferase from *S. japonicum*. *Protein Sci.* **10**, 572–580 (2001).
106. Meilleur, F., Contzen, J., Myles, D. A. A. & Jung, C. Structural stability and dynamics of hydrogenated and perdeuterated cytochrome P450cam (CYP101). *Biochemistry* **43**, 8744–8753 (2004).
107. Meilleur, F., Dauvergne, M.-T., Schlichting, I. & Myles, D. A. A. Production and X-ray crystallographic analysis of fully deuterated cytochrome P450cam. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **61**, 539–544 (2005).
108. Blakeley, M. P., Hasnain, S. S. & Antonyuk, S. V. Sub-atomic resolution X-ray crystallography and neutron crystallography: Promise, challenges and potential. *IUCrJ* **2**, 464–474 (2015).
109. Niimura, N. Neutrons expand the field of structural biology. *Curr. Opin. Struct. Biol.* **9**, 602–608 (1999).
110. Ng, J. D. *et al.* Large-volume protein crystal growth for neutron macromolecular crystallography. *Acta Crystallogr. Sect. F Structural Biol. Commun.* **71**, 358–370 (2015).
111. Kelpšas, V. *et al.* Perdeuteration, large crystal growth and neutron data collection of *Leishmania mexicana* triose-phosphate isomerase E65Q variant. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **75**, 260–269 (2019).
112. Tanaka, I. Necessity and Effectiveness of a Biological Diffractometer for Installation in Second Target Station as a New Neutron Pulsed Source in Japan. in *Proceedings of the 3rd International Symposium of Quantum Beam Science at Ibaraki University 'Quantum Beam Science in Biology and Soft Materials (ISQBSS2018)'* **011015**, 1–5 (Journal of the Physical Society of Japan, 2019).
113. Budayova-Spano, M., Koruza, K. & Fisher, Z. *Large crystal growth for neutron protein crystallography. Methods in Enzymology* **634**, (Elsevier Inc., 2020).

114. Snell, E. H. *et al.* Optimizing crystal volume for neutron diffraction: D-xylose isomerase. *Eur. Biophys. J.* **35**, 621–632 (2006).
115. Thaller, C. *et al.* [9] Seed enlargement and repeated seeding. *Methods Enzymol.* **114**, 132–135 (1985).
116. O'Dell, W. B., Agarwal, P. K. & Meilleur, F. Oxygen Activation at the Active Site of a Fungal Lytic Polysaccharide Monooxygenase. *Angew. Chemie - Int. Ed.* **56**, 767–770 (2017).
117. McPherson, A. & Gavira, J. A. Introduction to protein crystallization. *Acta Crystallogr. Sect. FStructural Biol. Commun.* **70**, 2–20 (2014).
118. Candoni, N., Grossier, R., Hammadi, Z., Morin, R. & Veessler, S. Practical Physics Behind Growing Crystals of Biological Macromolecules. *Protein Pept. Lett.* **19**, 714–724 (2012).
119. Feher, G. Mechanisms of nucleation and growth of protein crystals. *J. Cryst. Growth* **76**, 545–546 (1986).
120. Majeed, S. *et al.* Enhancing protein crystallization through precipitant synergy. *Structure* **11**, 1061–1070 (2003).
121. Nanev, C. N. Advancements (and challenges) in the study of protein crystal nucleation and growth; thermodynamic and kinetic explanations and comparison with small-molecule crystallization. *Prog. Cryst. Growth Charact. Mater.* **66**, 100484 (2020).
122. Rupp, B. Origin and use of crystallization phase diagrams. *Acta Crystallogr. Sect. FStructural Biol. Commun.* **71**, 247–260 (2015).
123. Haas, C. & Drenth, J. Understanding protein crystallization on the basis of the phase diagram. *J. Cryst. Growth* **196**, 388–394 (1999).
124. Asherie, N. Protein crystallization and phase diagrams. *Methods* **34**, 266–272 (2004).
125. Kuznetsov, Y. G., Malkin, A. J. & McPherson, A. The liquid protein phase in crystallization: A case study - Intact immunoglobulins. *J. Cryst. Growth* **232**, 30–39 (2001).
126. Krauss, I. R., Merlino, A., Vergara, A. & Sica, F. An overview of biological macromolecule crystallization. *Int. J. Mol. Sci.* **14**, 11643–11691 (2013).
127. Weber, P. C. A protein crystallization strategy using automated grid searches on successively finer grids. *Methods* **1**, 31–37 (1990).
128. DeLucas, L. J. *et al.* Efficient protein crystallization. *J. Struct. Biol.* **142**, 188–206 (2003).
129. Govada, L. & Chayen, N. E. Choosing the method of crystallization to obtain optimal results. *Crystals* **9**, (2019).
130. Benvenuti, M. & Mangani, S. Crystallization of soluble proteins in vapor diffusion for x-ray crystallography. *Nat. Protoc.* **2**, 1633–1651 (2007).
131. Zeppezauer, M., Eklund, H. & Zeppezauer, E. S. Micro diffusion cells for the growth of single protein crystals by means of equilibrium dialysis. *Arch. Biochem. Biophys.* **126**, 564–573 (1968).
132. Ducruix, A. (Arnaud) & Giegé, R. (Richard). *Crystallization of nucleic acids and proteins : a practical approach.* (Oxford ; New York : Oxford University Press, [1999], 1999).
133. Chayen, N. E. Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54**, 8–15 (1998).
134. Chayen, N. E., Shaw Stewart, P. D., Maeder, D. L. & Blow, D. M. An automated system for microbatch protein crystallization and screening. *J. Appl. Crystallogr.* **23**, 297–302 (1990).
135. Salemme, F. R. A free interface diffusion technique for the crystallization of proteins for X-ray crystallography. *Arch. Biochem. Biophys.* **151**, 533–539 (1972).
136. Ng, J. D., Gavira, J. A. & García-Ruiz, J. M. Protein crystallization by capillary counterdiffusion for applied crystallographic structure determination. *J. Struct. Biol.* **142**, 218–231 (2003).
137. Vahdatahar, E., Junius, N. & Budayova-Spano, M. Optimization of crystal growth for neutron macromolecular crystallography. *J. Vis. Exp.* **2021**, 1–24 (2021).
138. McPherson, A. & Cudney, B. Optimization of crystallization conditions for biological macromolecules. *Acta Crystallogr. Sect. FStructural Biol. Commun.* **70**, 1445–1467 (2014).
139. Jones, H. G. *et al.* Iterative screen optimization maximizes the efficiency of macromolecular crystallization. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **75**, 123–131 (2019).

140. Bergfors, T. Seeds to crystals. *J. Struct. Biol.* **142**, 66–76 (2003).
141. Hondoh, H. & Nakada, T. Impurity effect on defect formation of protein crystals. *J. Cryst. Growth* **275**, 1423–1429 (2005).
142. Shaw Stewart, P. D., Kolek, S. A., Briggs, R. A., Chayen, N. E. & Baldock, P. F. M. Random microseeding: A theoretical and practical exploration of seed stability and seeding techniques for successful protein crystallization. *Cryst. Growth Des.* **11**, 3432–3441 (2011).
143. Stura, E. A. & Wilson, I. A. Applications of the streak seeding technique in protein crystallization. *J. Cryst. Growth* **110**, 270–282 (1991).
144. Meilleur, F., Contzen, J., Myles, D. A. A. & Jung, C. Structural Stability and Dynamics of Hydrogenated and Perdeuterated Cytochrome P450cam (CYP101). *Biochemistry* **43**, 8744–8753 (2004).
145. Nickerson, D. P. & Wong, L. L. The dimerization of *Pseudomonas putida* cytochrome P450(cam): Practical consequences and engineering of a monomeric enzyme. *Protein Eng.* **10**, 1357–1361 (1997).
146. Auclair, K., Moënné-Loccoz, P. & Ortiz de Montellano, P. R. Roles of the proximal heme thiolate ligand in cytochrome P450cam. *J. Am. Chem. Soc.* **123**, 4877–4885 (2001).
147. Lee, Y. T., Wilson, R. F., Rupniewski, I. & Goodin, D. B. CYP450cam visits an open conformation in the absence of substrate. *Biochemistry* **49**, 3412–3419 (2010).
148. Laemmli, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **227**, 680–685 (1970).
149. Wittig, I., Braun, H. P. & Schägger, H. Blue native PAGE. *Nat. Protoc.* **1**, 418–428 (2006).
150. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
151. Winn, M. D. *et al.* Overview of the CCP 4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 235–242 (2011).
152. Adams, P. D. *et al.* PHENIX : a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010).
153. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
154. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 352–367 (2012).
155. Verderber, E. *et al.* Role of the hemA gene product and δ -aminolevulinic acid in regulation of *Escherichia coli* heme synthesis. *J. Bacteriol.* **179**, 4583–4590 (1997).
156. Soret, J. L. Analyse Spectrale: Sur le Spectre d’Absorption du Sang dans la Partie Violette et Ultra-Violette. *C. R. Acad. Sci.* **97**, 1269–1270 (1883).
157. Hanson, L. K. *et al.* Origin of the Anomalous Soret Spectra of Carboxycytochrome P-450. *Journal of the American Chemical Society* **98**, 2672–2674 (1976).
158. Jung, C. Quantum chemical explanation of the ‘hyper’ spectrum of the carbon monoxide complex of cytochrome P-450. *Chem. Phys. Lett.* **113**, 589–596 (1985).
159. Hodgson, E. 7 - Microsomal Mono-Oxygenases. in *Pharmacology* (eds. KERKUT, G. A. & GILBERT, L. I.) 225–321 (Pergamon, 1985). doi:10.1016/B978-0-08-030812-8.50012-5
160. Priestersbach, A., Kubicek, J., Schäfer, F., Block, H. & Maertens, B. Purification of His-Tagged Proteins. *Methods Enzymol.* **559**, 1–15 (2015).
161. Porath, J. & Flodin, P. Gel Filtration: A Method for Desalting and Group Separation. *Nature* **183**, 1657–1659 (1959).
162. Gunsalus, I. C. & Wagner, G. C. Bacterial P-450cam methylene monooxygenase components: Cytochrome m, putidaredoxin, and putidaredoxin reductase. *Methods Enzymol.* **52**, 166–188 (1978).
163. Hu, S., Schneider, A. J. & Kincaid, J. R. Resonance Raman Studies of Oxycytochrome P450cam: Effect of Substrate Structure on $\nu(\text{O-O})$ and $\nu(\text{Fe-O2})$. *J. Am. Chem. Soc.* **113**, 4815–4822 (1991).
164. Unno, M., Ishimori, K., Ishimura, Y. & Morishima, I. High-Pressure Flash Photolysis Study of Hemoprotein: Effects of Substrate Analogs on the Recombination of Carbon Monoxide to

- Cytochrome P450CAM. *Biochemistry* **33**, 9762–9768 (1994).
165. Aoki, M., Ishimori, K. & Morishima, I. NMR studies of putidaredoxin: Associations of putidaredoxin with NADH-putidaredoxin reductase and cytochrome P450cam. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1386**, 168–178 (1998).
 166. Aoki, M., Ishimori, K., Fukada, H., Takahashi, K. & Morishima, I. Isothermal titration calorimetric studies on the associations of putidaredoxin to NADH-putidaredoxin reductase and CYP450cam. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1384**, 180–188 (1998).
 167. Dang, M., Pochapsky, S. S. & Pochapsky, T. C. Spring-loading the active site of cytochrome P450cam. *Metallomics* **3**, 339–343 (2011).
 168. Lee, Y. T., Wilson, R. F., Rupniewski, I. & Goodin, D. B. CYP450cam visits an open conformation in the absence of substrate. *Biochemistry* **49**, 3412–3419 (2010).
 169. Tang, S. *et al.* Functional characterization of reductive dehalogenases by using blue native polyacrylamide gel electrophoresis. *Appl. Environ. Microbiol.* **79**, 974–981 (2013).
 170. Staikos, G. & Dondos, A. Study of the sodium dodecyl sulphate-protein complexes: Evidence of their wormlike conformation by treating them as random coil polymers. *Colloid Polym. Sci.* **287**, 1001–1004 (2009).
 171. Reynolds, J. A. & Tanford, C. Binding of dodecyl sulfate to proteins at high binding ratios. Possible implications for the state of proteins in biological membranes. *Proc. Natl. Acad. Sci. U. S. A.* **66**, 1002–7 (1970).
 172. Smith, B. J. SDS Polyacrylamide Gel Electrophoresis of Proteins. *Methods Mol. Biol.* **1**, 41–55 (1984).
 173. Tal, M., Silberstein, A. & Nusser, E. Why does Coomassie Brilliant Blue R interact differently with different proteins? A partial answer. *J. Biol. Chem.* **260**, 9976–9980 (1985).
 174. Schägger, H. & von Jagow, G. Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem.* **199**, 223–231 (1991).
 175. Fiala, G. J., Schamel, W. W. A. & Blumenthal, B. Blue native polyacrylamide gel electrophoresis (BN-PAGE) for analysis of multiprotein complexes from cellular lysates. *J. Vis. Exp.* 1–8 (2010). doi:10.3791/2164
 176. Fraser, J. S. *et al.* Accessing protein conformational ensembles using room-temperature X-ray crystallography. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 16247–16252 (2011).
 177. Keedy, D. A. *et al.* Crystal cryocooling distorts conformational heterogeneity in a model michaelis complex of DHFR. *Structure* **22**, 899–910 (2014).
 178. Keedy, D. A. *et al.* Mapping the conformational landscape of a dynamic enzyme by multitemperature and XFEL crystallography. *Elife* **4**, 1–26 (2015).
 179. Kneller, D. W. *et al.* Structural plasticity of SARS-CoV-2 3CL Mpro active site cavity revealed by room temperature X-ray crystallography. *Nat. Commun.* **11**, 7–12 (2020).
 180. Fukuda, Y. & Inoue, T. High-temperature and high-resolution crystallography of thermostable copper nitrite reductase. *Chem. Commun.* **51**, 6532–6535 (2015).
 181. Deis, L. N. *et al.* Multiscale conformational heterogeneity in staphylococcal protein a: Possible determinant of functional plasticity. *Structure* **22**, 1467–1477 (2014).

Chapter 4 : Conformational dynamics of CYP450cam investigated by small angle X-ray scattering

4.1 Introduction to cytochrome P450cam conformational dynamics

Proteins are complex macromolecular machines, catalyzing crucial biological processes in the cell. The advent of protein X-ray crystallography has provided in a surge of information regarding the structure and function of proteins, however the inherent static environment present in protein crystals provides an incomplete picture.¹ Crystallography can be described as the ultimate purification process which selects for a single conformer.² Proteins are highly dynamic molecules, traversing the energy landscape continuously to achieve efficient catalysis. Protein mobility plays an important role in the binding of substrate, release of product and stabilizing different transition states during catalysis. These essential conformational transitions can be induced by small molecule binding or interaction with a partner protein, as is the case in the hemoprotein cytochrome P450 (CYP450).^{3,4} CYP450s are a class of monooxygenases ubiquitous in most life forms that catalyze the insertion of oxygen into a non-activated C-H bond with the addition of molecular oxygen and two electrons obtained from reduced cofactors.^{5,6} Cytochrome P450cam (CYP101A1) from *Pseudomonas putida* is an archetypal CYP450 which catalyzes the regio- and stereo-selective hydroxylation of camphor.⁷⁻⁹ CYP450cam undergoes a variety of conformational changes that depend on substrate binding, substrate analog binding, interaction with its redox partner protein putidaredoxin (Pdx) as well as ligands such as CO and CN⁻.¹⁰⁻¹²

4.1.1 Catalysis by the CYP450cam-Pdx Couple

CYP450cam is responsible for the 5-*exo*-hydroxylation of D-camphor, the first step in camphor catabolism as an energy source.⁷ To achieve this hydroxylation CYP450cam interacts with putidaredoxin (Pdx) which is a 11.6 kDa redox partner protein with a Fe₂S₂ cluster.¹³ Pdx binds to CYP450cam on the proximal side of the heme and is believed to play an important effector role, inducing conformational changes essential for substrate hydroxylation.¹⁴⁻¹⁶ Pdx functions as a shuttle, ensuring the step-wise donation of electrons from NADH to the heme active site of CYP450cam. The redox potential of Pdx in its

reduced form is -240 mV.¹⁷ Upon camphor binding at the active site of CYP450cam, there is a heme iron spin state transition from ferric low spin to high spin, accompanied by a change in the reduction potential from -303 mV to -173 mV (Figure 4.1).^{18,19} This change in reduction potential makes electron transfer from reduced Pdx to CYP450cam feasible and ensures that Pdx only reduces CYP450cam when substrate is bound.^{20,21} Following active site heme reduction, catalysis can proceed with oxygen binding and eventual product formation as described in **Chapter 3.1**. The importance of Pdx as an effector is highlighted by the fact that while the first reduction can be achieved by non-native redox partners including small molecules, Pdx is essential for the second reduction.^{13,22} The absence of Pdx for a second reduction results in decomposition of the reduced CYP450cam-O₂ complex to the oxidized CYP450cam dissociation of a superoxide species.²²

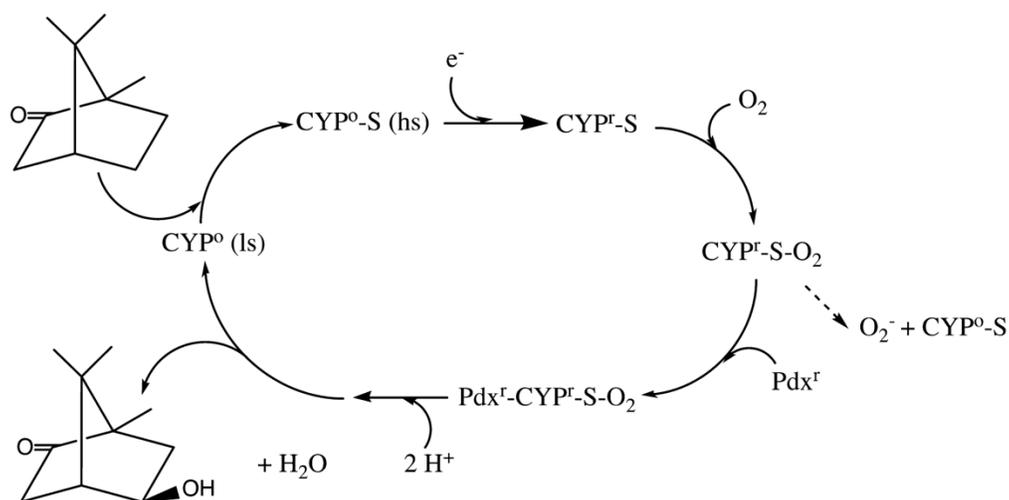


Figure 4.1 The catalytic cycle of CYP450cam. The oxidized CYP450cam (CYP⁰) in the low spin state (ls) binds the substrate, D-camphor and transitions to a high spin state (hs). The reduced CYP^f binds oxygen and reduced Pdx^f which can form the active complex to form the product 5-*exo*-camphor. If Pdx^f does not bind, the oxygen is reduced by the heme iron to superoxide and dissociates. Adapted from Pochapsky *et al.*² and reprinted with permission of the American Chemical Society (ACS). Please contact publisher for reuse.

4.1.2 Substrate-dependent CYP450cam conformational changes

The fundamental conformational change of CYP450cam from an open to a closed conformation upon camphor binding was characterized by Lee *et al.* using X-ray crystallography.²³ Prior to this, CYP450cam had been widely believed to be a static system.^{24,25} Lee *et al.* however established that in the

absence of camphor CYP450cam mostly adopts an open conformation in which the F- and G-helices and FG-loop retract forming a substrate binding channel and the B'-helix is in a disordered state (Figure 4.2). Camphor binding, results in closure of the substrate access channel formed by the F- and G-helices and the B'-helix adopts an ordered conformation. The FG-loop displayed the most movement from the open form, with a 9 Å displacement.

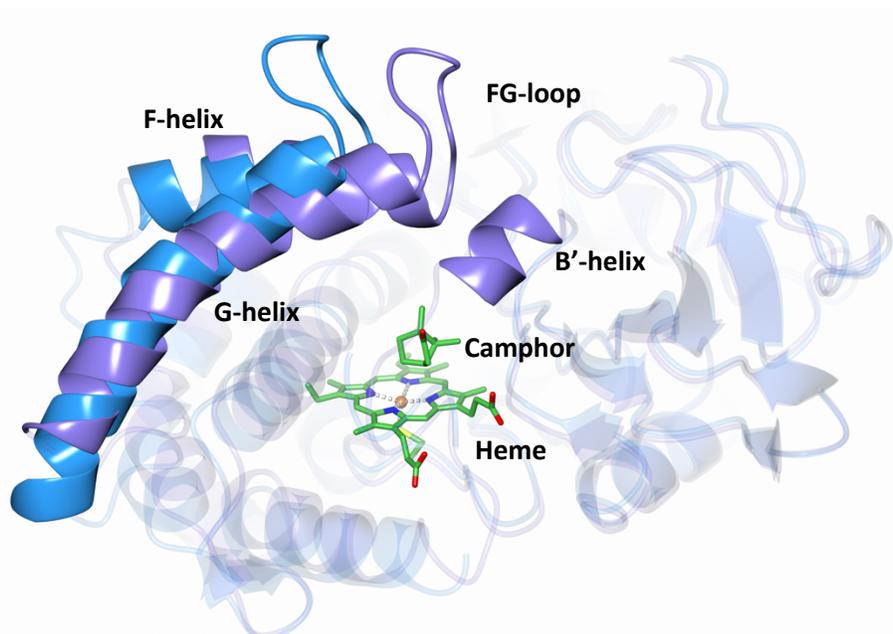


Figure 4.2 The open (PDB 3L61) and closed conformation (PDB 2CPP) of CYP450cam. Upon substrate binding, CYP450cam reverts to a closed conformation in which the substrate binding channel, delimited by the F-helix, G-helix and FG-loop contracts and the B'-helix adopts an ordered conformation.

The F- and G-helices showed a 5-6 Å movement when compared the open conformation. The open CYP450cam form a solvent-filled channel which permits substrate entry to the active site. These water molecules are absent in the closed camphor-bound conformation along with the access channel. This suggests that upon substrate binding, the CYP450cam active site closes to ensure productive substrate binding and prevent egress of substrate and reactive intermediates. On the proximal face it was noted that Arg112 and Leu358 move away from the heme in the closed conformation. The existence of an open and closed conformation were subsequently supported by computational, infrared spectroscopy and NMR studies.²⁶⁻³⁰

4.1.3 Structural characteristics of the CYP450cam-Pdx complex

Early attempts to determine the structure of Pdx were hampered by its intractability to crystallize. Solution NMR studies represented the first headway made to determine structure of Pdx by substitution of the iron with gallium to avoid paramagnetic broadening due to the two high-spin iron atoms.^{31,32} Sevrioukova *et al.* were subsequently able to solve the crystallographic structure of Pdx by generating a surface cysteine mutant that permitted crystallization.¹³ The structure of Pdx is made up of a hydrophobic core composed of a five stranded β -sheet which are flanked by α -helices and a protruding C-terminal involved in interaction with CYP450cam. The effector role of Pdx had been established, with kinetics studies indicating that substrate hydroxylation occurs when Pdx is added to oxy-ferrous substrate bound CYP450cam.^{22,33}

The structural basis of the role of Pdx as an effector was confirmed crystallographically by Tripathi *et al.* using a covalent CYP450cam-Pdx complex.³⁴ They determined the crystal structure of oxidized and reduced CYP450cam complexed to Pdx. In the complex, Pdx binds to the proximal side of CYP450cam (Figure 4.3 A). Interestingly, analysis of the reduced complex structure revealed that of the four molecules in the asymmetric unit, three contained the hydroxylated product, 5-*exo*-camphor which confirms the complex represents an active form of CYP450cam. Residues that had been predicted to be involved in CYP450cam-Pdx interactions by EPR were confirmed by the crystal structure.³⁵ The Asp38 and C-terminal Trp106 are involved in interactions with CYP450cam. Pdx Asp38 forms an electrostatic interaction with CYP450cam Arg112. Pdx Trp106, whose central role in interaction with CYP450cam has been confirmed by mutagenesis,³⁶ forms numerous interactions with CYP450cam: Pdx Trp106 forms hydrophobic interactions with CYP450cam Ala113 and CYP450cam Arg109 forms electrostatic interactions with the Ser44 carboxyl group (Figure 4.3 B). Their structure revealed that Pdx binding shifts CYP450cam toward an open conformation. The open conformation of CYP450cam in the substrate free form had been previously characterized by CYP450cam the retraction of the F- and G-helices and FG-loop to form a substrate binding channel and open the active site.

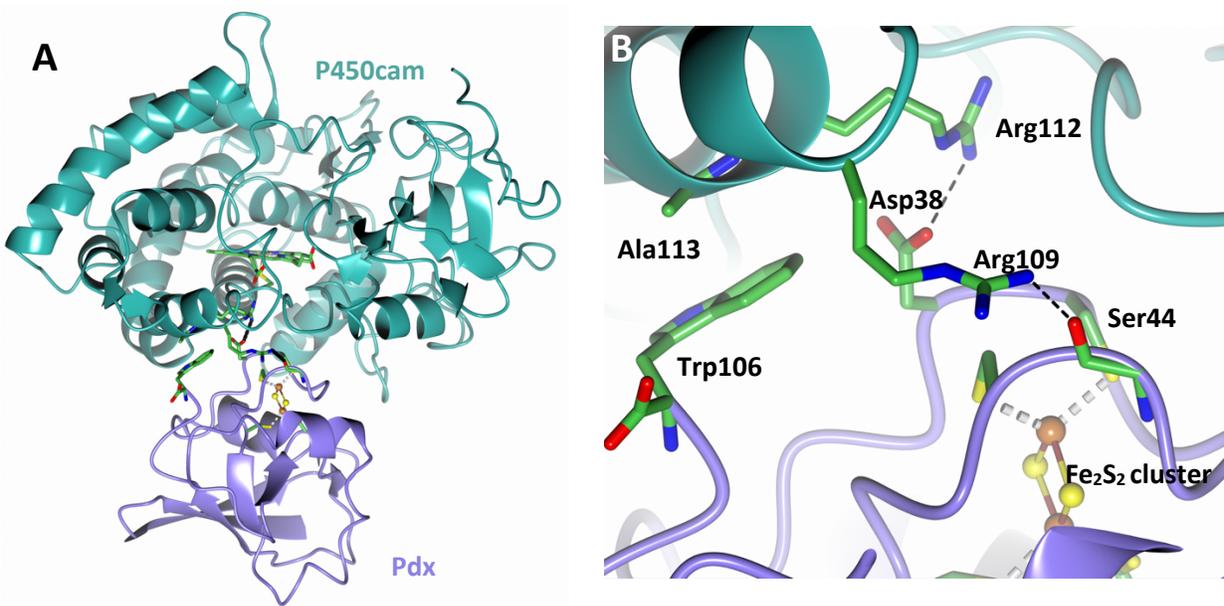


Figure 4.3 The CYP450cam-Pdx complex (PDB 4JWU). **(A)** Pdx binds on the proximal face of CYP450cam. **(B)** Pdx Trp106 forms hydrophobic interactions with CYP450cam Ala113 and CYP450cam Arg109 forms electrostatic interactions with the Ser44 carboxyl group.

The Pdx-dependent conformational change primarily involves movement around the Pdx docking site and the substrate access channel. The C-helix of CYP450cam moves ~ 3 Å to accommodate interactions with Pdx Trp106. The movement of the CYP450cam C-helix is coupled with movement of helices F, G, I and B' with the net effect of opening the active site. The large I-helix, that traverses the active site and contains the characteristic kink involved in oxygen binding,³⁷ also indicated structural rearrangement upon Pdx binding. When oxygen binds near the active site, the hydrogen bond pair formed by Gly248 and Thr252 in the I-helix is disrupted, with Gly248 moving closer towards the bound oxygen. The oxygen-bound CYP450cam conformation can be described as halfway between the open and closed conformation. This led the authors to propose that Pdx favors binding this “more” open conformation, resulting in the shift to a “fully” open conformation, widening the I-helix kink and enabling water molecules to enter, thereby forming the protonation pathway necessary for catalysis. The CYP450cam Asp251 residue which has been proposed to be involved in shuttling protons to the active site,³⁸ was also found to be freed from salt bridge interactions with Lys178 and Arg186 when Pdx binds which further supports the role of Pdx as an effector that induces the conformational changes necessary for catalysis. This has been termed the “Asp switch

hypothesis” in which the open conformation induced by Pdx frees Asp251 from a salt bridge interaction with Lys178 and Arg186 ensuring that water enters the active site providing the proton relay network required for camphor hydroxylation can form.³⁴ This Tyr33 has been shown to be have conformational flexibility and play a role in complex formation.^{39,40}

In another study published in the same year by Hiruma *et al.*, paramagnetic NMR spectroscopy and X-ray crystallography were used to determine the CYP450cam-Pdx complex structure.⁴¹ While the positioning of Pdx relative to CYP450cam is very similar to that found by Tripathi *et al.*, there are some differences including different orientations of CYP450cam Arg109, Asn116 and Leu358. There is also some discrepancy of the residues interacting with Pdx Trp106 and the authors believe that this may be due to differing redox states. The structure determined by NMR in solution also indicates that Pdx Tyr33 is highly mobile and distant from CYP450cam, while in the crystal structure it interacts with CYP450cam Pro 122 and Asp 125. The dynamics of Tyr33 may play a role during Pdx to CYP450cam electrons transfer. Similar to Tripathi *et al.* the CYP450cam displays open state characteristics with opening of the substrate access channel. The open conformation of CYP450cam due to Pdx binding was further confirmed by double electron-electron resonance (DEER).^{15,16}

There have, however, been studies that refute the open conformation that Pdx is said to induce. An NMR study on the same complex found that the CYP450cam-Pdx complex remains in a closed state.⁴² They postulated that the closed state is required to ensure substrate retention in the active site and regioselectivity. Given this disparity of the open and closed forms, Batabyal *et al.* performed a molecular dynamics (MD) and mutagenesis study to investigate the CYP450cam conformational dynamics.⁴³ Principle component analysis (PCA) of the MD simulations suggested that Pdx induces an intermediate state in CYP450cam, between the open and closed states. The authors establish a more nuanced role for Pdx, where a partially open conformation is induced that frees Asp251 to form a proton relay network, however the active site remains stabilized in a more closed conformation which ensures camphor retention

and correct positioning. The F/G helical region more resembles the closed state with less flexibility, ensuring that substrate remains bound. Camphor mobility is also decreased ensuring productive substrate binding by active site stabilization. This selective conformational change by Pdx targeting only Asp251 was confirmed by mutagenesis study in which Asp251 was freed from its salt bridge interaction with Arg186. The Arg186Ala mutant structure showed that Asp251 had been released from the salt bridge to form the proton relay network and structural analysis further revealed the product 5-*exo*-camphor in the active site, supporting a targeted conformational change by Pdx involving Asp251 to ensure catalysis. This intermediate conformation of CYP450cam and Pdx was observed in a complex with the O₂ mimic CN⁻, using DEER analysis.¹¹ The conformational changes involved the residues Arg186, Asp251 and Thr252 within the I-helix to enable water molecule entry and formation of the proton transfer pathway. The F-helix was however not in a fully open conformation which prevents opening of the active site and loss of substrate.

4.1.4 The L358P mutation

Numerous spectroscopic studies including NMR,⁴⁴ EPR,⁴⁵ IR⁴⁶ and Resonance Raman⁴⁷ were initially used to characterize the conformational changes induced on CYP450cam by Pdx. It was found that the same spectroscopic properties caused by perturbations due to Pdx were achieved by use of a Leu538 to Pro mutant (L358P).⁴⁸ This was confirmed by formation of the product 5-*exo*-camphor in the presence of oxygen and non-physiological, small molecule donors such as ascorbic acid and dithionite. Solution of the crystal structure of L358P indicated that the mutation induced several structural changes similar to the conformational changes caused by Pdx docking.⁴⁹ The structural changes included movements of CYP450cam Arg112, increased proximity of the heme to the substrate camphor and structural changes in the I-helix to aid protonation pathway formation.

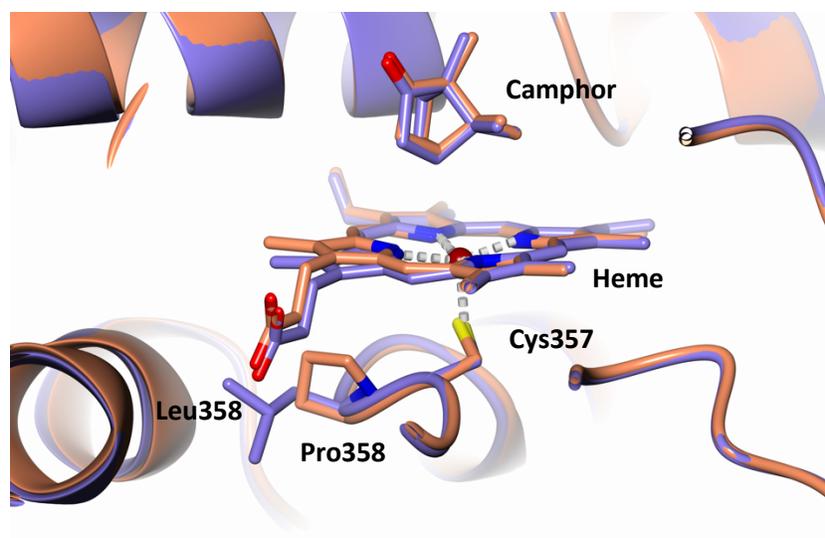


Figure 4.4 Position of the CYP450cam L358P mutation (PDB 1T86). The L358P mutation is located on the proximal face of the heme in the vicinity of the Pdx binding site. Wild type CYP450cam displayed in purple L358P mutant displayed in coral.

The CYP450cam Leu358 residue is positioned on the Pdx binding site after the heme coordinating Cys357 residue (Figure 4.4).⁴⁹ The effect of the L358P mutation is both electronic and steric. Firstly, in the wild-type, partial positive charge on the peptide amide connecting Cys357 to Leu358 stabilizes the negative charge localized on the cysteine sulfur. Mutation of Leu358 to proline decreases this electronic stabilization, increasing the negative charge distribution on the sulfur which promotes cleavage of the O-O bond of bound oxygen. Secondly, substitution of Leu358 for the rigid proline residue has numerous steric effects. Pro358 causes the C-helix to push against the heme, which in turn results in the I-helix converting from a closed conformation to partially open, with Thr252 adopting an alternate conformation which allows water to enter the active, ensuring formation of the proton transfer pathway for intermediate protonation.

4.1.5 An allosteric site for CYP450cam

The structural changes induced by substrate and redox partner protein interactions discussed above demonstrates that CYP450cam function is governed by a delicate conformational balance. The most recent facet of this dynamic picture of CYP450cam was introduced by the discovery of an allosteric site on CYP450cam.⁵⁰ Using molecular dynamics (MD), Follmer *et al.* defined an allosteric site on CYP450cam that results in opening of the substrate binding channel formed the the F- and G-helices, FG-loop and B'-

helix, designated “Channel 1”, and primes formation of a new channel thought to be involved in substrate egress and designated “Channel 2”.⁵⁰ Follmer *et al.* subsequently published a X-ray crystal structure of the CYP450cam-Pdx complex with CN⁻ bound to mimic O₂ binding at the active site iron.⁵¹ The CN⁻ bound complex displayed significant conformational changes including Pro89 undergoing a *cis-trans* isomerization and they observe formation of the theoretically proposed Channel 2.

The allosteric site proposed by Follmer *et al.* is located between the C-, D-, and E-helices near the residues Val118, Val123, Lys126, Leu127, Thr217, Ala219 and Leu166 (Figure 4.5).⁵⁰ Previously, NMR studies found that a similar region shows perturbations in the presence of high camphor levels.^{52,53} Formation of Channel 2 was observed in MD simulations near the residues spanning Ser83 – Ser102. The presence of a second camphor binding site within ~ 16 Å of the heme was suggested by Yao *et al.* with a K_D of ~ 43 μM.⁵⁴ Furthermore, a X-ray crystal structure by Lee *et al.* of a substrate-free CYP450cam which was subsequently soaked in camphor appears to have a *F_o-F_c* peak near the proposed allosteric site at 7.σ but was not considered definitive structural evidence of an allosteric site.^{23,50}

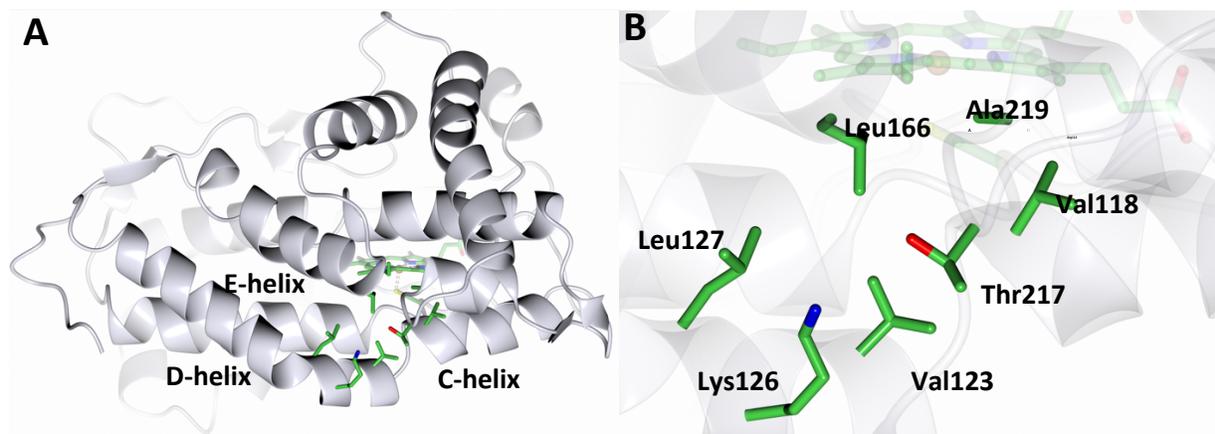


Figure 4.5 Proposed CYP450cam allosteric binding site (PDB 1DZ4). (A) The CYP450cam allosteric site where a second camphor binds to the surface of CYP450cam is located near the C-, D-, and E-helices. (B) The residues comprise the camphor binding site comprise Val118, Val123, Lys126, Leu127, Thr217, Ala219 and Leu166.

Follmer *et al.* propose that after binding of one camphor molecule to the allosteric site, there is opening of Channel 1 by movement of the FG-loop, allowing substrate to bind and Channel 2 is formed

soon after by mechanical coupling, permitting product egress.⁵⁰ A cooperative model is proposed in which camphor binding to the allosteric site favors a more open conformation allowing a second camphor to enter the active site which is followed by dissociation of camphor from the allosteric site and formation of Channel 2 for product egress. MD simulations with the Pdx binding mimic L358P suggest that since both Pdx binding and camphor allosteric binding promote formation of the open conformation, their binding is synergistic. Additionally, allosteric camphor binding was found to make the Leu358 rotamer more flexible. Since Pdx binding results in an altered Leu358 dihedral angle, allosteric camphor binding may support Pdx binding and subsequent induction of an open conformation. In terms of a biological context, the authors propose that this allosteric site ensures that camphor is only utilized as a carbon source when it is in abundance, at high concentrations. At low concentrations CYP450cam is camphor bound, but in a closed conformation. Higher camphor concentrations result in camphor binding to the allosteric site, resulting in the open conformation of CYP450cam which promotes Pdx binding. This can also be associated with the CamR repressor of CYP450cam expression which only dissociates from the CamR regulatory sequence at high camphor concentrations.⁵⁵

To further examine the role of the CYP450cam allosteric site and camphor binding in solution, where there are no crystal packing artefacts, Skinner *et al.* performed paramagnetic NMR pseudocontact shift (PCS) analysis along with MD simulations.⁵⁶ The MD simulations of only CYP450cam indicate that in excess camphor, a camphor molecule binds to the allosteric site, opening Channel 1 and priming Channel 2 to open. This is followed by a second camphor binding in the active site near the first camphor which widens Channel 2. Formation of Channel 2 was however not observed for the CYP450cam-Pdx complex in MD simulations. PCS analysis supports the presence of an allosteric site involving the residues Lys126, Leu166, Gly168, Thr217 and Ala219 as proposed by Follmer *et al.*⁵⁰ Additionally, distance distribution analysis of the F-, G- and B'-helices indicate that Pdx binding results in an intermediate open-closed conformation as has been proposed by Batabyal *et al.*⁴³ Given these results, the authors propose the following mechanism: at low camphor concentrations CYP450cam is present in the closed camphor-bound

conformation. At high concentrations of camphor, a second camphor binds to the allosteric site, resulting in a partially open CYP450cam conformation in which the substrate entry channel is partially opened and the Arg186-Asp251 salt bridge is destabilized. Pdx preferentially binds to this partially open conformation resulting in breakage of the Arg186-Asp251 bridge and formation of the protonation pathway as well as formation of Channel 2 to permit product egress.

4.1.6 Small angle X-ray scattering studies of CYP450cam

Lewis *et al.* performed a small angle X-ray scattering (SAXS) study on CYP450cam to investigate the CYP450cam conformation in the presence and absence of camphor.⁵⁷ SAXS is a solution scattering technique that provides information on the size, shape, conformational flexibility and macromolecular architecture in native conditions.⁵⁸ SAXS measure the intensity of X-rays scattered by a sample as a function of the angle at which they are scattered.⁵⁹ On a basic level, a SAXS experiment permits the characterization of the molecular weight, extended particle volume, overall size of the sample by determining the radius of gyration (R_g) and the sample's maximum dimension (D_{max}).⁶⁰ The SAXS studies performed by Lewis *et al.* found that the CYP450cam showed characteristics of an elongated protein. The R_g of CYP450cam, i.e., the average weighted distance to the center of the molecule, was found to be $24.14 \text{ \AA} \pm 0.20 \text{ \AA}$ and $23.93 \text{ \AA} \pm 0.21 \text{ \AA}$ for the camphor-free and bound forms, respectively. This led the authors to conclude that, within experimental error, there was no change between the camphor-free and camphor-bound forms. Given the deviation of the R_g from what is expected for a perfect sphere ($R_g = 18.5 \text{ \AA}$), the authors predict an ellipsoid shape with maximum dimensions of 80 \AA . Given the occurrence at that time of tetragonal CYP450cam crystals with unit cells of $64 \times 64 \times 255 \text{ \AA}$ (8 asymmetric units per unit cell) and orthorhombic crystals with unit cell dimensions of $36 \times 104 \times 109 \text{ \AA}$ (4 asymmetric units per unit cell),⁶¹ negated the possibility of an oblate ellipsoid or short cylinder. It was therefore concluded that CYP450cam forms an approximate cylinder 80 \AA in length 30 \AA wide. The solution of the crystal structure of CYP450cam two years later at a resolution of 2.6 \AA , revealed that CYP450cam forms a triangular prism with a maximum dimension of 60 \AA and a minimum dimension of 30 \AA .^{9,24,62}

4.2. Aim of research

CYP450cam is a highly dynamic system with a delicate conformational balance regulating activity. Studies of CYP450cam in isolation as well as in complex with its redox partner protein Pdx have revealed how these effectors may regulate catalysis. It is essential, however, to take into account the artefacts that may be introduced by crystal packing in X-ray diffraction studies. Crystallization may only capture the conformations that will fit into the crystal lattice or reduce the mobility of regions that have a dynamic role in catalysis. Solution studies of CYP450cam are therefore necessary to elucidate catalytic insights that a crystal structure will not provide. The aim of this research is to study the conformation of CYP450cam using small angle X-ray scattering (SAXS). SAXS provides information of the conformational state, shape and size of a protein or assembly in native conditions in solution. Using SAXS, we will be able to determine the effects that substrate binding and redox partner protein interaction have on the structural flexibility and conformational rearrangements of CYP450cam. The role of substrate as an allosteric effector has recently been postulated, with the second binding site for camphor at an allosteric providing a more nuanced picture of the reaction. The conformation of CYP450cam in solution will be studied by SAXS to determine the effect of increasing camphor concentrations, the Leu358Pro mutation which results in an open conformation of CYP450cam an interaction with Pdx. SAXS studies were performed on CYP450cam in the early 1980s, however these were performed at limited camphor concentration ranges and did not investigate the role redox partner binding.⁵⁷ We therefore aim to create a more complete picture of CYP450cam dynamics.

4.3 Materials and methods

4.3.1 Reagents

All reagents, unless otherwise stated were obtained from Fischer Scientific (Hampton, New Hampshire, USA) and Sigma-Aldrich (St. Louis, Missouri, USA) and purifications columns were from GE Healthcare Life Sciences (Chicago, Illinois, USA).

4.3.2 Mutant constructs

4.3.2.1 CYP450cam C334A

The CYP450cam C334A mutant was synthesized (Genscript) into the pET-15b vector (Novagen). This CYP450cam contains no His₆-tag and will be termed CYP450cam(C334A).

4.3.2.2 CYP450cam L358P

The CYP450cam with the mutations C334A and L358P mutant was synthesized (Genscript) into the pET-28a vector (Novagen). This CYP450cam contains a N-terminal His₆-tag and will be termed CYP450cam(L358P).

4.3.2.3 CYP450cam crosslink mutant

The CYP450cam with the mutations C59S, C86S, C137S, C286S, C334A and K344C was synthesized (Genscript) into the pET-28a vector (Novagen). This CYP450cam will be termed CYP450cam(Crosslink).

4.3.2.4 Pdx crosslink mutant

The Putidaredoxin (Pdx) gene with the mutations D19C, C73S and C85S was synthesized (Genscript) into the pET-28a vector (Novagen). This Pdx contain a N-terminal His₆-tag and will be termed Pdx(Crosslink).

4.3.3 Protein Expression

CYP450cam(C334A), CYP450cam(L358P), CYP450cam(Crosslink) and Pdx(Crosslink) were overexpressed in *E.coli* BL21(DE3) grown in Luria-Bertani broth for to an OD₆₀₀ of 0.7 after which expression was induced with 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 24 h at 25°C. The CYP450cam mutants were additionally supplemented with 1 mM 5-aminolevulinic acid and 50 μM FeCl₃ upon induction. Pdx(Crosslink) was additionally supplemented FeCl₃ upon induction.

Mutant CYP450cam cell pellets were resuspended in 50 mM Tris pH 7.8 and 0.5 mg/mL lysozyme. Additionally, the Lee *et al.* protocol was followed and 0.1 mM hemin chloride added to the *E. coli* harvested cells.⁶³ Cells were lysed by sonication and centrifuged at 30 000g for 45 min. Pdx(Crosslink) cell pellets were resuspended in either 50 mM KP_i pH 7.6, 50 mM KP_i pH 6 and 0.5 mg/mL lysozyme. Cells were lysed by sonication and centrifuged at 30 000g for 45 min.

4.3.4 Purification

4.3.4.1 CYP450cam

Anion exchange purification

P450cam mutants cell lysate was loaded onto a either a HiTrap Q XL column equilibrated with 50 mM Tris pH 7.8. Elution was performed using 50 mM Tris pH 7.8 and 1 M KCl in a linear gradient. The eluate was pooled from a peak obtained monitoring the absorbance peak at 391 nm and fractions were concentrated using a 30 kDa cutoff centrifugal filter (Amicon).

Size exclusion chromatography

The CYP450cam mutants from the anion exchange purification were loaded onto a Superdex 75 prep grade, HiLoad 16/60, 120 mL column equilibrated with a wash buffer composed of 50 mM KP_i pH 7.6 and 250 mM KCl. The eluate was pooled from a peak obtained monitoring the absorbance peak at 391 nm and fractions were concentrated using a 30 kDa cutoff centrifugal filter (Amicon).

4.3.4.2 Putidaredoxin

Ni-affinity purification

Pdx(Crosslink) lysate was loaded onto a HisTrap HP column equilibrated with 50 mM KP_i pH 7.6, 250 mM KCl, and 20 mM imidazole pH 7.6. Elution was performed using 50 mM KP_i pH 7.6, 250 mM KCl and a linear gradient of 500 mM imidazole pH 7.6. The eluate was pooled from a peak obtained monitoring the absorbance peak at 391 nm and fractions were concentrated using a 5 kDa cutoff centrifugal filter (Amicon).

Size exclusion chromatography

Pdx from the Ni-affinity purification was loaded onto a Superdex 75 prep grade, HiLoad 16/60, 120 mL column equilibrated with a wash buffer composed of 50 mM KP_i pH 7.6 and 250 mM KCl. The eluate was pooled from a peak obtained monitoring the absorbance peak at 391 nm and fractions were concentrated using a 5 kDa cutoff centrifugal filter (Amicon).

4.3.5 Complex formation

In order to form the CYP450cam-Pdx complex, the chemical crosslink approach utilized by Tripathi *et al.* was followed.³⁴ The CYP450cam(Crosslink) mutant with surface cysteines removed (C59S, C86S, C137S, C286S and C334A) and a cysteine near the Pdx binding site (K344C) was used. The Pdx(Crosslink) mutant with surface cysteines (C73S and C85S) removed and a cysteine near the CYP450cam binding site (D19C) was used.

Complex formation was performed using purified CYP450cam(Crosslink) and Pdx(Crosslink). To perform the crosslinking reaction, 100 μM homobifunctional crosslinker bis(maleimido)hexane (BMH), was added to 5 μM of CYP450cam and 50 μM Pdx in 20 mM KP_i pH 7.6 pH 7.4 and 50 mM KCl. This mixture was stirred overnight at 4°C.

Purification of the complex from the reaction mixture was first performed using a 50 kDa cutoff centrifugal filter (Amicon) to ensure removal free Pdx(crosslink). The complex was then loaded onto a a

HisTrap HP column equilibrated with 50 mM KP_i pH 7.6, 250 mM KCl, and 20 mM imidazole pH 7.6. Elution was performed using 50 mM KP_i pH 7.6, 250 mM KCl and a linear gradient of 500 mM imidazole pH 7.6. The eluate was pooled from a peak obtained monitoring the absorbance peak at 391 nm and fractions were concentrated using a 50 kDa cutoff centrifugal filter (Amicon).

4.3.6 Characterization

4.3.6.1 SDS-PAGE

Gel electrophoresis was performed on whole cells following expression as well as purification fractions using sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE).⁶⁴ Samples were run on Mini-PROTEAN TGX Stain-Free Precast Gels (BioRad Laboratories, Hercules, California, USA) and performed according to the manufacturer's protocol using 5 µL of the Precision Plus Protein Unstained Standards (BioRad Laboratories, Hercules, California, USA).

4.3.6.2 CO-difference spectrum

Whole cells of CYP450cam mutants were used for CO difference spectroscopy. CYP450cam was reduced with a few grains of sodium dithionite in duplicate and one of the samples was exposed to CO in a sealed environment for 10 min. Absorbance of the CO-bound and free CYP450cam was measured from 360 – 500 nm in a microtiter plate using a SpectraMax i3 spectrophotometer (Molecular Devices, San Jose, California, USA).

4.3.7 Small angle X-ray scattering sample preparation

4.3.7.1 Concentration series

CYP450cam(C334A) was prepared in 1 mg/mL, 2 mg/mL, 3 mg/mL, 4 mg/mL, 5 mg/mL, 6 mg/mL in 50 mM KP_i pH 7.6, 250 mM KCl and 1 mM D-camphor. Buffer containing KP_i pH 7.6, 250 mM KCl and 1 mM D-camphor was used for buffer subtraction.

4.3.7.2 Camphor binding

All CYP450cam mutants and the CYP450cam-Pdx complex were analyzed at a 6 mg/mL concentration in 50 mM KP_i pH 7.6 and 250 mM KCl. CYP450cam(L358) samples in 0 μM, 1.6 μM, 43 μM and 1000 μM camphor concentrations were measured. CYP450cam(Crosslink) samples in 0 μM, 1 μM, 1.6 μM, 2 μM, 10 μM, 20 μM, 43 μM, 60 μM, 100 μM and 1000 μM camphor concentrations were measured. The CYP450cam-Pdx complex in 0 μM, 1.6 μM, 43 μM and 1000 μM camphor concentrations were measured. All buffers containing camphor at the assigned concentrations and KP_i pH 7.6 and 250 mM KCl were prepared for buffer subtraction.

4.3.8 Small angle X-ray scattering data collection

SAXS data was collected on a Rigaku Bio-SAXS 2000 instrument which uses Cu K_α radiation at $\lambda=1.54 \text{ \AA}$. The X-ray beam was collimated to provide a scattering vector q -range of $0.008 - 0.67 \text{ \AA}^{-1}$, where q is defined as $q = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle and λ is the incident wavelength.⁶⁵ Samples were measured for a total of 90 min with 10 min scans and all data were corrected for transmission and sample background. Reduction and averaging were performed with Rigaku SAXSLab 3.1.0b14.

To perform buffer subtraction, Guinier fits and Pair-Distribution $P(r)$ function calculations, the BioXTAS RAW program was used.⁶⁶ Guinier analysis was performed at low q ,⁶⁷ and the radius of gyration (R_g) was derived using the Guinier approximation:

$$I(q) = I(0) \left(-\frac{q^2 R_g^2}{3} \right)$$

with an upper limit defined by $q \cdot R_g = 1.3$. Molecular mass was determined from the SAXS data using the volume of correlation based molecular weight calculation, V_c ,⁶⁸ which uses $I(0)$ and R_g and is independent of protein concentration. D_{\max} were manually chosen in the RAW interface using GNOM to optimize the $P(r)$ function.⁶⁹ To determine how well the experimental data fit theoretical curves from atomistic models, the FOXS server was used.^{70,71} The PDB structures used for the atomistic theoretical models included 3L61 - camphor free, 2CPP - camphor bound, 1T86 - L358P mutant camphor bound, 4JWS - CYP450cam-Pdx complex camphor free and 4JX1 - CYP450cam-Pdx complex camphor bound. In order to refine the high-resolution crystallographically determined CYP450cam structures and ensure improved agreement with experimentally determined SAXS data, normal mode analysis (NMA) was used to explore their conformational space in SREFLEX.⁷²

4.4 Results and discussion

4.4.1 Protein expression and purification

CYP450cam(C334A), CYP450cam(L358P) and CYP450cam(Crosslink) were successfully expressed in the absence of camphor to ensure no camphor content for subsequent small angle X-ray scattering (SAXS) studies (Figure 4.6). Expression was performed with the addition of 5-aminolevulinic acid, a heme precursor,⁷³ and FeCl₃ to ensure heme incorporation as discussed in **Chapter 3.2**. Additionally, the Pdx(crosslink) mutant was successfully expressed with addition of FeCl₃ during expression to ensure formation of the Fe₂S₂-cluster.

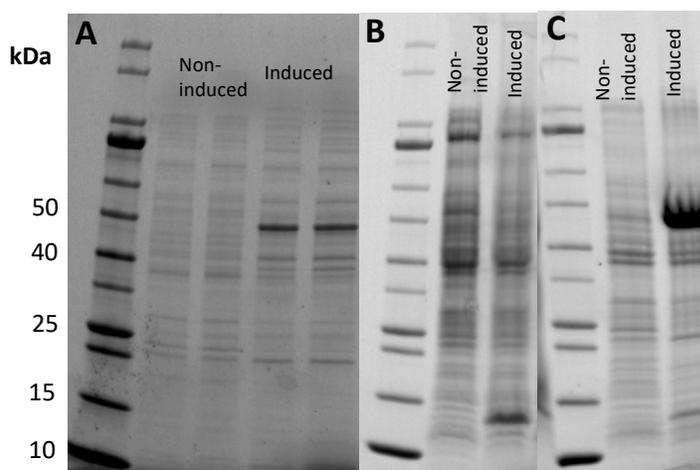


Figure 4.6 SDS-PAGE of CYP450cam mutant expression. **(A)** CYP450cam(L358P) expression. **(B)** Pdx(Crosslink) expression. **(C)** CYP450cam(Crosslink) expression.

The first step of purification of all the CYP450cam mutants was performed using a strong Q-column anion exchange column. This was followed by two rounds of SEC to obtain a pure sample (Figure 4.7). To ensure that the CYP450cam(Crosslink) mutant had been expressed and purified in an active form, a CO-difference spectrum was performed. The CO-difference spectrum displayed a peak at 450nm suggesting that the surface cysteine mutations had not interfered with heme incorporation at the active site (Figure 4.8).

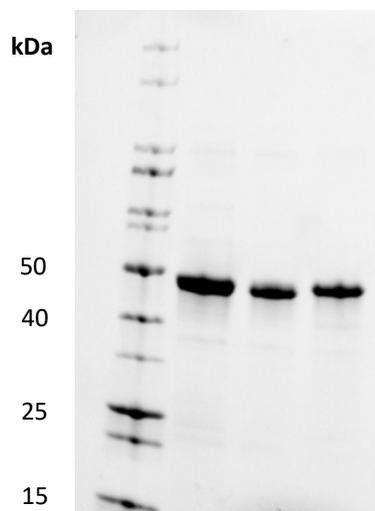


Figure 4.7 Representative SDS-PAGE of pure CYP450cam(Crosslink) samples.

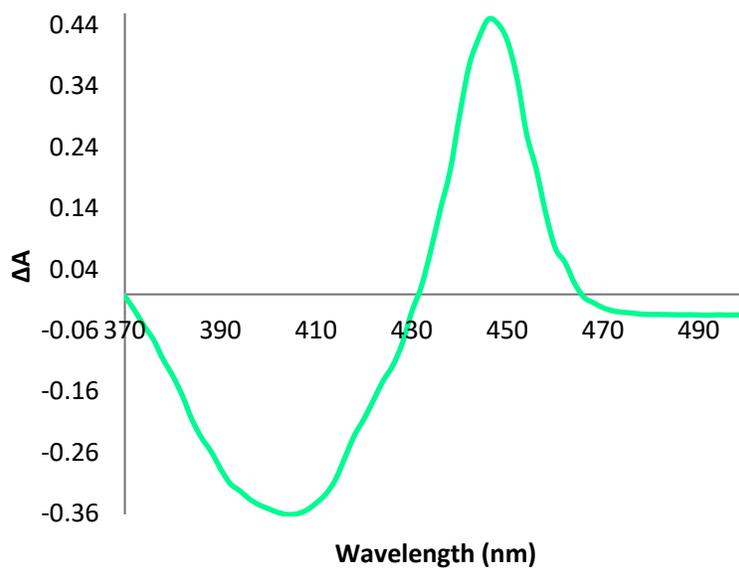


Figure 4.8 Representative CO-difference spectrum of CYP450cam(Crosslink).

Purification of the His6-tagged Pdx(Crosslink) was performed using Ni-affinity chromatography followed by SEC to ensure that a pure sample was obtained (Figure 4.9).

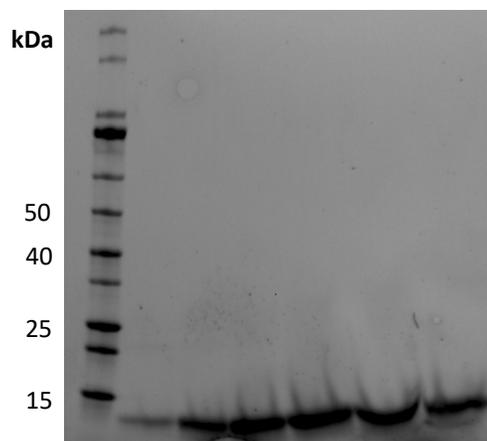


Figure 4.9 SDS-PAGE of purified Pdx(Crosslink) shown in a concentration range.

4.4.2 Complex formation

The role of Pdx as an effector during CYP450cam catalysis was studied by Tripathi *et al.* in 2013 using a covalent CYP450cam-Pdx complex.³⁴ This complex was generated using the homobifunctional maleimide crosslinker, bis(maleimido)hexane (BMH) (Figure 4.10). X-ray crystallographic studies indicated successful formation of the CYP450cam-Pdx complex in a 1:1 ratio, with Pdx binding at the correct proximal site.

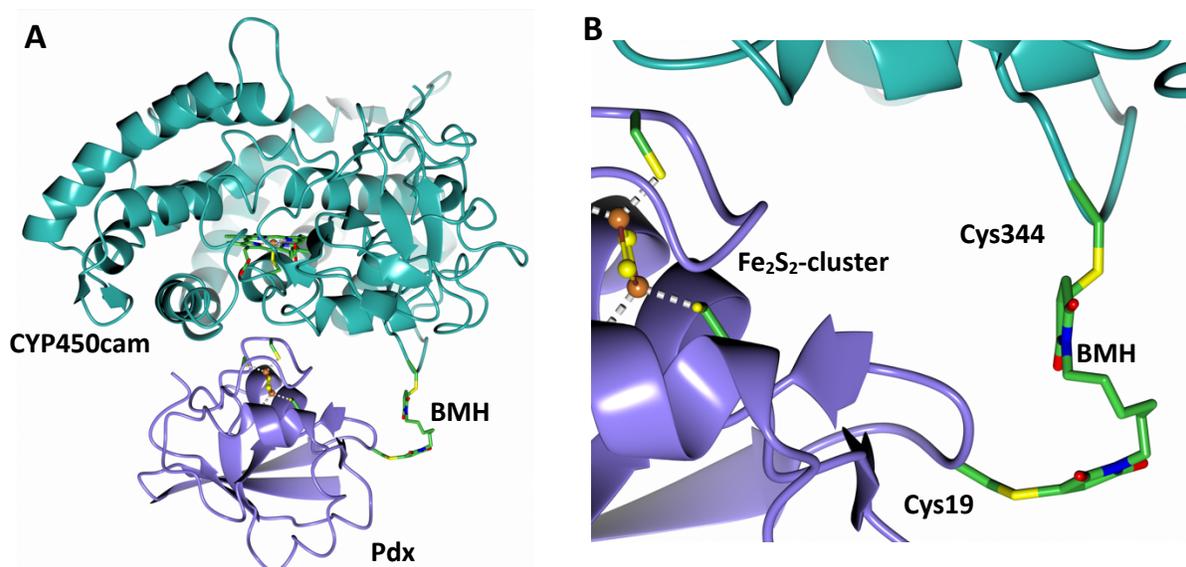


Figure 4.10 Structure of the covalent CYP450cam-Pdx complex (PDB 4JWU). **(A)** The CYP450cam(Crosslink) mutant is covalently bound to the Pdx(Crosslink) mutant *via* BMH. **(B)** The mutated Cys344 CYP450cam(Crosslink) is covalently bound to the mutated Cys19 of Pdx(Crosslink) through a BMH linker.

In order to study the effect of Pdx binding on CYP450cam, this BMH linked covalent complex was formed. A CYP450cam(Crosslink) mutant in which the surface cysteine residues have been mutated to serine to avoid nonspecific crosslinking to Pdx(Crosslink). A new cysteine residue was introduced at Lys344 as a primary site for crosslinking (K344C) of CYP450cam to Pdx. Additionally, a cysteine was introduced at Asp19 of Pdx (D19C) to crosslink to the Cys344 residue *via* BMH (Figure 4.10 B). The complex has been successfully formed as is illustrated in the SDS-PAGE gel (Figure 4.11).

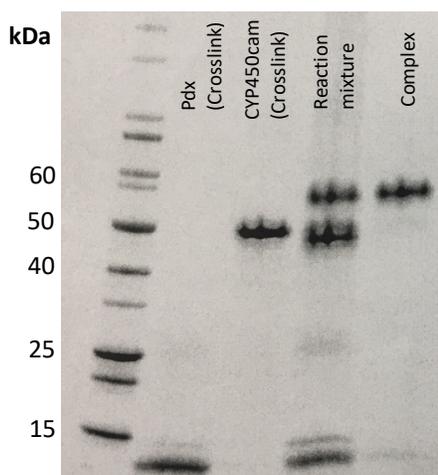


Figure 4.11 SDS-PAGE of components of the crosslink reaction mixture and the formed CYP450cam-Pdx complex.

4.4.3 SAXS data evaluation

Guinier analysis

Evaluation of the particle size, data quality and aggregation was performed using a Guinier analysis.⁶⁷ Two parameters are obtained from Guinier analysis: the radius of gyration (R_g), which is a measure of the particle size, represents the average root-mean-square distance to the center of particle density and $I(0)$, the intensity at zero scattering angle $q = 0$, which can be related to the molecular weight of the particle.

The Guinier analysis is a plot of $\ln I(q)$ vs q^2 in the low q region of the scattering profile. The Guinier approximation is represented by

$$I(q) = I(0) \left(-\frac{q^2 R_g^2}{3} \right)$$

which is linearized to

$$\ln I(q) = \ln I(0) - \frac{R_g^2}{3} \cdot q^2$$

where $I(q)$, $I(0)$, R_g , and q are intensity as a function of the scattering vector, intensity at zero scattering angle, radius of gyration, and the scattering vector, respectively. Guinier analysis was performed with an upper limit defined by $q \cdot R_g = 1.3$ for a globular system.

Kratky plot

To ascertain the degree of folding of the macromolecule, a Kratky plot is utilized.⁷⁴ In the Kratky plot $q^2 I(q)$ vs q is plotted.⁷⁵ This plot provides information about sample flexibility, where unfolded samples are characterized by a plateau at high q , while globular proteins are characterized by a bell-shape. A Kratky plot can be expressed as

$$\frac{q^2 I(q)}{I(0)} \text{ vs } q$$

which normalizes the scattering profiles by mass and concentration by dividing by $I(0)$.⁶ Additionally, the Kratky plot can be made dimensionless by multiplying the q -vector by R_g and multiplying $I(q)$ by $(q R_g)$

$$\frac{(q R_g)^2 I(q)}{I(0)} \text{ vs } q R_g$$

which allows comparison of particles with different mass and conformation.⁷⁷

The R_g was determined in addition to the molecular weight using the volume of correlation based molecular weight calculation, V_c ,⁶⁸ which uses $I(0)$ and R_g and is independent of protein concentration.

The volume of correlation, V_c , is defined as

$$V_c = \frac{I(0)}{\int_{\infty}^0 q I(q) dq}$$

where $I(q)$, $I(0)$, and q are intensity as a function of the scattering vector, intensity at zero scattering angle, and the scattering vector, respectively.⁶⁸ This integral can be used to determine the molecular weight as

$$M_w = \left(\frac{V_c^2}{\frac{R_g}{c}} \right)^k$$

upon convergence in which $\frac{V_c^2}{R_g}$ is logarithmically proportional to the molecular weight, M_w . c and k are sample dependent constants, in the case of proteins, $c = 0.1231$ and $k = 1$.

Pair distribution function

A pair distribution or $P(r)$ function was used to determine the maximum dimension of the particles, D_{\max} , and evaluate the shape of the particle.^{78,79} The $P(r)$ function provides information of the distances between all atom pairs in a molecule and is defined as follows:

$$P(r) = \frac{r}{2\pi} \int_0^{\infty} I(q)q \sin(qr) \cdot dq$$

in which r is the radius between scattering atoms in Å. Solution of this equation is performed by inverse Fourier transform (IFT), as implemented in GNOM.⁶⁹ The $P(r)$ function provides qualitative information about R_g and $I(0)$ and comparison with the R_g and $I(0)$ obtained from the Guinier fit can be made to ascertain reliability of the result. D_{\max} is determined $P(r)$ plot of

$$\frac{P(r)}{I(0)} \text{ vs } r$$

in which the $P(r)$ function should approach zero gradually at the appropriate D_{\max} . Additionally, the $P(r)$ function varies according to particle shape, with a globular protein (approximate spherical shape) displaying a bell-shaped curve.⁸⁰

4.4.4 Effect of camphor concentration on CYP450cam

Previous studies of CYP450cam at in the presence and absence of camphor using SAXS had not yielded any significant changes in conformation.⁵⁷ These studies were performed at 0 μM and 500 μM camphor. This concentration of camphor however far exceeds the binding constant of camphor at $K_D \sim 1.6 \mu\text{M}$.⁸¹⁻⁸³ A range of camphor concentrations including these lower levels of camphor were therefore employed. Additionally, given the presence of a postulated second binding site on CYP450cam with a $K_D \sim 43 \mu\text{M}$,⁵⁴ SAXS data was additionally collected at this concentration. Data was collected with a q_{\min} of 0.008 \AA^{-1} and a q_{\max} of 0.696 \AA^{-1}

Initial analysis of CYP450cam(C334A) at a range of concentrations was performed to determine linearity of the data to ensure monodispersity of the system. This ensures that the protein molecules behave as free particles in solution (i.e., without interparticle interferences). The Guinier plot for CYP450cam(C334A) maintained linearity throughout the 1 mg/mL – 6 mg/mL concentration range, indicating an absence of aggregation or interparticle interactions. Characterization of CYP450cam(Crosslink) at 1000 μM camphor indicated a typical small angle scattering profile comprised of a plateau region at low- q which obeys the laws of Guinier, a sloping mid- q range sensitive to the shape of the CYP450 and a steep high- q area dependent on the interface between CYP450cam and the surrounding solvent environment (Figure 4.12 A). A Kratky plot indicated a bell-shaped curve indicating a globular shape, however a small shoulder at high- q suggests a degree of flexibility (Figure 4.12 B). Guinier analysis was performed and the Guinier plot was linear at low- q indicating an absence of aggregation or interparticle repulsion and a calculated R_g of $23.85 \text{ \AA} \pm 0.03 \text{ \AA}$ (Figure 4.12 C).

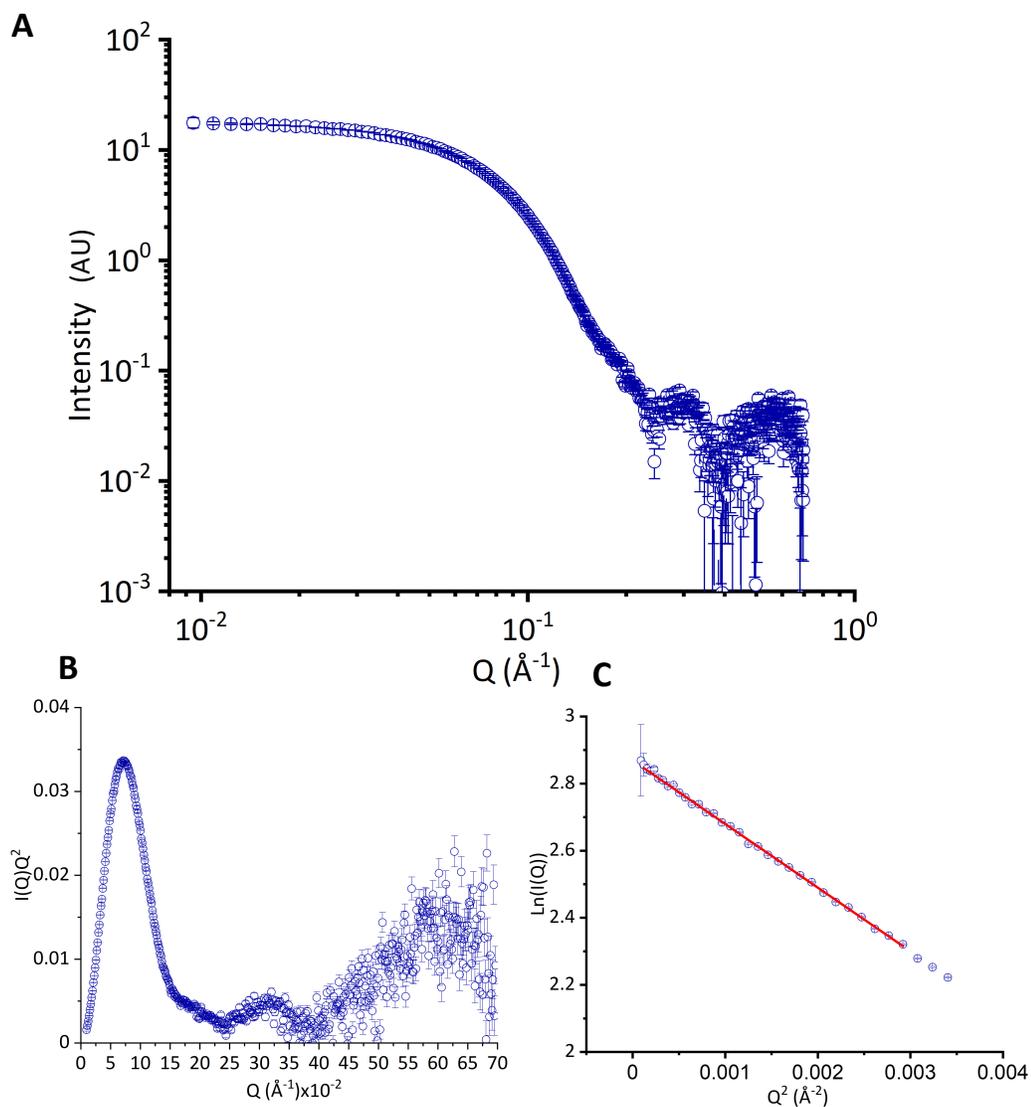


Figure 4.12 SAXS analysis of CYP450cam at 1000 μM camphor. **(A)** SAXS curve of intensity of scattering over a q -range. **(B)** Kratky plot indicating a globular protein with some flexibility. **(C)** Guinier analysis plot to determine radius of gyration (R_g).

Analysis of the CYP450cam at different camphor concentrations indicated that there is a change of R_g dependent on camphor concentration. Initial studies by Lewis *et al.* concluded that the CYP450cam conformation and dynamics were independent of camphor, citing an R_g of $24.14 \text{\AA} \pm 0.20 \text{\AA}$ and $23.93 \text{\AA} \pm 0.21 \text{\AA}$ for the camphor-free and bound forms, respectively.⁵⁷ These studies were performed at 0 μM camphor and at an excess of 500 μM camphor. The R_g for the collected data is $24.31 \text{\AA} \pm 0.03 \text{\AA}$ at 0 μM camphor which is in good agreement with the previous studies. SAXS data was collected at a range of concentrations from 0 μM to 1000 μM camphor to provide a comprehensive picture of camphor-dependent

CYP450cam conformational changes (Table 4.1 and Figure 4.13). In contrast to the study by Lewis *et al.*, we observe a marked decrease in R_g upon substrate addition. At 1 μM camphor there is a decrease in the R_g to $23.01 \text{ \AA} \pm 0.05 \text{ \AA}$. In addition to this, there is a decrease in the D_{max} from 81 \AA to 68 \AA suggesting compaction of the structure as can be seen from the $P(r)$ function (Figure 4.14 A).

Table 4.1 Size analysis of CYP450cam(Crosslink) as a function of camphor concentration.

Camphor (μM)	R_g (\AA)	D_{max} (\AA)	M_w (kDa)
0	24.31 ± 0.03	81.0	47.0
1	23.01 ± 0.05	68.0	44.0
1.6	23.55 ± 0.06	72.0	45.4
2	23.78 ± 0.03	76.0	45.9
10	23.75 ± 0.03	75.0	46.0
20	23.48 ± 0.03	72.0	45.7
43	24.10 ± 0.03	80.0	46.5
60	23.67 ± 0.03	74.0	46.0
100	23.60 ± 0.03	75.0	45.5
1000	23.85 ± 0.03	78.0	46.1

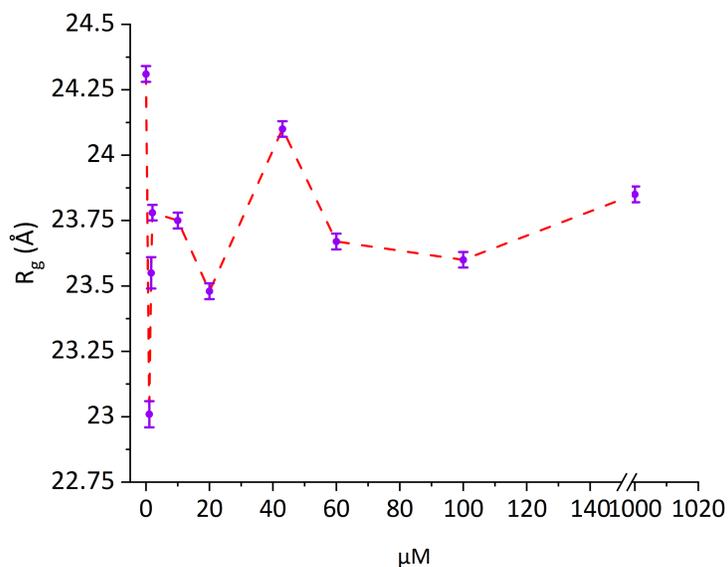


Figure 4.13 The radius of gyration (R_g) of CYP450cam as a function of camphor concentration.

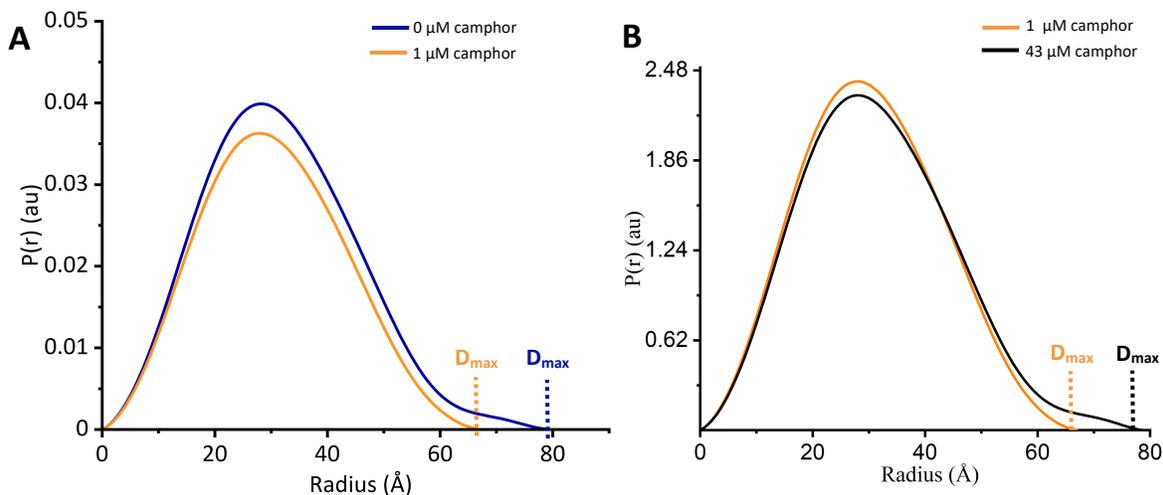


Figure 4.14 Pair distribution function $P(r)$ of CYP450cam. **(A)** $P(r)$ with 0 μM and 1 μM camphor. **(B)** $P(r)$ with 1 μM and 43 μM camphor.

The R_g remains below the initial R_g at 0 μM camphor at 1.60 μM , 2 μM , 10 μM and 20 μM camphor (Figure 4.13). This drop in R_g supports the structural transition from CYP450cam in an open form in the absence of camphor which transitions to a closed form in the presence of camphor.²³ This transition is characterized by closure of the substrate access channel formed by the F- and G-helices and ordering of the B'-helix when camphor binds. These structural changes observed in the X-ray crystal structure may play a role in the solution structure studied here by SAXS.

A further observation was an increase in the R_g at 43 μM camphor to $24.10 \text{ \AA} \pm 0.03 \text{ \AA}$ and an increase of the D_{max} to 80 \AA as can be seen from the $P(r)$ function (Figure 4.14 B). This concentration is significant because NMR studies have proposed that a second molecule binds to CYP450cam at a 43 μM concentration. Furthermore, molecular dynamics (MD) and paramagnetic nuclear magnetic resonance (NMR) studies have confirmed the presence of an allosteric site on camphor.^{50,56} Upon camphor binding to this site, it is postulated that CYP450cam adopts an open conformation which primes it for redox partner binding as well as formation of a channel for product egress termed Channel 2. This site is delimited by the residues Lys126, Leu166, Gly168, Thr217 and Ala219.⁵⁰ The SAXS data presented here suggest that at 43 μM camphor there is indeed a conformational change that occurs resulting in an increased R_g .

Interestingly, the increase in R_g at 43 μM camphor is followed by a decrease in R_g and D_{max} at higher camphor concentrations (Figure 4.13). This may initially seem counter-intuitive since at higher camphor concentrations, the allosteric site should remain bound. The dependence of the active site heme iron spin state on camphor binding may however play a role in CYP450cam catalysis since upon camphor binding, there is a heme iron spin state transition from ferric low spin to high spin. This is accompanied by a change in the reduction potential from -303 mV to -173 mV to facilitate reduction by the redox partner protein Pdx.^{18,19} Another spin state transition is however observed from high-spin back to low spin at high concentrations of camphor. This was first observed by Lange *et al.* in 1977 where changes in the absorbance spectrum indicated that from more than 0 μM to approximately 100 μM camphor, the heme iron is in a high spin state.⁸⁴ Above 100 μM camphor, however, this shifts back to a low-spin state. This spin transition was confirmed by electron paramagnetic resonance (EPR) which concluded that camphor binding alters the active site electronic environment.⁴⁵ Analyses of CYP450cam spin states at high and low pressure indicated a similar transition to low spin at high pressures.⁸⁵ The spectral changes were initially related to binding modes by Schenkman in 1970 in liver microsomes as Type I and Type II binding modes.⁸⁶ Type I binding in CYP450cam is characterized by camphor ligand binding to a hydrophobic pocket in the active site which is accompanied by the characteristic blue shift in the Soret band which indicates a shift from low spin to high spin.^{18,19} In this case, water is displaced, and the heme iron is pentacoordinate. When Type II binding occurs, there is a sixth ligand bound to the heme iron resulting in a low-spin heme iron.⁸⁷ Given that that transition is observed at high concentrations, there is a possibility that high levels of camphor result in a similar hexacoordinate, low spin species. The high levels of camphor may saturate the active site causing camphor to coordinate the heme iron in the resting state.

The correlation of the decrease in R_g at camphor concentrations greater than 60 μM may therefore be related to changes in conformation that occur at high levels of camphor which affect not only the electronic state of the heme active site, but also the conformational dynamics of the rest of the CYP450cam structure.

Analysis of the R_g of the CYP450cam(L358P) mutant as a function of camphor concentration, does not indicate a significant change of the protein conformation with an R_g of ~ 23.5 Å to ~ 23.8 Å across the camphor concentrations (Table 4.2, Figure 4.15). The D_{max} also remains at ~ 73.0 Å at all camphor concentrations as can be seen from the $P(r)$ function (Figure 4.16).

Table 4.2 Size analysis of CYP450cam(Crosslink) as a function of camphor concentration.

Camphor (μM)	P450cam(Crosslink)			P450cam(L358P)			P450cam-Pdx complex		
	R_g (Å)	D_{max} (Å)	M_w (kDa)	R_g (Å)	D_{max} (Å)	M_w (kDa)	R_g (Å)	D_{max} (Å)	M_w (kDa)
0	24.31 ± 0.03	81.0	47.0	23.48 ± 0.05	71.0	43.8	23.9 ± 0.04	66.0	51.9
1.6	23.55 ± 0.06	72.0	45.4	23.76 ± 0.04	75.0	43.8	23.94 ± 0.03	64.0	51.6
43	24.10 ± 0.03	80.0	46.5	23.68 ± 0.03	71.0	43.6	25.11 ± 0.03	74.0	52.6
1000	23.85 ± 0.03	78.0	46.1	23.66 ± 0.03	73.0	43.2	23.53 ± 0.03	65.0	50.6

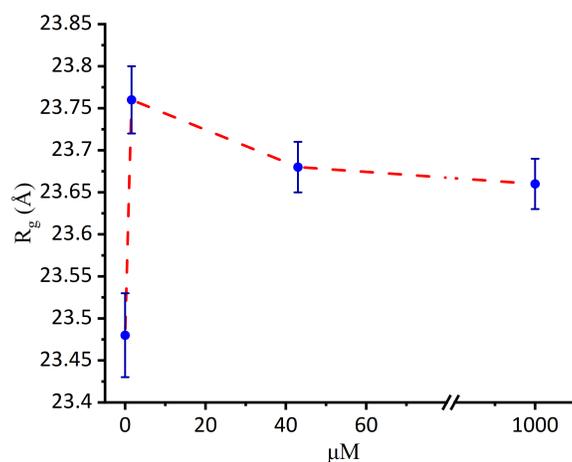


Figure 4.15 The radius of gyration (R_g) of CYP450cam(L358P) as a function of camphor concentration.

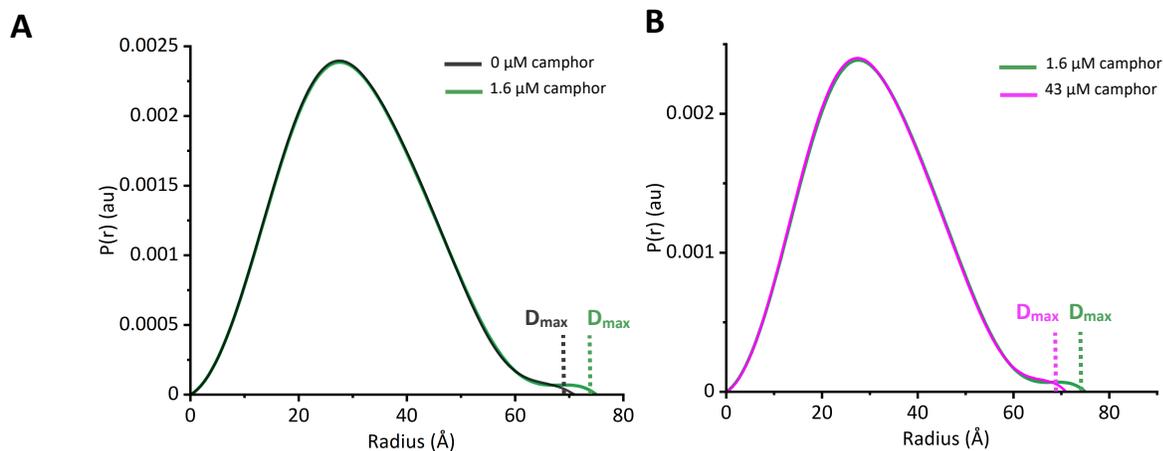


Figure 4.16 Pair distribution function $P(r)$ of CYP450cam(L358P). **(A)** $P(r)$ with 0 μM and 1.6 μM camphor. **(B)** $P(r)$ with 1.6 μM and 43 μM camphor.

The Leu358Pro mutation induces several structural changes similar to the conformational changes caused by Pdx docking.⁴⁹ This results in a CYP450cam with an open conformation in which the introduced Pro358 pushes against the heme, opening up the active site. This opening of the active site is predominantly a steric effect, which may explain why the observed R_g is independent on camphor concentration. CYP450cam(L358P) is locked in an open conformation with the steric changes obstructing any further conformational changes.

Comparison of the SAXS profile of the free CYP450cam and the CYP450cam-Pdx complex indicates a slight change in the profile at mid- q indicating a difference in shape from the uncomplexed form (Figure 4.17 A). Additionally, the Kratky plot has a more pronounced shoulder indicating an additional domain and potentially increased flexibility (Figure 4.17 B). The Kratky plot of the CYP450cam-Pdx complex has a similar shape to the free CYP450cam, indicating that there is little change in the overall globular shape.

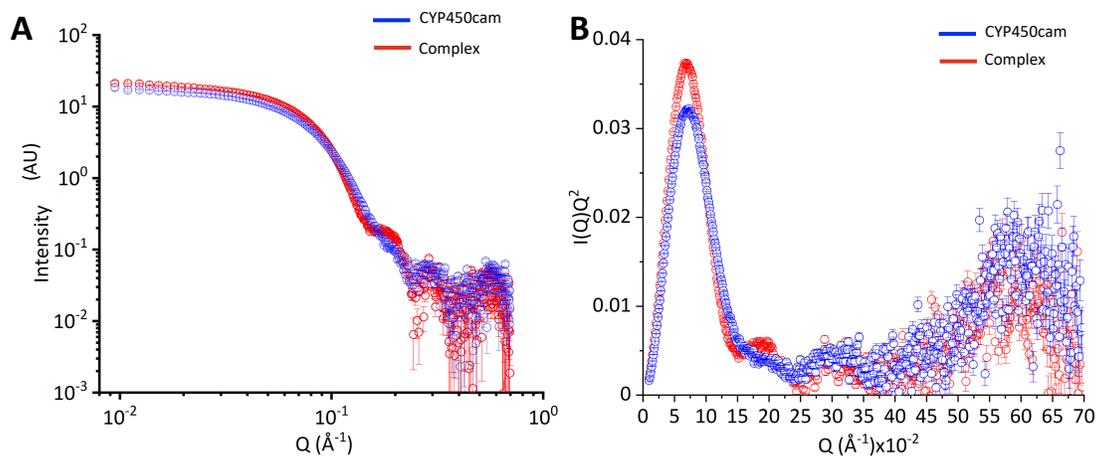


Figure 4.17 Comparison SAXS data from CYP450cam and the CYP450cam-Pdx complex. **(A)** SAXS curve of intensity of scattering over the q -range. **(B)** Kratky plot indicating a slight change in flexibility.

SAXS analysis of the CYP450cam-Pdx complex indicates a particle with an R_g of $23.9 \pm 0.04 \text{ \AA}$ at 0 \mu M camphor, not significantly larger than the R_g of the CYP450cam(Crosslink) mutant (Table 4.2, Figure 4.18). This suggests that upon complexation with Pdx, the CYP450cam may undergo structural rearrangements and compaction to accommodate Pdx in which there is not a large increase in R_g . The D_{\max} determined is smaller than that of the uncomplexed CYP450cam. Further modelling using *ab initio* modeling methods are required to determine the conformation that has resulted in this decrease in D_{\max} .

There is not a significant change in R_g upon addition of camphor at the 1.6 \mu M with an R_g of $23.9 \text{ \AA} \pm 0.04 \text{ \AA}$ at 0 \mu M and $23.94 \text{ \AA} \pm 0.04 \text{ \AA}$ at 1.6 \mu M camphor (Figure 4.18). The D_{\max} does also not significantly change as can be seen from the $P(r)$ function (Figure 4.19 A). This confirms that binding of Pdx to CYP450cam results in a closed to open conformational change and that this open conformation is stabilized even in the presence of substrate. Similar to free CYP450cam, the CYP450cam-Pdx covalent complex indicates an increased R_g from $23.94 \text{ \AA} \pm 0.03 \text{ \AA}$ at 1.6 \mu M camphor to $25.11 \text{ \AA} \pm 0.03 \text{ \AA}$ at 43 \mu M camphor, the concentration at which camphor is believed to bind to the allosteric binding site resulting in the more open conformation (Figure 4.18). This is accompanied by an observed increase in the D_{\max} from 64.0 \AA to 74.0 \AA as seen from the $P(r)$ function (Figure 4.19 B). The allosteric site is occupied at higher camphor concentrations ($\sim 43 \text{ \mu M}$) and it has been found that camphor remains stably bound to the allosteric

site in the presence of Pdx.⁵⁶ These data indicate that allosteric binding of camphor can result in CYP450cam conformational changes when in complex with Pdx.

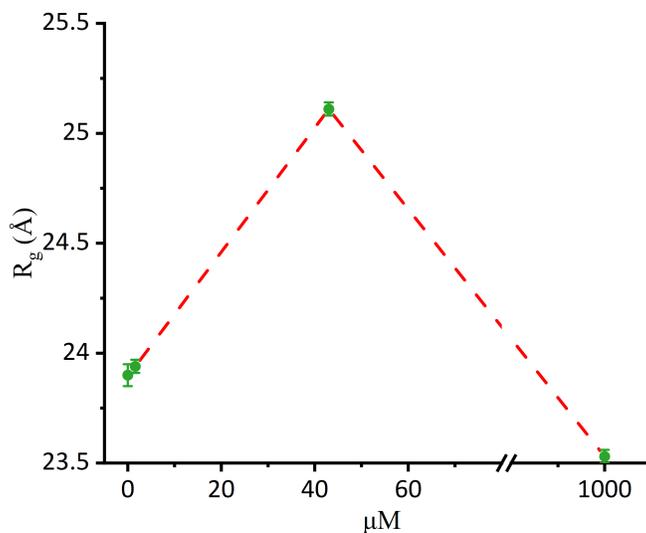


Figure 4.18 The radius of gyration (R_g) of the CYP450cam-Pdx complex as a function of camphor concentration.

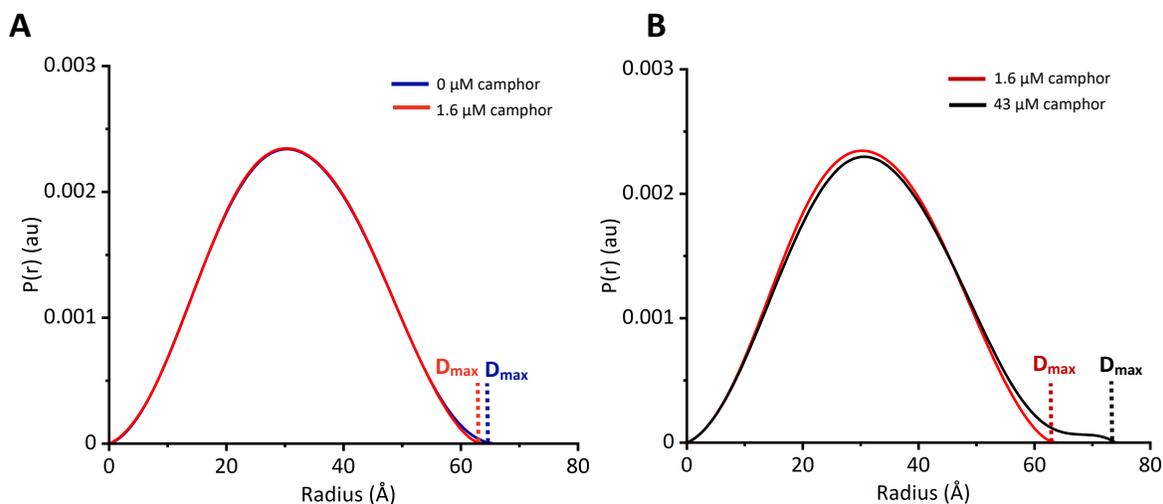


Figure 4.19 Pair distribution function $P(r)$ of the CYP450cam-Pdx complex. **(A)** $P(r)$ with 0 μM and 1.6 μM camphor. **(B)** $P(r)$ with 1.6 μM and 43 μM camphor.

4.4.5 FOXS crystal structure and data comparison

In order to determine how well the experimentally obtained data fit a SAXS curve generated from the CYP450cam protein atomic coordinates, analysis was performed using FOXS (Fast X-ray Scattering).

^{70,71} FOXS is a web-based server that rapidly computes a SAXS profile when a molecular structure is provided. FOXS utilizes rigid body modelling to compute the theoretical SAXS profile. A profile is calculated using the Debye formula to predict the intensity $I(q)$ as a function of the form factor $f_i(q)$. The calculated profile is subsequently fitted to the experimental profile and the quality of fit, χ^2 is determined.

Analysis of the experimentally obtained R_g was larger than the calculated R_g from the crystal structure for all the profiles (Table 4.3). Given that the experimental R_g determined for CYP450cam was similar to that determined in early SAXS experiments, it can be concluded that crystal packing has a significant effect on the conformational flexibility of CYP450cam. Indeed, overlay of the experimental and theoretical plots indicate high χ^2 during fitting and a mismatch in the mid- q range that is sensitive to the shape of the particle (Figure 4.20). This suggest that the shape of CYP450cam is different in solution that that which is observed in the more rigid crystal packing environment.

Table 4.3 Comparison of the experimental and calculated dimensions using the FOXS server.

Camphor (μ M)	P450cam(Crosslink)			P450cam(L358P)			P450cam-Pdx complex		
	R_g^{exp} (\AA)	R_g^{calc} (\AA) (PDB 3L61 - camphor free, 2CPP - camphor bound)	χ^2	R_g^{exp} (\AA)	R_g^{calc} (\AA) (PDB, 1T86 - camphor bound)	χ^2	R_g^{exp} (\AA)	R_g^{calc} (\AA) (PDB 4JWS - camphor free, 4JX1 - camphor bound)	χ^2
0	24.31 ± 0.03	21.50	53.78	23.48 ± 0.05	21.12	34.99	23.9 ± 0.04	23.27	42.20
1	23.01 ± 0.05	21.18	21.28						
1.6	23.55 ± 0.06	21.18	34.06	23.76 ± 0.04	21.12	31.26	23.94 ± 0.03	22.93	13.40
43	24.10 ± 0.03	21.18	62.10	23.68 ± 0.03	21.12	40.01	25.11 ± 0.03	22.93	16.98
1000	23.85 ± 0.03	21.18	56.02	23.66 ± 0.03	21.12	37.83	23.53 ± 0.03	22.93	21.54

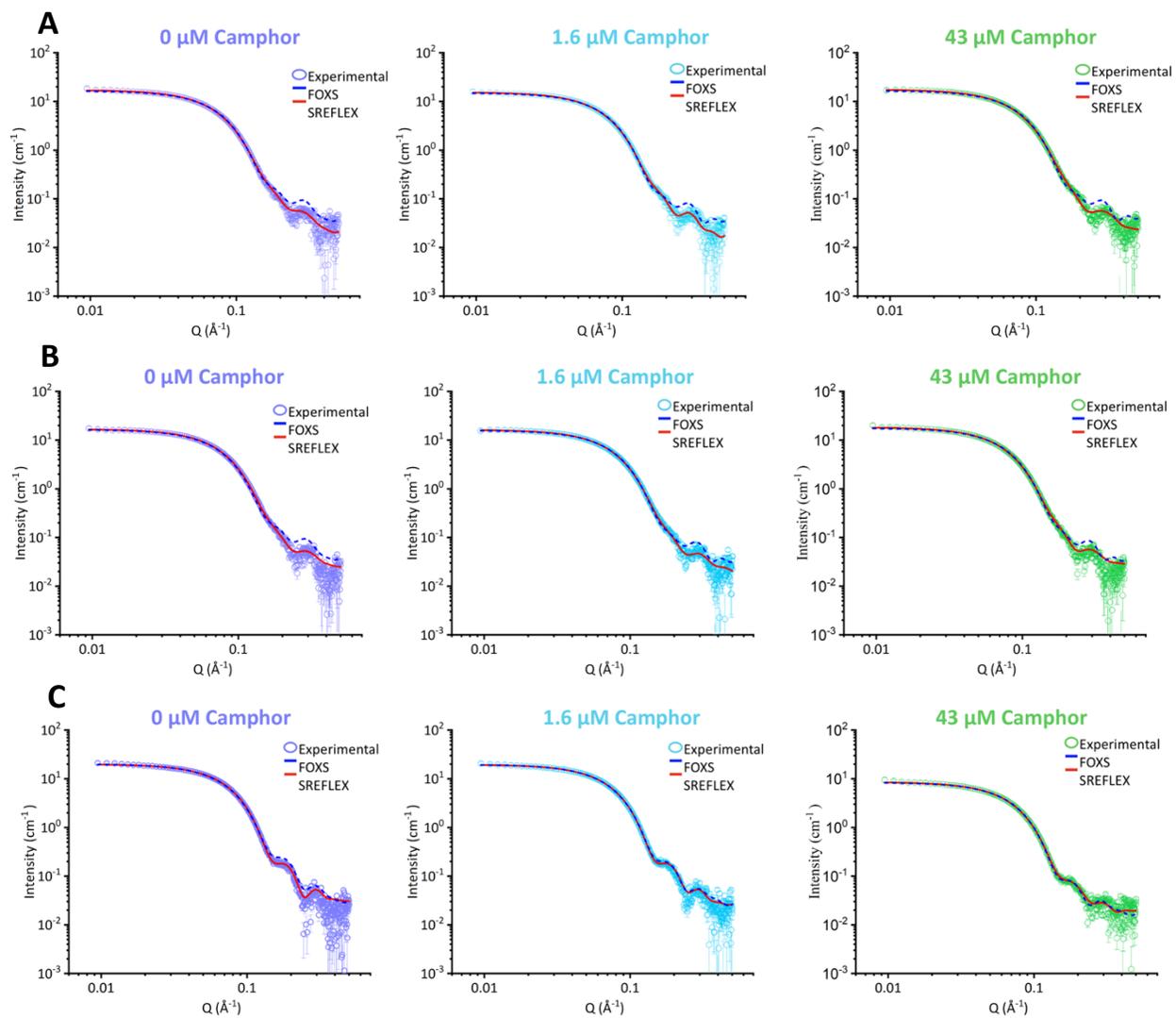


Figure 4.20 Overlay of calculated and experimental SAXS curves determined by FOXS and SREFLEX for (A) CYP450cam, (B) CYP450cam(L358P) and (C) CYP450cam-Pdx(Complex) at 0 μM, 1.6 μM and 43 μM camphor.

4.4.6 SREFLEX conformational distribution calculations

To establish the regions of the CYP450cam structure that may be contributing to the observed increased R_g , conformational analysis was performed using SREFLEX. SREFLEX represents a modeling approach that explores the conformational space of atomic macromolecular models using normal mode analysis (NMA) and refines the obtained conformations guided by experimentally obtained SAXS data.⁷² In order to model the structural flexibility, the input model (generally high-resolution crystallographic model) is partitioned into pseudo-domains after which NMA is used to probe large and subsequently smaller movements in a hierarchical approach. NMA permits a coarse-grained approach to study the conformational changes since SAXS data provides information at a lower resolution than X-ray crystallography.^{88,89} The work flow of SREFLEX can be summarized as follows: (1) A high resolution input structure is used as the starting model (2) Conformational changes are predicted by NMA (3) Domain assignments are performed to reduce computational search space (4) Guided by experimental SAXS data, conformations are hierarchically sampled (5) A full-atom model is generated and the fit is evaluated using χ^2 values.⁷²

Evaluation of conformational flexibility was performed using the SREFLEX web-based server (http://www.embl-hamburg.de/bio_saxs/online.html). As the starting model, PDBs of the conditions of interest were utilized. The PDB structures used for the atomistic theoretical models included 3L61 – camphor free, 2CPP – camphor bound, 1T86 – L358P mutant camphor bound, 4JWS – CYP450cam-Pdx complex camphor free and 4JX1 – CYP450cam-Pdx complex camphor bound. The χ^2 were lower than those observed for in comparison of the experimentally and theoretically generated curved in FOXS, indicating that conformational flexibility does contribute to the SAXS scattering observed in solution in comparison to the X-ray structure (Table 4.4, Figure 4.20).

Table 4.4 Comparison of the experimental and calculated dimensions using SREFLEX normal mode analysis.

Camphor (μM)	P450cam(Crosslink)				P450cam(L358P)				P450cam-Pdx complex			
	R_g^{exp} (\AA)	R_g^{calc} (\AA) (PDB 3L61 – camphor free, 2CPP – camphor bound)	χ^2	RMSD	R_g^{exp} (\AA)	R_g^{calc} (\AA) (PDB, 1T86 – camphor bound)	χ^2	RMSD	R_g^{exp} (\AA)	R_g^{calc} (\AA) (PDB 4JWS – camphor free, 4JX1 – camphor bound)	χ^2	RMSD
0	24.31 ± 0.03	24.02	7.89	3.25	23.48 ± 0.05	23.40	8.49	5.58	23.9 ± 0.04	24.09	12.69	5.50
1.6	23.55 ± 0.06	23.58	6.39	3.59	23.76 ± 0.04	23.31	6.76	5.34	23.94 ± 0.03	24.15	10.42	2.92
43	24.10 ± 0.03	24.01	7.77	4.87	23.68 ± 0.03	23.34	11.53	5.41	25.11 ± 0.03	24.70	12.69	8.81

Structural overlay of the structures generated from modelling using NMA in SFREFLEX and the crystal structure revealed specific regions that displayed and altered conformation. In particular, the region defined by residues 114-150 and 200-250 displayed positions different from that observed in the crystal structure (Table 4.5). In particular, the $\beta 2$ -sheet which connects the H- and I-helices as well as the H-helix, G-helix, D-helix and C-helix displayed increased mobility when compared to the crystal structure (Figure 4.21). This region is adjacent to the allosteric site that has been proposed by Follmer *et al.*,^{50,56} where CYP450cam is thought to bind a second molecule of camphor.⁵⁴ The allosteric site is defined by the residues Val118, Val123, Lys126, Leu127, Leu166, Gly168, Thr217 and Ala219.

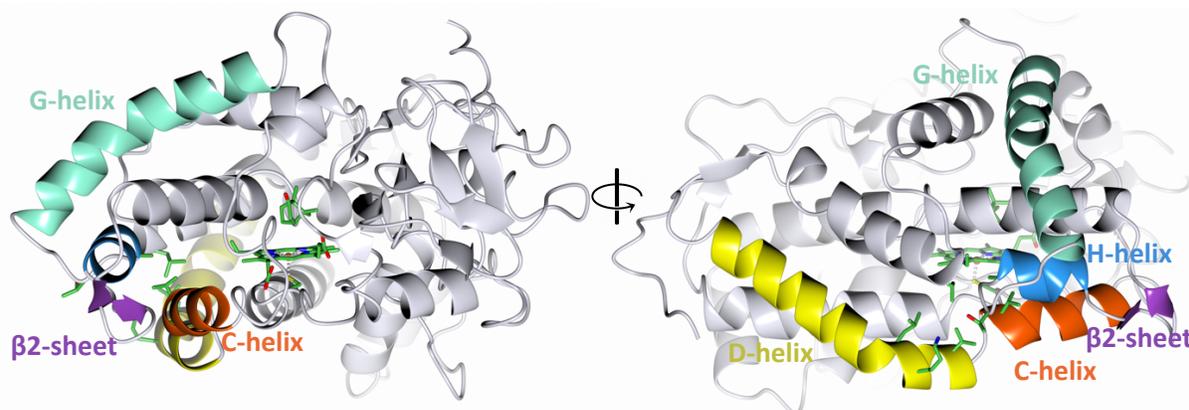


Figure 4.21 Structural elements in CYP450cam (PDB 1DZ4) with increased mobility found by SREFLEX normal mode analysis.

Structural alignment of the crystal structure with the calculated structure indicates a consistency with increased mobility near this allosteric site, with the β 2-sheet being the prevalent element that displays movement for CYP450cam, CYP350cam(L358P) and the complex (Figure 4.22 – 4.24). Calculation of the RMSD of a structural alignment indicates a 10 – 20 Å displacement for these regions. Modeling of the CYP450cam-Pdx complex, however, indicated some inconsistencies. The NMA simulations indicated movement of Pdx that is not feasible for the 0 μ M camphor and 43 μ M camphor SAXS curves. The extent of Pdx displacement is too large to ensure correct electron transfer. Future SREFLEX studies in which improved restraints of the Pdx relative to CYP450cam are implemented will be necessary to reflect productive binding in the complex. While these initial simulations offer a coarse-grained window into the increased mobility that CYP450cam has in solution, it is highly informative to note that the region predicted by SAXS to be mobile is the same region that exerts allosteric control from a closed to an open conformation.

Table 4.5 Structural elements with increased mobility after introduction of structural flexibility using SREFLEX normal mode analysis.

SAXS Curve	Input PDB	Residues with increased mobility	Structural element with increased mobility
0 μ M Camphor CYP450cam(Crosslink)	3L61	A224 – T234	β 2-sheet
1.6 μ M Camphor CYP450cam(Crosslink)	2CPP	E209 – K214 D218 – N225 A224 – T234	G-helix H-helix β 2-sheet
43 μ M Camphor CYP450cam(Crosslink)	2CPP	E198 – K214 D218 – N225 A224 – T234	G-helix H-helix β 2-sheet
0 μ M Camphor CYP450cam(L358P)	1T86	E209 – K214 D218 – N225 A224 – T234	G-helix H-helix β 2-sheet
1.6 μ M Camphor CYP450cam(L358P)	1T86	E209 – K214 D218 – N225 A224 – T234	G-helix H-helix β 2-sheet
43 μ M Camphor CYP450cam(L358P)	1T86	E209 – K214 D218 – N225 A224 – T234	G-helix H-helix β 2-sheet
0 μ M Camphor CYP450cam-Pdx Complex	4JWS	L114 – K126 E192 – K214 D218 – N225 A224 – T234	C-helix G-helix H-helix β 2-sheet
1.6 μ M Camphor CYP450cam-Pdx Complex	4JX1	A224 – T234	β 2-sheet
43 μ M Camphor CYP450cam-Pdx Complex	4JX1	R130 – Q145 A224 – T234	D-helix β 2-sheet

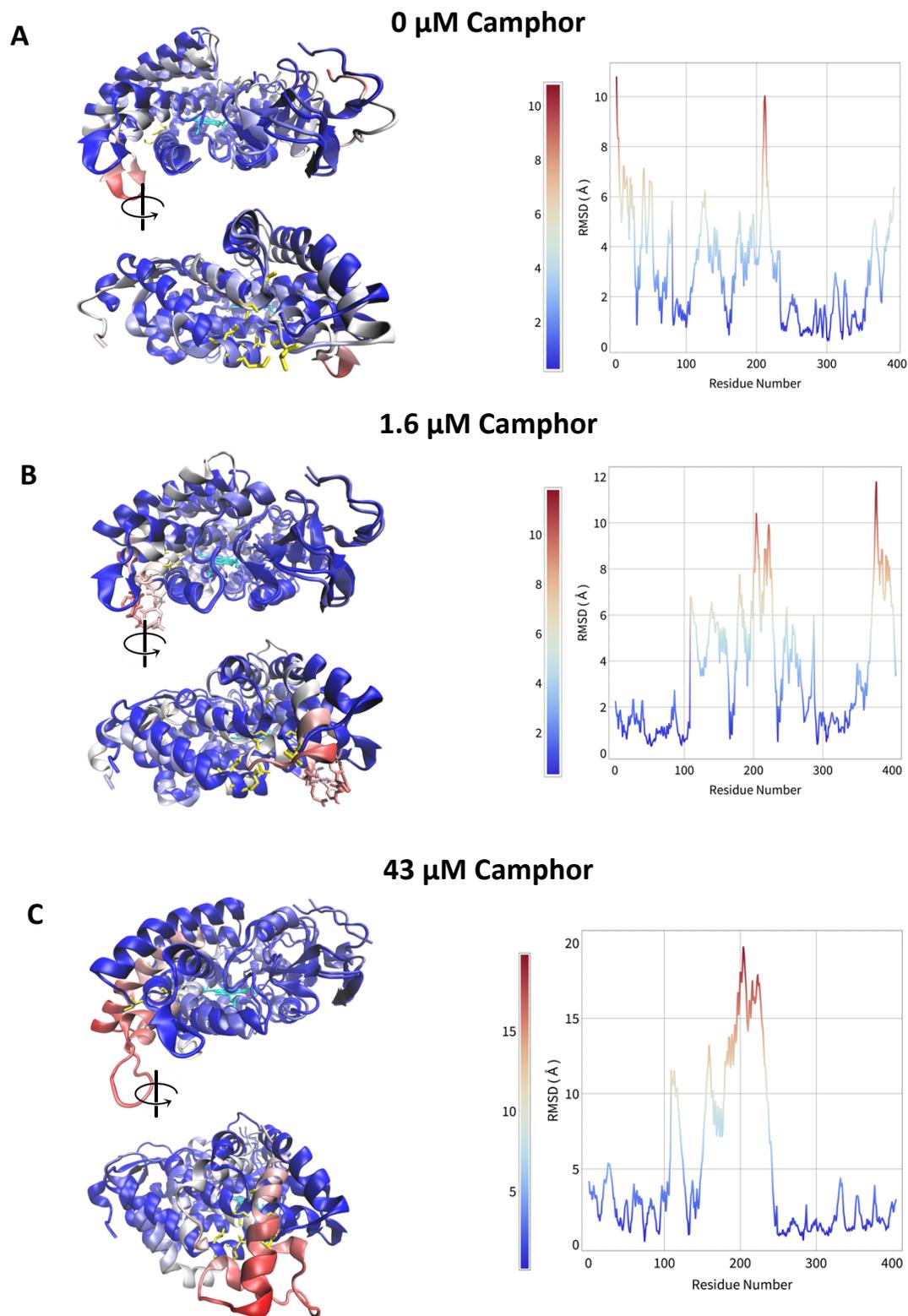


Figure 4.22 Structural alignment of SREFLEX model with original CYP450cam crystal structure. RMSD per residue on the right panel for each comparison. Alignment performed with PDB (A) 3L61 (B) 2CPP (C) 2CPP.

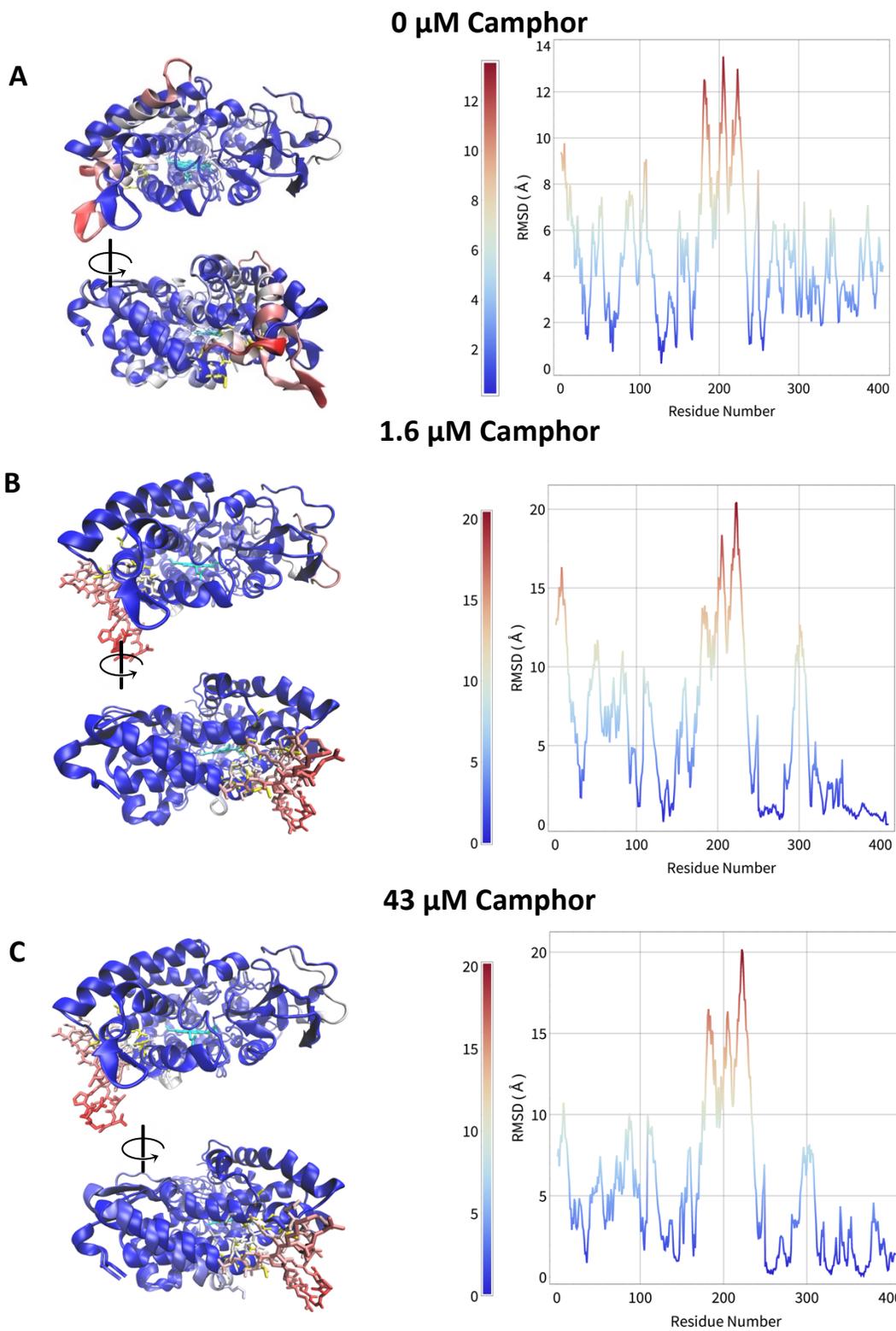


Figure 4.23 Structural alignment of SREFLEX model with original CYP450cam(L358P) crystal structure. RMSD per residue on the right panel for each comparison. Alignment performed with PDB (A) 1T86 (B) 1T86 (C) 1T86.

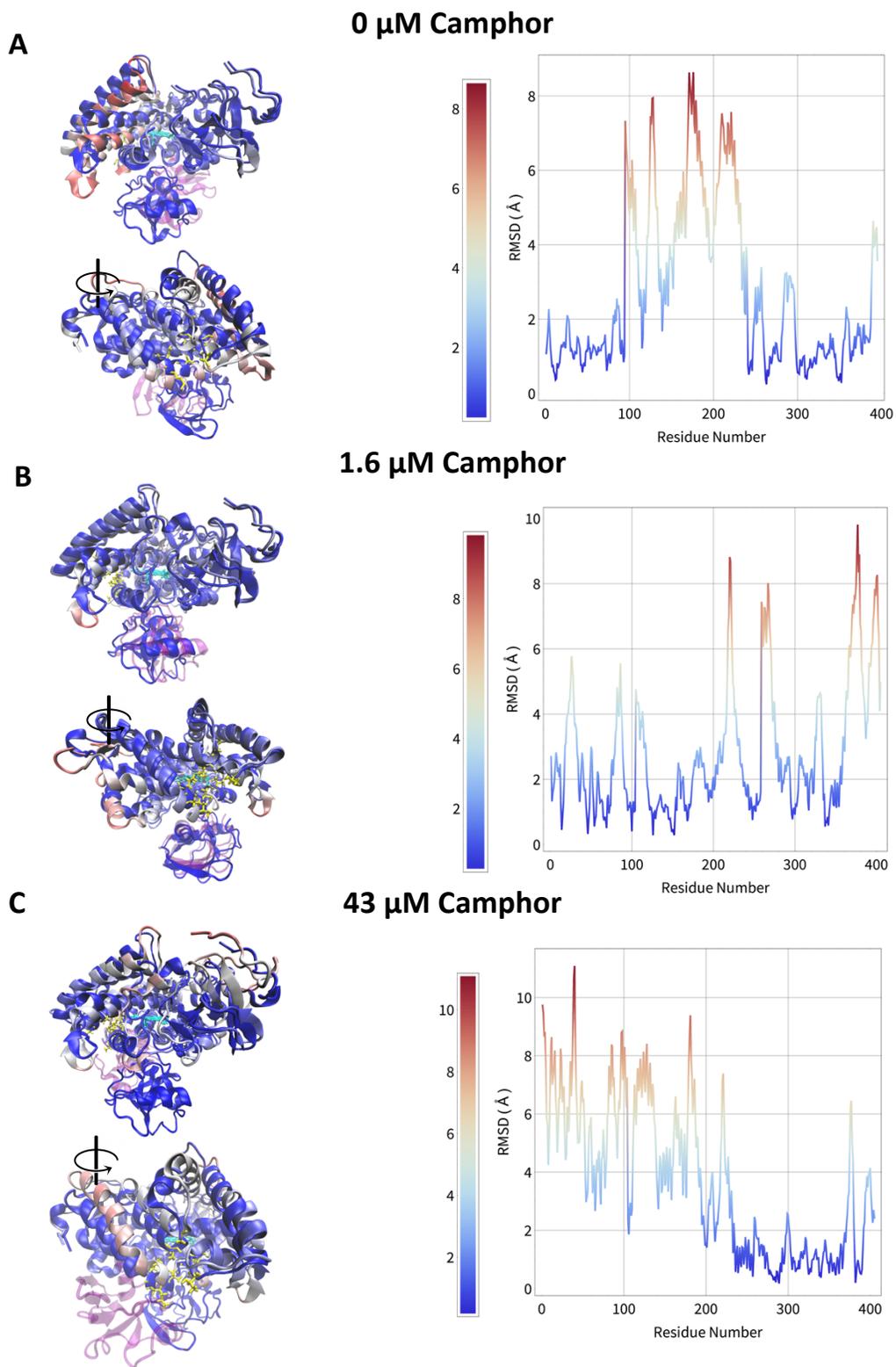


Figure 4.24 Structural alignment of SREFLEX model with original CYP450cam-Pdx complex crystal structure. RMSD per residue on the right panel for each comparison. Alignment performed with PDB (A) 4JWS (B) 4JX1 (C) 4JX1.

4.5 Conclusions

While the structural information gained from SAXS is low resolution, it provides significant information about the conformational flexibility of a macromolecule in solution.⁷⁷ X-ray crystallography is a high-resolution technique that provides atomistic information about the protein structure, however the inherent static environment present in protein crystals provides an incomplete picture.¹ Proteins are highly dynamic molecules, traversing the energy landscape continuously to achieve efficient catalysis. CYP450cam is no exception, and spectroscopy and solution NMR have revealed a dynamic system in which structural perturbations ensure efficient catalysis.² Interaction of CYP450cam with camphor has revealed several layers of control, from an open conformation, to a closed conformation upon substrate binding which shifts to an open conformation in the presence of high levels of camphor due to allosteric regulation. These conformational changes are coupled to delicate tuning of the electronic properties of the active site heme iron. It stands to reason that while experimentally, conclusions may be drawn on the iron spin state and redox potential, the supporting architecture of the protein structure contributes to these transitions. Early SAXS data concluded that “It is clear that if cytochrome P-450 undergoes a conformational change upon binding camphor, it is smaller than that which we can accurately detect using solution-scattering methods. Thus the structural changes inferred from temperature jump spin relaxation and pressure studies most likely involve only local rearrangements near the heme.”⁵⁷ The concentration of substrate utilized however, far exceeded the physiological levels that may exert conformational regulation. SAXS is a probe that permits interrogation of a protein structure in solution under native conditions. The results here have confirmed X-ray crystallography findings that CYP450cam undergoes an open to closed transition upon substrate binding. Furthermore, the recent evidence for an allosteric site in camphor is supported by opening of the structure. The apparent transition observed at high camphor concentrations by a decrease in R_g illustrates the artefacts that can be observed in the artificially generated environment of an experiment. In experimental conditions, excess camphor can be added which can lead to a “crowded” active site with camphor resulting in a Type II low spin hexacoordinated species. Modelling of the SAXS data with theoretical SAXS curves indicates a discrepancy in R_g , with the R_g for the solution data being larger than

that of the crystallographic structure. This suggests that CYP450cam is more flexible in solution than in the packed environment of the crystal lattice. Initial conformational modeling has attributed this conformational flexibility to the β 2-sheet region in proximity to the allosteric site proposed. The data presented here provide a perspective on the increased conformational flexibility of CYP450cam. This flexibility plays an important role in catalysis, with substrate and redox partner proteins affecting nuanced regulation on CYP450cam chemistry.

4.6 References

1. Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964–972 (2007).
2. Pochapsky, T. C. & Pochapsky, S. S. What Your Crystal Structure Will Not Tell You about Enzyme Function. *Acc. Chem. Res.* (2019). doi:10.1021/acs.accounts.9b00066
3. Orozco, M. A theoretical view of protein dynamics. *Chem. Soc. Rev.* **43**, 5051–5066 (2014).
4. Ortiz de Montellano, P. R. *Cytochrome P450*. Springer (Springer International Publishing, 2015). doi:10.1007/978-3-319-12108-6
5. Poulos, T. L. Heme Enzyme Structure and Function. *Chem. Rev.* **114**, 3919–3962 (2014).
6. Nelson, D. R. *et al.* CYP450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* **6**, 1–42 (1996).
7. Gunsalus, I. C. & Wagner, G. C. Bacterial P-450cam methylene monooxygenase components: Cytochrome m, putidaredoxin, and putidaredoxin reductase. *Methods Enzymol.* **52**, 166–188 (1978).
8. Gunsalus, I. C. & Sligar, S. G. Equilibrium states and dynamic reactions of iron in the camphor monooxygenase system. *Adv. Exp. Med. Biol.* **74**, 254–62 (1976).
9. Poulos, T. L., Finzel, B. C. & Howard, A. J. High-resolution crystal structure of cytochrome P450cam. *J. Mol. Biol.* **195**, 687–700 (1987).
10. Lee, Y. T., Glazer, E. C., Wilson, R. F., Stout, C. D. & Goodin, D. B. Three clusters of conformational states in CYP450cam reveal a multistep pathway for closing of the substrate access channel. *Biochemistry* **50**, 693–703 (2011).
11. Chuo, S. W., Wang, L. P., Britt, R. D. & Goodin, D. B. An Intermediate Conformational State of Cytochrome P450cam-CN in Complex with Putidaredoxin. *Biochemistry* **58**, 2353–2361 (2019).
12. Liou, S. H. *et al.* Linkage between Proximal and Distal Movements of CYP450cam Induced by Putidaredoxin. *Biochemistry* **59**, 2012–2021 (2020).
13. Sevrioukova, I. F., Garcia, C., Li, H., Bhaskar, B. & Poulos, T. L. Crystal structure of putidaredoxin, the [2Fe-2S] component of the CYP450cam monooxygenase system from *Pseudomonas putida*. *J. Mol. Biol.* **333**, 377–92 (2003).
14. Gunsalus, I. C. & Sligar, S. G. Redox regulation of cytochrome P450cam mixed function oxidation by putidaredoxin and camphor ligation. *Biochimie* **58**, 143–147 (1976).
15. Myers, W. K., Lee, Y.-T., Britt, R. D. & Goodin, D. B. The Conformation of CYP450cam in Complex with Putidaredoxin Is Dependent on Oxidation State. *J. Am. Chem. Soc.* **135**, 11732–11735 (2013).
16. Liou, S. H., Mahomed, M., Lee, Y. T. & Goodin, D. B. Effector Roles of Putidaredoxin on Cytochrome P450cam Conformational States. *J. Am. Chem. Soc.* **138**, 10163–10172 (2016).
17. Reipa, V., Holden, M. J., Mayhew, M. P. & Vilker, V. L. Temperature dependence of the formal reduction potential of putidaredoxin. *Biochim. Biophys. Acta - Bioenerg.* **1459**, 1–9 (2000).
18. Sligar, S. G. Coupling of Spin, Substrate, and Redox Equilibria in Cytochrome P450. *Biochemistry* **15**, 5399–5406 (1976).
19. Honeychurch, M. J., Hill, H. A. O. & Wong, L. L. The thermodynamics and kinetics of electron transfer in the cytochrome P450(cam) enzyme system. *FEBS Lett.* **451**, 351–353 (1999).
20. Sligar, S. G. & Gunsalus, I. C. A thermodynamic model of regulation: modulation of redox equilibria in camphor monooxygenase. *Proc. Natl. Acad. Sci.* **73**, 1078–1082 (1976).
21. Reipa, V., Mayhew, M. P. & Vilker, V. L. A direct electrode-driven CYP450 cycle for biocatalysis. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 13554–13558 (1997).
22. Lipscomb, J. D., Sligar, S. G., Namtvedt, M. J. & Gunsalus, I. C. Autooxidation and hydroxylation reactions of oxygenated cytochrome P-450cam. *J. Biol. Chem.* **251**, 1116–24 (1976).
23. Lee, Y. T., Wilson, R. F., Rupniewski, I. & Goodin, D. B. CYP450cam visits an open conformation in the absence of substrate. *Biochemistry* **49**, 3412–3419 (2010).
24. Poulos, T. L., Finzel, B. C. & Howard, A. J. Crystal Structure of Substrate-Free *Pseudomonas putida* Cytochrome P-450. *Biochemistry* **25**, 5314–5322 (1986).
25. Pylypenko, O. & Schlichting, I. Structural aspects of ligand binding to and electron transfer in

- bacterial and fungal CYP450s. *Annu. Rev. Biochem.* **73**, 991–1018 (2004).
26. Ascitutto, E. K., Dang, M., Pochapsky, S. S., Madura, J. D. & Pochapsky, T. C. Experimentally restrained molecular dynamics simulations for characterizing the open states of cytochrome P450cam. *Biochemistry* **50**, 1664–1671 (2011).
 27. Ascitutto, E. K., Young, M. J., Madura, J., Pochapsky, S. S. & Pochapsky, T. C. Solution structural ensembles of substrate-free cytochrome P450 cam. *Biochemistry* **51**, 3383–3393 (2012).
 28. Rydzewski, J. & Nowak, W. Thermodynamics of camphor migration in cytochrome P450cam by atomistic simulations. *Sci. Rep.* **7**, 1–10 (2017).
 29. Stoll, S. *et al.* Double electron-electron resonance shows cytochrome P450cam undergoes a conformational change in solution upon binding substrate. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 12888–12893 (2012).
 30. Basom, E. J., Maj, M., Cho, M. & Thielges, M. C. Site-Specific Characterization of Cytochrome P450cam Conformations by Infrared Spectroscopy. *Anal. Chem.* **88**, 6598–6606 (2016).
 31. Kazanis, S. *et al.* Conversion of a Fe₂S₂ ferredoxin into a Ga(3+) rubredoxin. *J. Am. Chem. Soc.* **117**, 6625–6626 (1995).
 32. Kazanis, S. & Pochapsky, T. C. Structural features of the metal binding site and dynamics of gallium putidaredoxin, a diamagnetic derivative of a CyS₄Fe₂S₂ ferredoxin. *J. Biomol. NMR* **9**, 337–346 (1997).
 33. Sligar, S. G., Debrunner, P. G., Lipscomb, J. D., Namtvedt, M. J. & Gunsalus, I. C. A role of the putidaredoxin COOH-terminus in P-450cam (cytochrome m) hydroxylations. *Proc. Natl. Acad. Sci. U. S. A.* **71**, 3906–10 (1974).
 34. Tripathi, S., Li, H. & Poulos, T. L. Structural basis for effector control and redox partner recognition in cytochrome P450. *Science*. **340**, 1227–1230 (2013).
 35. Shimada, H., Nagano, S., Hori, H. & Ishimura, Y. Putidaredoxin-cytochrome P450cam interaction. *J. Inorg. Biochem.* **83**, 255–260 (2001).
 36. Stayton, P. S. & Sligar, S. G. Structural Microheterogeneity of a Tryptophan Residue Required for Efficient Biological Electron Transfer between Putidaredoxin and Cytochrome P-450cam. *Biochemistry* **30**, 1845–1851 (1991).
 37. Schlichting, I. *et al.* The catalytic pathway of cytochrome P450cam at atomic resolution. *Science* **287**, 1615–22 (2000).
 38. Gerber, N. C. & Sligar, S. G. A role for Asp-251 in cytochrome P-450cam oxygen activation. *J. Biol. Chem.* **269**, 4260–6 (1994).
 39. Kuznetsov, V. Y. *et al.* The putidaredoxin reductase-putidaredoxin electron transfer complex: Theoretical and experimental studies. *J. Biol. Chem.* **280**, 16135–16142 (2005).
 40. Kuznetsov, V. Y., Poulos, T. L. & Sevrioukova, I. F. Putidaredoxin-to-cytochrome P450cam electron transfer: Differences between the two reductive steps required for catalysis. *Biochemistry* **45**, 11934–11944 (2006).
 41. Hiruma, Y. *et al.* The structure of the cytochrome P450cam-putidaredoxin complex determined by paramagnetic NMR spectroscopy and crystallography. *J. Mol. Biol.* **425**, 4353–4365 (2013).
 42. Skinner, S. P. *et al.* Delicate conformational balance of the redox enzyme cytochrome P450cam. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 9022–9027 (2015).
 43. Batabyal, D., Richards, L. S. & Poulos, T. L. Effect of Redox Partner Binding on Cytochrome P450 Conformational Dynamics. *J. Am. Chem. Soc.* **139**, 13193–13199 (2017).
 44. Tosha, T. *et al.* NMR study on the structural changes of cytochrome P450cam upon the complex formation with putidaredoxin. Functional significance of the putidaredoxin-induced structural changes. *J. Biol. Chem.* **278**, 39809–39821 (2003).
 45. Lipscomb, J. D. Electron Paramagnetic Resonance Detectable States of Cytochrome P-450Cam. *Biochemistry* **19**, 3590–3599 (1980).
 46. Nagano, S. *et al.* Infrared Spectroscopic and Mutational Studies on Putidaredoxin-Induced Conformational Changes in Ferrous CO-P450cam. *Biochemistry* **42**, 14507–14514 (2003).
 47. Unno, M. *et al.* Complex formation of cytochrome P450cam with putidaredoxin. Evidence for

- protein-specific interactions involving the proximal thiolate ligand. *J. Biol. Chem.* **277**, 2547–2553 (2002).
48. Tosha, T., Yoshioka, S., Ishimori, K. & Morishima, I. L358P mutation on cytochrome P450cam simulates structural changes upon putidaredoxin binding: The structural changes trigger electron transfer to oxy-P450cam from electron donors. *J. Biol. Chem.* **279**, 42836–42843 (2004).
 49. Nagano, S., Tosha, T., Ishimori, K., Morishima, I. & Poulos, T. L. Crystal structure of the cytochrome P450cam mutant that exhibits the same spectral perturbations induced by putidaredoxin binding. *J. Biol. Chem.* **279**, 42844–42849 (2004).
 50. Follmer, A. H., Mahomed, M., Goodin, D. B. & Poulos, T. L. Substrate-Dependent Allosteric Regulation in Cytochrome P450cam (CYP101A1). *J. Am. Chem. Soc.* **140**, 16222–16228 (2018).
 51. Follmer, A. H., Tripathi, S. & Poulos, T. L. Ligand and Redox Partner Binding Generates a New Conformational State in Cytochrome P450cam (CYP101A1). *J. Am. Chem. Soc.* **141**, 2678–2683 (2019).
 52. Colthart, A. M. *et al.* Detection of substrate-dependent conformational changes in the CYP450 fold by nuclear magnetic resonance. *Sci. Rep.* **6**, 1–12 (2016).
 53. Ascitutto, E. K. & Pochapsky, T. C. Some Surprising Implications of NMR-directed Simulations of Substrate Recognition and Binding by Cytochrome P450cam (CYP101A1). *J. Mol. Biol.* **430**, 1295–1310 (2018).
 54. Yao, H., McCullough, C. R., Costache, A. D., Pallela, P. K. & Sem, D. S. Structural evidence for a functionally relevant second camphor binding site in CYP450cam: model for substrate entry into a CYP450 active site. *Proteins* **69**, 125–38 (2007).
 55. Aramaki, H. *et al.* Formation of repressor-inducer-operator ternary complex: Negative cooperativity of d-camphor binding to CamR. *Genes to Cells* **16**, 1200–1207 (2011).
 56. Skinner, S. P. *et al.* Partial Opening of Cytochrome P450cam (CYP101A1) Is Driven by Allostery and Putidaredoxin Binding. *Biochemistry* [acs.biochem.1c00406](https://doi.org/10.1021/acs.biochem.1c00406) (2021). doi:10.1021/acs.biochem.1c00406
 57. Lewis, B. A. & Sligar, S. G. Structural studies of cytochrome P-450 using small angle x-ray scattering. *J. Biol. Chem.* **258**, 3599–3601 (1983).
 58. Brosey, C. A. & Tainer, J. A. Evolving SAXS versatility: solution X-ray scattering for macromolecular architecture, functional landscapes, and integrative structural biology. *Curr. Opin. Struct. Biol.* **58**, 197–213 (2019).
 59. Grishaev, A. Sample preparation, data collection, and preliminary data analysis in biomolecular solution X-ray scattering. *Curr. Protoc. Protein Sci.* **1**, 1–18 (2012).
 60. Kikhney, A. G. & Svergun, D. I. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett.* **589**, 2570–2577 (2015).
 61. Poulos, T. L., Perez, M. & Wagner, G. C. Preliminary crystallographic data on cytochrome P-450CAM. *J. Biol. Chem.* **257**, 10427–10429 (1982).
 62. Poulos, T. L., Finzel, B. C. & Gunsalus, I. C. The 2.6-Å crystal structure of *Pseudomonas putida* cytochrome P-450. *J. Biol. Chem.* **260**, 16122–16130 (1985).
 63. Lee, Y. T., Wilson, R. F., Rupniewski, I. & Goodin, D. B. CYP450cam visits an open conformation in the absence of substrate. *Biochemistry* **49**, 3412–3419 (2010).
 64. Laemmli, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **227**, 680–685 (1970).
 65. Singh, P. S. *Small-Angle Scattering Techniques (SAXS/SANS). Membrane Characterization* (Elsevier B.V., 2017). doi:10.1016/B978-0-444-63776-5.00006-1
 66. Hopkins, J. B., Gillilan, R. E. & Skou, S. BioXTAS RAW: Improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. *J. Appl. Crystallogr.* **50**, 1545–1553 (2017).
 67. Guinier, A. & Fournet, G. *Small angle scattering of X-rays*. (Wiley Interscience, 1955).
 68. Rambo, R. P. & Tainer, J. A. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **496**, 477–481 (2013).

69. Svergun, D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495–503 (1992).
70. Schneidman-Duhovny, D., Hammel, M. & Sali, A. FoXS: a web server for rapid computation and fitting of SAXS profiles. *Nucleic Acids Res.* **38**, 540–544 (2010).
71. Schneidman-Duhovny, D., Hammel, M., Tainer, J. A. & Sali, A. Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys. J.* **105**, 962–974 (2013).
72. Panjkovich, A. & Svergun, D. I. Deciphering conformational transitions of proteins by small angle X-ray scattering and normal mode analysis. *Phys. Chem. Chem. Phys.* **18**, 5707–5719 (2016).
73. Verderber, E. *et al.* Role of the hemA gene product and δ -aminolevulinic acid in regulation of *Escherichia coli* heme synthesis. *J. Bacteriol.* **179**, 4583–4590 (1997).
74. Mertens, H. D. T. & Svergun, D. I. Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J. Struct. Biol.* **172**, 128–141 (2010).
75. Pilz, I., Glatter, O. & Kratky, O. Small-angle X-ray scattering. *Methods Enzymol.* **61**, 148–249 (1979).
76. Burchard, W. & Schmidt, M. Static and dynamic structure factors calculated for flexible ring macromolecules. *Polymer (Guildf)*. **21**, 745–749 (1980).
77. Rambo, R. P. & Tainer, J. A. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers* **95**, 559–571 (2011).
78. Glatter, O. Data evaluation in small angle scattering: calculation of the radial electron density distribution by means of indirect Fourier transformation. *Acta Phys. Austriaca* **47**, 83–102 (1977).
79. Boldon, L., Laliberte, F. & Liu, L. Review of the fundamental theories behind small angle X-ray scattering, molecular dynamics simulations, and relevant integrated application. *Nano Rev.* **6**, 25661 (2015).
80. Svergun, D. I. & Koch, M. H. J. Small-angle scattering studies of biological macromolecules in solution. *Reports Prog. Phys.* **66**, 1735–1782 (2003).
81. Li, S. & Wackett, L. P. Reductive Dehalogenation by Cytochrome P450CAM: Substrate Binding and Catalysis. *Biochemistry* **32**, 9355–9361 (1993).
82. Mouro, C., Bondon, A., Jung, C., De Certaines, J. D. & Simonneaux, G. Assignment of heme methyl ¹H-NMR resonances of high-spin and low-spin ferric complexes of cytochrome P450cam using one-dimensional and two-dimensional paramagnetic signals enhancement (PASE) magnetization transfer experiments. *Eur. J. Biochem.* **267**, 216–221 (2000).
83. Schulze, H., Hui Bon Hoa, G., Helms, V., Wade, R. C. & Jung, C. Structural Changes in Cytochrome P-450 cam Effected by the Binding of the Enantiomers (1 R)-Camphor and (1 S)-Camphor. *Biochemistry* **35**, 14127–14138 (1996).
84. Lange, R., Bonfils, C. & Debey, P. The low-spin/high-spin transition equilibrium of camphor-bound cytochrome P-450. Effects of medium and temperature on equilibrium data. *Eur. J. Biochem.* **79**, 623–8 (1977).
85. Marden, M. C. & Hui Bon Hoa, G. P-450 binding to substrates camphor and linalool versus pressure. *Arch. Biochem. Biophys.* **253**, 100–107 (1987).
86. Schenkman, J. B. Nature of the type I and II spectral changes in liver microsomes. *Biochemistry* **9**, 2081–2091 (1970).
87. Kumaki, K., Sato, M., Kon, H. & Nebert, D. W. Correlation of type I, type II, and reverse type I difference spectra with absolute changes in spin state of hepatic microsomal cytochrome P-450 iron from five mammalian species. *J. Biol. Chem.* **253**, 1048–1058 (1978).
88. Mahajan, S. & Sanejouand, Y. H. On the relationship between low-frequency normal modes and the large-scale conformational changes of proteins. *Arch. Biochem. Biophys.* **567**, 59–65 (2015).
89. Frezza, E. & Lavery, R. Internal Coordinate Normal Mode Analysis: A Strategy to Predict Protein Conformational Transitions. *J. Phys. Chem. B* **123**, 1294–1301 (2019).

Chapter 5 : Conclusions and perspectives

5.1 The role of active site and second shell residues in oxygen activation by *NcLPMO9D*

Using neutron protein crystallography, we were successfully able to cryo-trap an activated dioxygen species at the *NcLPMO9D* copper active site following ascorbate reduction. As discussed in **Chapter 2**, this activated dioxygen species was modeled as a mixed state superoxo and hydroperoxo species and indicated that a second-shell His157 functions as the proton donor. The crystal packing resulted in a unique occluded active site environment analogous to substrate binding that permitted stabilization of the reactive oxygen intermediates. The insolubility of the LPMO polysaccharide substrate has made solution of crystal structures of a substrate bound form of LPMO challenging. The packing conformation observed in the *LPMO9D* crystal structure is therefore, while not functionally relevant in solution, a representation of how substrate may close off the copper active site environment and prevent dissociation of a reactive oxygen species as shown by Kjaergaard *et al.*¹

As a consequence of the limited substrate bound forms of LPMOs, most of the mechanistic insights are derived from density functional theory (DFT) and quantum mechanics/molecular mechanics (QM/MM) studies. While these provide a good first approximation of the energetic favorability of reaction intermediates and the role of second-shell residues during catalysis, the results are often dependent on the functional used, the residues included at the QM region and the charge and spin states assigned to the active site. A study by Larsson *et al.* has found that comparison of the TPSS, TPSSh and B3LYP methods predict very different stabilities of a range of potential activated oxygen species.² It is therefore essential to take experimental results into consideration during a computation to limit the bias a applied functional may introduce. The results presented in **Chapter 2** have highlighted a role for His157 in oxygen activation directly following copper reduction and also provided experimental evidence of a hydroperoxo species as an early intermediate in the LPMO reaction mixture. While our structural studies have permitted trapping

of an activated dioxygen species, the electronic state of the hydroperoxo intermediate cannot be determined solely by structural studies.

Our preliminary DFT studies showed that superoxo and single protonated His157 and hydroperoxo and doubly deprotonated His157 states are stable, however more detailed analysis of the charge and spin states of the copper and bound intermediates will be required in future work. QM/MM calculations to determine the electronic state of copper in the early stages of dioxygen activation with an accurate representation of the number of electrons will be needed. Our proposed role of His157 as a proton donor, and its capture in the imidazolate form is unusual, however not unprecedented. A neutral histidine acting as a general acid has been mechanistically proposed for the His95 residue in triose phosphate isomerase and the His98 residue in methylglyoxal synthase.³⁻⁵ Furthermore, the electropositive copper active site environment may contribute to the stabilization of histidine in an imidazolate form as observed for the second shell His147 residue in copper amine oxidase in a neutron crystallography study by Murakawa *et al.*⁶

We have further established that the second shell His157 residue displays conformational flexibility aided by mining minima free energy calculations. Mutagenesis studies of a the corresponding second shell His161 in *MtPMO3* by Span *et al.* found that His 161 plays a role as a proton donor during oxygen activation.⁷ It would informative to perform similar mutagenesis studies on His157 in LPMO9D to determine whether this second shell residue is essential for dioxygen activation. Kinetic analysis as well as structural studies of His157 mutants would provide a more complete picture of the role of this residue in catalysis. Indeed, the dynamics of the LPMO-polysaccharide system may play an essential role during catalysis permitting shuttling of protons for catalysis but also ensuring an occluded environment for formation of reactive intermediates for catalysis.

5.2 Large crystal growth of CYP450cam

Numerous expression and purification trials of CYP450cam were performed to optimize sample quality and purity. CYP450cam samples with sufficient purity and heme incorporation levels were obtained which were used for crystallization. As discussed in **Chapter 3**, CYP450cam was successfully crystallized and X-ray diffraction data collection and structure solution indicated that CYP450cam had correct heme incorporation and the substrate was correctly bound to the active site. The morphology of the obtained crystals presented primarily as needles and plates which are not of sufficient volume for neutron protein crystallography studies. Further optimization will therefore be necessary to ensure the growth of large, voluminous crystals. Crystallization trials employed vapor diffusion techniques as well as microbatch and microdialysis crystallization. Since these techniques yielded primarily needle-like crystals, further exploration of micro- and macroseeding will be necessary since it is apparent that initial crystals have a similar morphology. An alternative technique that was not applied that has been found to yield large crystals for neutron protein crystallography is free-interface capillary counter diffusion.⁸ Furthermore, since the needles observed in in CYP450cam crystallization trials form rapidly, often after 24 hours, the rate of nucleation may be too rapid to ensure the formation of large crystals. It has been found that use of a semi-solid agarose gel (SSAG) can slow nucleation and promote large crystal growth.⁹ Therefore, future trials using such a SSAG system as well as application of macroseeding techniques may assist large crystal growth for neutron protein crystallography.

5.3 Conformational dynamics of CYP450cam

Solution small angle X-ray scattering (SAXS) studies were performed on CYP450cam to study its conformational flexibility in solution as discussed in **Chapter 4**. While an early study had been performed on CYP450cam, it concluded that conformational changes could not be detected upon substrate camphor binding.¹⁰ Solution of the X-ray crystal structure of CYP450cam subsequently obviated the need of the comparatively low resolution SAXS technique to study CYP450cam.¹¹ While SAXS can be considered, low-resolution, it is however an invaluable technique for the study of the conformational flexibility of proteins since it permits their characterization in their native conformation in solution. Crystallographic structural analysis has the drawback of providing a static picture of the protein structure, since only the conformations that fit into the crystal lattice are selected for.

SAXS analysis of CYP450cam in a range of camphor concentrations confirmed conformational transition from an open to a closed state as has been observed in crystallographic studies. Very low camphor concentrations of 1 μM near the CYP450cam K_D for camphor indicated a conformational change by a decrease in the radius of gyration (R_g) and the particle maximum dimension (D_{max}). A Leu358Pro CYP450cam mutant that mimics binding of the redox partner protein Putidaredoxin (Pdx) and the covalent CYP450cam-Pdx complex did not display this decrease in R_g confirming their stabilization in the open conformation. We further observed a change in R_g and D_{max} at 43 μM camphor, a concentration at which camphor has been proposed to bind to a second site on CYP450cam.¹² This second binding site of camphor on CYP450cam has been proposed as an allosteric site using MD and paramagnetic NMR and has been found to cause the partial opening of CYP450cam.^{13,14} The SAXS studies presented here support a change in conformation of CYP450cam near the concentration of camphor at which is it believed to bind to the allosteric site. Furthermore, normal mode analysis of the flexibility of CYP450cam informed by the experimental SAXS data indicate that the allosteric region delimited by the β 2-sheet which connects the H- and I-helices as well as the H-helix, G-helix, D-helix and C-helix displayed increased mobility when compared to the crystal structure.

These studies bring to the forefront a more dynamic nature of CYP450cam that is carefully modulated by substrate and redox partner protein interactions. Future SAXS studies of CYP450cam with greater q -resolution and subsequent modeling may provide a more detailed description of the conformational flexibility of CYP450cam. Mutagenesis of the residues involved in camphor allosteric binding and studies of the resulting conformational flexibility using SAXS and catalysis using kinetic studies will further inform on the role of allostery and conformational change in CYP450cam.

A final insight can be gleaned when taking the results from the crystallization attempts in **Chapter 3** and putting them into context with the insights on conformational flexibility in **Chapter 4**. It has been observed that crystallization is in itself a purification step in which only certain protein conformations are selected and effectively “purified” into a homogenous protein crystal.¹⁵ Our SAXS studies indicated a surprising finding that at higher ($> 60 \mu\text{M}$) of camphor, there is a decrease in the R_g and D_{max} from the more open conformation at $43 \mu\text{M}$ camphor which may result from camphor allosteric binding. While it has been shown that CYP450cam undergoes a spin state transition at high levels of camphor, our SAXS results indicate that CYP450cam also undergoes a conformational change. At high levels of camphor CYP450cam adopts a low-spin hexacoordinated conformation (Type II species). The crystallization trials performed in **Chapter 3** were conducted at high levels of camphor in excess of $1000 \mu\text{M}$. These high levels of camphor may have resulted in a conformational change in the CYP450cam preventing it from crystallizing and altering the thermodynamics of crystallization of the population in the substrate-bound high-spin conformation. Indeed, all the crystallization trials of CYP450cam indicate the formation of precipitate even when crystals grew (Figure 5.1).

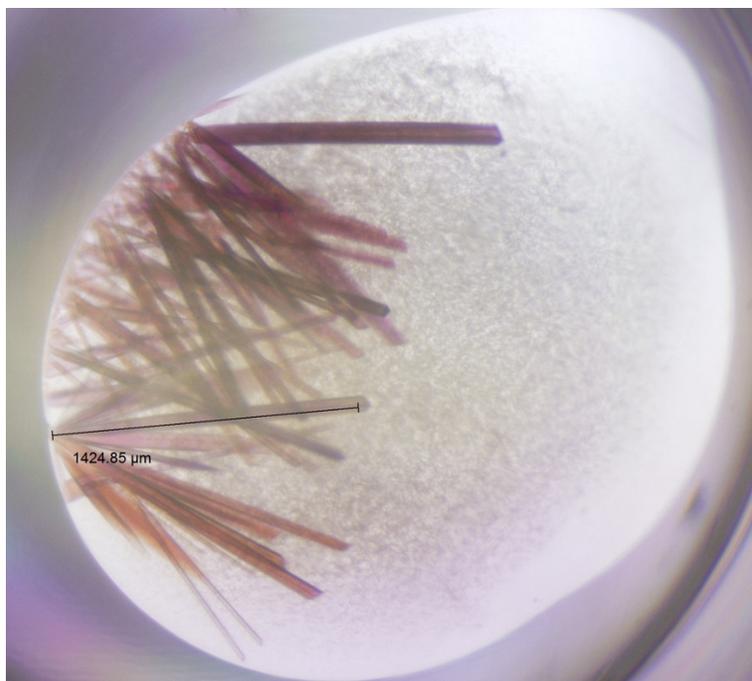


Figure 5.1 Sitting drop crystallization of CYP450cam with needle-like crystals and precipitate formation.

This reiterates the role of crystallization as a final purification step and also highlights the delicate conformational balance of CYP450cam dependent on substrate concentration. Future crystallization trials informed by this knowledge of the conformational flexibility of CYP450cam will be performed in which low camphor concentrations at the K_D for active site camphor binding or in the absence of camphor are used. This will allow us to determine whether the addition of excess camphor negatively impacted past crystallization efforts.

5.4 References

1. Kjaergaard, C. H. *et al.* Spectroscopic and computational insight into the activation of O₂ by the mononuclear Cu center in polysaccharide monooxygenases. *Proc. Natl. Acad. Sci.* **111**, 8797–8802 (2014).
2. Larsson, E. D., Dong, G., Veryazov, V., Ryde, U. & Hedegård, E. D. Is density functional theory accurate for lytic polysaccharide monooxygenase enzymes. *Dalt. Trans.* **49**, 1501–1512 (2020).
3. Lodi, P. J. & Knowles, J. R. Neutral imidazole is the electrophile in the reaction catalyzed by triosephosphate isomerase: structural origins and catalytic implications. *Biochemistry* **30**, 6948–6956 (1991).
4. Lennartz, C., Schäfer, A., Terstegen, F. & Thiel, W. Enzymatic reactions of triosephosphate isomerase: A theoretical calibration study. *J. Phys. Chem. B* **106**, 1758–1767 (2002).
5. Marks, G. T., Susler, M. & Harrison, D. H. T. Mutagenic Studies on Histidine 98 of Methylglyoxal Synthase: Effects on Mechanism and Conformational Change. *Biochemistry* **43**, 3802–3813 (2004).
6. Murakawa, T. *et al.* Neutron crystallography of copper amine oxidase reveals keto/enolate interconversion of the quinone cofactor and unusual proton sharing. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 10818–10824 (2020).
7. Span, E. A., Suess, D. L. M., Deller, M. C., Britt, R. D. & Marletta, M. A. The Role of the Secondary Coordination Sphere in a Fungal Polysaccharide Monooxygenase. *ACS Chem. Biol.* **12**, 1095–1103 (2017).
8. Vahdatahar, E., Junius, N. & Budayova-Spano, M. Optimization of crystal growth for neutron macromolecular crystallography. *J. Vis. Exp.* **2021**, 1–24 (2021).
9. Matsumura, H. *et al.* Approach for growth of high-quality and large protein crystals. *J. Synchrotron Radiat.* **18**, 16–19 (2011).
10. Lewis, B. A. & Sligar, S. G. Structural studies of cytochrome P-450 using small angle x-ray scattering. *J. Biol. Chem.* **258**, 3599–3601 (1983).
11. Poulos, T. L., Finzel, B. C. & Howard, A. J. High-resolution crystal structure of cytochrome P450cam. *J. Mol. Biol.* **195**, 687–700 (1987).
12. Yao, H., McCullough, C. R., Costache, A. D., Pallela, P. K. & Sem, D. S. Structural evidence for a functionally relevant second camphor binding site in CYP450cam: model for substrate entry into a CYP450 active site. *Proteins* **69**, 125–38 (2007).
13. Follmer, A. H., Mahomed, M., Goodin, D. B. & Poulos, T. L. Substrate-Dependent Allosteric Regulation in Cytochrome P450cam (CYP101A1). *J. Am. Chem. Soc.* **140**, 16222–16228 (2018).
14. Skinner, S. P. *et al.* Partial Opening of Cytochrome P450cam (CYP101A1) Is Driven by Allostery and Putidaredoxin Binding. *Biochemistry* [acs.biochem.1c00406](https://doi.org/10.1021/acs.biochem.1c00406) (2021). doi:10.1021/acs.biochem.1c00406
15. Pochapsky, T. C. & Pochapsky, S. S. What Your Crystal Structure Will Not Tell You about Enzyme Function. *Acc. Chem. Res.* (2019). doi:10.1021/acs.accounts.9b00066