

**Investigating the Catalytic Mechanisms of
Bio-degrading Copper Proteins:
Multi-copper Oxidases (MCOs) and Lytic Polysaccharide Monooxygenases (LPMOs)**

Thesis by
Jieun Shin

In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Chemistry

The logo for the California Institute of Technology (Caltech), featuring the word "Caltech" in a bold, orange, sans-serif font.

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2021

(Defended May 27th, 2021)

© 2021

Jieun Shin

ORCID: 0000-0003-3817-6282

All Rights Reserved

To my family, friends, and mentors throughout my life

ACKNOWLEDGEMENTS

For PhD students, graduate school is a long journey taking up more than half of our twenties in most cases. During my PhD, Caltech has indeed been a second home (if not the first) for me. Being away from family as an international student, the Caltech community has been a family to me, and completing my PhD would have been impossible with the help from such a great group of people I have got to know here at Caltech over the years. I have sincere gratitude toward all of you.

First of all, I would like to thank my advisor and my committee. I thank Harry Gray and Jay Winkler so much for helping me to learn and pursue my dream. There have been so many challenges during my PhD, but Harry and Jay have patiently guided me along the way like a grandpa and a dad, and interacting with them has been a great part of my life. I would not have been able to get through this hard journey without their kind support. Whenever I have problems, I could always run to either one of them or both, and the problem was most of the times magically gone. Harry probably is one of the most responsive and patient mentors. I will never be able to forget all of his quick email replies with dozens of exclamation marks and the biggest smile in the world he gives to people. In addition to science and academic skills, I learned from Harry the “people skills” required in leadership positions and how to stay positive and motivated even in this serious academic world. I really thank him for being so understanding and caring. I will never be able to forget our back-and-forth paper editing processes from which I got to learn a lot as well as our weekend meetings and our trip together to Ventura for the conference (I will make sure I learn how to drive so that I can drive you one day!). I’d also like to thank Jay for his devoted mentorship over the years. I would not have been able to survive my PhD without all the hands-on trainings I have been able to get from

Jay. He taught me colloquial English (or American, more correctly) expressions, MATLAB coding, how to use and fix tools and instruments, laser spectroscopy and all the other lab techniques. Because of Jay's fatherly help, I feel like I have become a much more functional human being. Whenever I have interactions with people these days, I feel much more confident about myself after years of training from him, because I can always relate to something he taught me before. In my mind, I always go like 'oh, Jay taught me that!' whenever I do something. Jay, I will never be able to forget all of our daily conversations on both life and science, and most of all, the warmest smiles you used to give me when I try things even when I was very incompetent in the lab in the beginning. Thank you so much for being there for me when I was taking all the baby steps. Harry and Jay: you have been more than just mentors to me during my stay at Caltech.

I also feel very lucky that I have had a privilege of having such a wonderful group of people in my committee: Michael Hoffmann, Woody Fischer and Doug Rees (chair). I thank them so much for being such amazing mentors and supporters over the years.

Mike is the very first Caltech professor I got to know. My memory from his group during my first year still takes up a big place in my heart, and I will never be able to forget his kindness and caring support. Mike always tried to see potentials and values in me and warmly encouraged me whatever I chose to do. It really meant a lot to me. I really thank Mike and Rod Kiewiet for giving me the opportunity to do summer research in Mike's group. Both Mike and Rod have always been so encouraging and motivational that I always try to think about them whenever I feel frustrated with grad school life to remind myself of how desperately I wanted to do research when I was getting started. The research experience I got to have in Hoffmann group really helped me to build a solid foundation to be a better scientist. Thank you, Mike, for all your kind advice on both research and life throughout my PhD.

Woody is also one of the most considerate and communicative professors I have met throughout my life. I really thank Woody for always welcoming me into meetings and discussions, and for always giving me great research suggestions and helpful advice to get things going. The weekly Mn meeting was really like an inspirational shelter for me to recharge and refresh myself, and I got to learn so much from him and his group. Woody, thank you for being so understanding of grad students' difficulties and frustrations. You once said you are the eldest among your five Fischer siblings, and I can see where your responsible and affectionate character comes from. You really have this big brother or father-like presence, and I thank you so much for always caring about my progress and for listening to all my whining.

I also thank Doug so much for being such a wonderful chair of my committee and for always providing me great advice. Because of his help, I got to keep things going with completing the PhD requirements. Although Doug had millions of other duties as a dean of graduate studies, he always made time for students with his prompt responses. Whenever I had to check things concerning my thesis progress, he was very understanding and supportive, so I really thank him for all of his kind help and encouragement. Moreover, Doug, I really enjoyed working as your TA for one of your classes, and learned a lot from both your teaching style and the course content. I will never be able to forget your kindness from all of my interactions with you.

Harry, Jay, Mike, Woody and Doug, all of you are my role models, and hopefully I can be a great mentor like you all one day.

I would also like to acknowledge all the help I have been getting from amazing staff scientists and the PIs at various Caltech research facilities. Mona Shahgholi was one of the kindest and the most responsible staff scientists I got to know at Caltech. I really thank Mona for always being so caring and understanding of my research life. She always empathized with my worries, and helped me to make progress with my research as efficiently as possible with her Mass-spec

expertise. Whenever I had difficulties with my samples, she really tried to help me to get going by taking her time to think about what other options we could try. We need more staff scientists like you, Mona! I also thank Brett Lomenick, Jeff Jones and Tsui-Fen Chou at the proteomics center. Their expertise in peptide analysis and mass-spec really helped me to get going with my project. It was fun preparing the samples with the kind help from Brett, and the help from Jeff on data interpretation was also so valuable. I really learned a lot from all the discussions I had with Brett, Jeff and Tsui-Fen. Thank you so much for always making the time for me! I also thank Annie Moradian and Michael Sweredoski for helping me with peptide mapping to confirm the expression of my enzymes for my cell samples, and I thank Jens Kaiser for helping me with protein crystallization for X-ray crystallography. It was fun to set up the screening plates with Jens, and I learned a lot from his explanations. For characterizing my enzymes, Paul Oyala helped me with the EPR spectroscopy and Nathan Dalleska helped me with the ICP-MS. I thank them so much for their time and all the helpful discussions. In addition, I got to try the XRF with the help from George Rossman. I really thank him for taking his time in the lab to try out the measurements with me and for all his kind explanations.

To get my LPMO projects going, the help I got from Zarah Forsberg and Vincent Eijssink from Norwegian university of life sciences (NMBU) was crucial. I happened to get started with working on one of the poorly expressing LPMOs, so I was struggling a lot in the beginning to optimize the expression and purification conditions for over six months. When I reached out to figure out what the problems were, Vincent was very willing to help me out and got me connected with Zarah who had had over ten years of experience in working with LPMOs. I really thank him for his kindness and openness to share. Zarah kindly took time to send me the plasmids of some of the better expressing LPMOs she had been working on, and helped me until I got to successfully express active enzymes. All the weekly zoom meetings I got to have

with her for about three months really helped me to learn more about LPMOs and to get going with my project. I really thank her for sharing her experience and expertise.

All the learning experiences I have been able to get in Gray group indeed helped me to grow both as a person and as a scientist. During my stay in Gray group over the years, I have overlapped with Astrid Mueller, Brad Brennan, Brendon McNicholas, Brian Sanders, Bryan Hunter, Carl Blumenfeld, Dana Levine, Emmanuel Despagnet-Ayoub, Ethan Simonoff, Javier Fajardo Jr., Jill Clinton, Joseph Schwan, Julian West, Katharina Brinkert, Maryann Morales, Oliver Shafaat, Patrycja Kielb, Raheleh Ravanfar, Sarah Del Ciello, Shabnam Hematian, Wes Kramer and Yuling Shen. There were also visiting students from Denmark and Poland: Anders Bjerrum, Mette Luetzen Hoff Sorensen and Rafal Orłowski, who brought in more diverse experience to the group. We have also had an amazing group of undergrads, Alessio Amaolo, Alessandra Mondello, Cherish Nie, Danh Ngo, Gyu Bin Jang, Janice Kang, Jayce Miller, Perry Hicks, Peter Saghy and Tom Sheridan, who made the group atmosphere livelier and fresher with their pure spirit! I really thank all of you for creating a great work environment from which I got to learn and grow. I guess you are the people I got to spend most of my time with during my PhD, and I will always remember all the positive experience I have had. All my experience in Gray group really meant a lot to me.

My gratitude also goes to my friends in Rees, Barton, Tirrell and Clemons Group: Maiko Obana, Hyun-Gi Yun, Bekah Silva, Jeff Lai, Ailienna Maggiolo, Chengcheng Fan, Sam Ho, Miguel Pinto, Kausar Raza and Siobhan MacArdle, Adela Nano, Stephanie Threatt, Rebecca Warmack and Naima Sharaf. Due to the similar nature of our work, we got to share the experience with each other a lot. Whenever something stops working or runs out in the lab under time pressure, I could always run to these people to borrow things and talk about

frustrations. Thank you so much for letting me use the instruments, lending me lab supplies and all the helpful discussions to broaden my perspective in my research. Moreover, since I did not get to overlap with any other grad student in the bio sub-group of Gray group since I officially joined the group, being able to connect with Maiko, Hyun-Gi, Bekah, Ailienna and Siobhan over the years really meant a lot to me. I miss going to conferences together with you in particular, and thank you so much for all the emotional support and all our chats about research and grad school experiences. You were like sisters to me during my PhD.

I also had unforgettable experiences in Hoffmann group during my first year of grad school. I thank Kangwoo Cho, Yang Yang, Siwen Wang, Yanzhe Zhu, Alison Wu, Eunkyung Kim, Su Young Ryu, Cody Finke, Nina Bahnemann, Clement Cid, Xing Xie and Justin Jasper for making my transition from undergrad to grad school easier when I was feeling a little lost right after I came from Korea. I really learned a lot from you all, and thank you for your friendship and for creating such a supportive and inclusive environment. Going through the course work and qual preparation together with you people made all the hard work so much more endurable. Moreover, Kangwoo and Yang were such amazing mentors to me, and thank you for training me and helping me with my experiments.

I also really thank our Mn meeting group: Joan Valentine, John Magyar, Usha Lingappa, Hope Johnson, Calvin Rusley and Anna Li. I thank Woody for letting me join the meetings, and I thank all of the Mn meeting members for always warmly welcoming me and giving me the opportunities to learn and share. Weekly Mn meetings have really been joys of my life, and I thank you so much for all your support and encouragement especially during the last year of my PhD. Joan and John always kindly gave me good advice since my first year in grad school, and Usha has been a great friend to me. I really thank their help and support over the years.

Moreover, Goddard group members, Soo Kyoung Kim, Soonho Kwon, Moon Young Yang, Hye Young Shin and Yalu Chen, have also been good friends to me over the years. Since their offices are also in the sub-basement of Beckman Institute where I usually work, I was able to always run to them whenever I feel frustrated with my experiments. Although they are computational people, they always tried to empathize with me and encouraged me to keep going. I thank them not only for their emotional support but also for sometimes pulling me out of the lab in the evenings or in the weekends for lunches and dinners to help me not to drown myself in the lab.

My grad school experience would have been much more difficult without the help from so many people all across campus. I thank Daniel Yoder and Laura Kim from the international student programs office for helping me with transitions. I remember sitting in Daniel's office on the first day of my arrival at Caltech back in the summer of 2015 for the check-in procedures. I was telling him how scary it feels being away from my family all of a sudden across the ocean. He calmed me down by sharing his experiences of visiting foreign countries. I still can't forget his kind words. My interaction with Daniel and Laura continued in various events for international students and whenever I needed travel documents or visa support. I really thank them for always being so open and kind. A tremendous amount of thanks also goes to Alison Ross and Rebecca Fox from the chemistry office. Keeping track of all the PhD requirements and timelines would have been impossible without Alison's help. She has been so understanding and kind, and I really thank her for caring about my progress and always calming me down with her encouragement. It really meant a lot to me. I also thank Rebecca for her kind help when there was a scheduling conflict in the seminar talk. I almost freaked out right before my talk, but she resolved the situation so professionally and kindly, so it was a great relief to

me. She always kindly helped me with my inquiries and needs. I also thank Brian Stoltz for being such an amazing option representative. He was also always so supportive and encouraging. I really thank the graduate studies office and Caltech library as well for all of their efforts to give better learning experiences to students.

On my daily research life, my research would not have gone efficiently without the help from Greg Rolette, Armando Villasenor and Joe Drew from the chemistry stock room, and Denise Rodriquez, Jason Salib, Jamie Garofalo and many other VWR staffs from the biology VWR stockroom. They always ensured that research supplies arrive and get delivered on time, and kindly helped me whenever there were shipping delays or missing packages. I thank Greg, Armando and Joe for their help with all the incoming packages, and I thank Denise, Jason and Jamie for always helping me with efficient purchases of lab supplies. They always tried their best to make sure my research does not get delayed due to any out-of-stock materials. Their positive energy really helped me to get going despite of many frustrations in research. I will never be able to forget all the chats we had in the stockroom especially with Denise. Even when the supplies I was waiting for arrived after her regular work hours, she was willing to stay longer to help me to get the materials on time. These kinds of sincere help from people across campus gave me energy to keep going. Moreover, I thank our building managers Rick Jackson and Susan Ruffins, and our administrative assistant Pat Anderson. Rick, Susan and Pat were the problem solvers of the building and many other related administrative issues, always being so supportive of students. Thank you for creating such a great working environment in the building. I also thank Victor Rivera and Elisa Brink from Caltech procurement office for helping me with efficient purchases and payment of lab supplies and for running the Techmart so efficiently. Also, I thank all the janitors of the building for creating such a pleasant atmosphere and the Caltech facility for efficient gas deliveries.

Furthermore, I would also like to thank all my other Caltech friends who have been such amazing emotional supporters. There is no doubt that graduate studies are at times very frustrating. Without the encouragement from Byung-Kuk Yoo, Larry Henling, Serim Ryu, Areum Kim, Albert Chung, Hyeongchan Jo, Heejung Ahn, Jonghoon Lim, Jeongmin Kim, Seu Sim, Seola Lee and Chong Sun, I would not have been able to get through difficulties. Whenever I whine about things out of frustration, they always helped me not to give up and keep going. I also thank Jeong Hoon Ko and Sonjong Hwang for their good advice about grad school life, and I thank Linus Eng, Katharina Brinkert, Alexandra Barth and Ellen Yan for their emotional support. Moreover, I thank Taleen Dilanyan for TA-ing together with me. Because of Taleen, my TA-ing experience was more fun, and I learned a lot from working with her. I would also like to thank the Pasadena LEARNS Program team: Alexander, Vidhya, Yi and Andrew. Although we could not continue the mentoring activities longer due to pandemic, it was a very meaningful experience for me.

My deep gratitude goes to my family as well, of course. I talk with my parents every day after work. Although I try not to let out my work stress on them, I sometimes couldn't hide my fatigue and frustration. Because of their emotional support and love, I was able to reflect on things from different perspective even when there were many difficulties. Thank you for believing in me and thank you for caring about me so much. I also thank my little brother (25 years old right now, but still is and always will be a little brother to me). All the daily heart emojis my family send to each other in our group chat room have been the best medication for healing me from stress and anxiety. I miss you so much!

I also thank all my friends back in Korea, who always worry about me whenever they hear

from news reports that something happened in California. (Although California is even much bigger than South Korea, they automatically relate whatever they hear about California with me.) I sometimes find it pretty funny, but I thank them for always thinking of me and caring about my well-being.

In addition, I would like to thank all my friends and mentors who made me become who I am right now. I think I have been able to get to this point because of all the help I've got from so many people throughout my life. Whenever there are difficulties, there were always mentors and friends who offer me unexpected opportunities or sincere help without wanting any credit. To reciprocate all the help I've got, I wish to become someone who can make positive impacts to the world.

I don't know how many times I had to cry while writing this. I will never be able to forget all the memories I have with you all. Thank you so much for all the good memories over the years.

ABSTRACT

Lignin and cellulose comprise a large portion of the renewable biomass on Earth. However, substantially due to laborious course of processing, the conversion efficiency of these biomaterials to accessible biofuel is very low. Therefore, effective depolymerization and utilization of these biopolymers are requirements for environmentally friendly and sustainable energy development. In the hope of finding solutions to these biomass utilization challenges, there have been growing interests in using biodegrading metalloenzymes as active biocatalysts. However, there still remain many questions regarding mechanistic details of enzyme catalysis and effective application of these enzymes. This thesis focuses on investigating the redox chemistry involved in the catalytic mechanisms of two main lignin- and cellulose- degrading copper enzymes: multicopper oxidases (MCOs) and lytic polysaccharide monooxygenases (LPMOs).

MCOs are capable of aerobic oxidation of lignin as their primary function, but the nature of their substrate variability also allows the oxidation of not only diverse high potential organic and inorganic complexes, but also earth abundant divalent metal ions such as manganese. LPMOs, on the other hand, enable the cleavage of glycosidic bonds in recalcitrant insoluble cellulosic substances, which are not degradable by other hydrolytic enzymes such as endoglucanases and cellulobiohydrolases.

It is remarkable that nature has created such versatile enzymes with specific active site metals and redox-active amino acids involved in electron transfer, which contribute to substrate oxidation as well as enzyme survival against oxidative damage during catalysis. By gaining a deeper understanding of how these enzymes work, we could greatly enhance current usage efficiencies and develop more energy-efficient biocatalysts.

Chapter I gives an introduction to biological coppers, two groups of bio-degrading copper enzymes: multicopper oxidases (MCOs) and lytic polysaccharide monooxygenases (LPMOs), and the role of redox-active amino acids in electron transfer and enzyme catalysis. For the MCO work, a thermophilic laccase (*Tth-lac*) from *Thermus thermophilus* HB27 and a CotA laccase (CotA-lac) from *Bacillus Subtilis* were studied. For the LPMO work, two cellulose active LPMOs (ScLPMO10B and ScLPMO10C) and a chitin active LPMO (BILPMO10A) were studied.

Chapter II describes thermodynamic aspects of *Tth-lac* catalysis. The temperature dependence of the formal potential of type I copper (Cu_{TI}) in *Tth-lac* is reported, and the interplay between many competing dynamic and thermodynamic factors which results in thermostability and activity of *Tth-lac* is discussed.

Chapter III reports the electron transfer (ET) kinetics data obtained with *Tth-lac* using the transient absorption spectroscopy. The results of photochemical electron/hole transfer studies indicate that the chains of Trp and Tyr can participate in electron transfer through *Tth-lac*, which could potentially have a role in enzyme catalysis as well.

Chapter IV discusses the protective role of a Trp/Tyr pair positioned close to the trinuclear copper cluster (TNC) in *Tth-lac*. It is indeed remarkable that laccases are capable of utilizing the power of oxygen to catalyze the oxidation of diverse high-potential substrates. But, as a tradeoff, the utilization of dioxygen can make the enzyme highly susceptible to oxidative damage. Chapter IV provides supporting evidence that led us to conclude that the TNC-

proximal Trp/Tyr pair functions as an internal antioxidant for prolonging the enzyme lifetime.

Chapter V describes investigations on the factors that affect MCO catalysis, which include the potentials of the active site coppers, possible reactive intermediates, and common structural motifs. Based on the structural homology between *Tth*-lac and CotA-lac, some preliminary work done on CotA-lac is also reported.

Chapter VI outlines the work on LPMOs. After the successful expression and purification of ScLPMO10B, ScLPMO10B and BILPMO10A, standard activity assays were done with insoluble cellulose and chitin substrates to confirm the enzyme activity. The results are compared with that from the photo-degradation experiments to investigate if the photochemically generated Cu(III) species are active intermediates in LPMO catalysis.

Chapter VII reports the results on bioinformatics analysis on the distribution of vicinal amino acids in different enzyme classes. This study was to examine the biological significance of amino acid pairs and clusters existing in many different enzyme classes, with vicinal surface tyrosines in CotA-lac as an underlying motivation behind the work.

This thesis demonstrates that MCOs and LPMOs are truly versatile enzymes which can oxidize such diverse refractory substrates, and there could be multiple pathways that the enzymes achieve this task. As shown so far, not only the active site metals but also the chain of redox-active amino acids as well as metal coordinating residues can contribute to enzyme catalysis.

PUBLISHED CONTENT AND CONTRIBUTIONS

Chapter 2:

Reprinted with permission from Springer Nature Customer Service Centre GmbH:

Shin J., Gray H.* and Winkler J.* (2020) Stability/Activity Tradeoffs in *Thermus thermophilus* HB27 Laccase. *J Biol Inorg Chem.* 25, 233-238. Doi: 10.1007/s00775-020-01754-7.

J.S. performed all experiments, wrote the initial draft, and participated in modifying the manuscript.

Chapter 4:

Shin J., Gray H.* and Winkler J.* (2021) Trp/Tyr pair protects *Thermus thermophilus* HB27 laccase from oxidative damage. - *Manuscript submitted to ACS Chem Biol.*

J.S. performed all experiments, wrote the initial draft, and participated in modifying the manuscript.

TABLE OF CONTENTS

Acknowledgements.....	iv
Abstract	xi
Published Content and Contributions	xii
Table of Contents	xiii
List of Illustrations and/or Tables	vii
Nomenclature.....	viii
Chapter I: Bio-degrading Cu Enzymes and Reactive Intermediates	1
1.1. Biological copper centers	2
1.2. Bio-degrading copper enzymes	3
1.3. Multicopper Oxidases (MCOs)	4
<i>Thermus thermophilus</i> HB27 laccase (<i>Tth-lac</i>)	
CotA laccase from <i>Bacillus Subtilis</i> (<i>CotA-lac</i>)	
1.4. Lytic Polysaccharide Monooxygenase (LPMO).....	7
1.5. Roles of redox-active Trp/Tyr in electron transfer	10
1.6. References	12
Chapter II: Stability/Activity Tradeoffs in <i>Thermus thermophilus</i> HB27 Laccase.....	16
2.1. Abstract	17
2.2. Introduction	17
2.3. Methods.....	19
Sample preparation	
Circular dichroism spectra	

Estimation of the Cu _{T1} potential from redox equilibria	
2.4. Results and discussion	21
Temperature dependence of the Cu _{T1} formal potential	
2.5. Concluding remarks	28
2.6. Acknowledgment	28
2.7. References	29
2.8. Supplementary material	33
S2.1. Protein preparation	
S2.2. Circular dichroism spectra	
S2.3. Synthesis of [Ru(NH ₃) ₄ (bpy)](PF ₆) ₂	
S2.4. Estimation of the Cu _{T1} potential from redox equilibria	
S2.5. Thermodynamic parameters	
S2.6. Hydrophobicity, polarity and solvent accessible surface area (SASA)	
S2.7. References	
Chapter III: Electron Transfer through Laccase from <i>Thermus thermophilus</i> HB27.....	46
3.1. Introduction	47
3.2. Laser sample preparation	47
Protein expression	
Protein purification	
Surface labeling with [Ru(bpy) ₂ (Phen)IA] ²⁺ complex	
Synthesis of the Ru-photosensitizer	
Ru-labeling mechanism	
Metalation with CuSO ₄	
Protein quantification	

Copper quantification	
Laser samples	
Quantum yield calculation	
Other reduction methods	
3.3. Flash-quench laser experiments	61
Reductive quenching pathway	
Oxidative quenching pathway	
Quencher studies	
Laser setup	
3.4. Transient absorption kinetics.....	66
Ru-C117 enzyme	
Ru-C117-W118F enzyme	
Ru-C117-W96F enzyme	
Investigating the pH effects	
Photo-reduction of laser samples	
Ru-C102 enzyme	
Activity check after the laser experiments	
Sodium azide inhibition effects	
3.5. Intramolecular long-range copper oxidation.....	79
3.6. Appendices.....	80
Kinetics fitting	
Temperature effects on the single-step electron transfer with Ru-C390 enzyme	
3.7. References	92

Chapter IV: Trp/Tyr Pair Protects <i>Tth</i>-lac from Oxidative Damage	95
4.1. Abstract	96
4.2. Introduction	96
4.3. Results and discussion	99
Enzyme activity	
Enzyme deactivation	
4.4. Concluding remarks	103
4.5. Materials and methods	103
Enzyme preparation	
Activity assay	
Total turnover measurements	
LC-MS and peptide analysis	
4.6. Acknowledgment	105
4.7. Funding sources	105
4.8. References	106
4.9. Supporting information.....	110
Chapter V: Investigating the Factors Affecting Laccase Catalysis: Active-site Potentials & Structural Motifs	122
5.1. Potential discrepancy	123
5.2. The Cu _{T1} -depleted (T1D) laccase and T1D-W118F.....	128
5.3. Met-rich loop mutant	131
5.4. CotA laccase from <i>Bacillus Subtilis</i> (CotA-lac)	136
5.5. References	139

Chapter VI: Investigating the Mechanistic Details of Lytic Polysaccharide Monooxygenase

(LPMO)	145
6.1. Introduction	146
6.2. Enzyme preparation	147
6.3. Activity assay	152
6.4. Photodegradation experiment	154
6.5. Cu ³⁺ transient absorption spectrum	156
6.6. Appendices	158
6.7. Acknowledgments	163
6.8. References	163

Chapter VII: Occurrence of Amino Acid Pairs and Clusters in Biological Systems 165

7.1. Motivation	166
7.2. Surface vicinal tyrosines	166
7.3. Expected and observed amino acid distribution	171
7.4. Distribution of internal vicinal Y-Y in different enzyme classes	176
7.5. More refined list of PDB codes	177
7.6. Biological assemblies	178
7.7. Conclusion	182
7.8. Appendices	183

LIST OF FIGURES

Chapter I:

Figure 1.1. Lignin and cellulose degradation by laccases and LPMOs.	3
Figure 1.2. Structure of a multicopper oxidase from <i>Thermus thermophilus</i> HB27 (PDB: 2YAE).	5
Figure 1.3. Coordinating residues of the active-site coppers in <i>Tth</i> -lac.	5
Figure 1.4. Structure of a multicopper oxidase from <i>Bacillus Subtilis</i> (PDB: 1GSK).	6
Figure 1.5. Proposed mechanisms of LPMO catalysis: O ₂ and H ₂ O ₂ mechanisms.	7
Figure 1.6. Enzyme structures of ScLPMO10B (PDB: 4OY6) and ScLPMO10C (PDB: 4OY7) with the active-site copper, conserved aromatic residue, and Trp and Tyr throughout the protein	9
Figure 1.7. Structural homology between <i>Tth</i> -lac and CotA-lac.	10

Chapter II:

Figure 2.1. Temperature dependence of the catalytic efficiency of <i>Tth</i> -lac oxidation of substrate ABTS	19
Figure 2.2. UV-Vis spectra (20 to 65°C) of a deoxygenated sample containing both wild type <i>Tth</i> -lac and four equivalents of [Ru(NH ₃) ₄ (bpy)](PF ₆) ₂	22
Figure 2.3. Plots of $\Delta E^\circ/T$ vs $1/T$ (left) and ΔE° vs T (right)	23
Figure 2.4. Plot of the Cu _{T1} ^{2+/+} reduction entropy (ΔS°_{rc} , J mol ⁻¹ K ⁻¹) versus the solvent accessible surface area (Å ²) of active-site binding residues in blue copper proteins.....	25
Figure S2.1. Circular dichroism spectra of <i>Tth</i> -lac from 20 to 65°C.....	35
Figure S2.2. UV-vis spectra (20 to 65°C) for deoxygenated samples of [Ru(NH ₃) ₄ (bpy)](PF ₆) ₂ and of wild type <i>Tth</i> -lac	37

Figure S2.3. Equilibrium concentrations of [Ru ²⁺], [Ru ³⁺], [Cu ²⁺] and [Cu ⁺]	38
Figure S2.4. ΔS°_{rc} vs. Hydrophobicity of residues within 8 Å from Cu _{T1}	42
Figure S2.5. E° of Cu _{T1} vs. Hydrophobicity of residues within 8 Å from Cu _{T1}	42
Figure S2.6. ΔS°_{rc} vs Polarity of residues within 8 Å from Cu _{T1}	43
Figure S2.7. The correlation between reduction entropy (ΔS°_{rc} , J mol ⁻¹ K ⁻¹) and relative solvent accessible surface area (RSA).....	43

Chapter III:

Figure 3.1. Ru-photosensitizer surface labeling sites on <i>Tth</i> -lac	50
Figure 3.2. Potential of Ru-species (E° vs. NHE in water) generated by the photosensitizer	51
Figure 3.3. Ru-photosensitizer labeling to a surface Cys	52
Figure 3.4. MALDI-TOF spectrum of the purified protein	53
Figure 3.5. Protein after labeling (orange) and after labeling and metalation (green).....	54
Figure 3.6. Problems with [Ru(NH ₃) ₆]Cl ₃ as a reversible quencher	58
Figure 3.7. Quantum yield calculation for the Ru(NH ₄) ₆ Cl ₃ quencher	59
Figure 3.8. Spectra showing the reduction of coppers in Ru-labeled <i>Tth</i> -lac	60
Figure 3.9. Flash-quench cycle of the Ru-photosensitizer.....	61
Figure 3.10. Cage escape yield of 1, 4-benzoquinone	63
Figure 3.11. Photochemical instability of 1,4-benzoquinone	64
Figure 3.12. Photo-stability of benzoquinone with 455 nm (left) and 365 nm (right) excitation	64
Figure 3.13. Quencher mechanisms of Methyl Viologen (MV) and Benzyl Viologen (BV)	65

Figure 3.14. Transient absorption kinetics of Ru-C117 at 440 nm	66
Figure 3.15. Transient absorption kinetics of Ru-C117 at 510 nm	67
Figure 3.16. Transient absorption kinetics of Ru-C117 at 555 nm	67
Figure 3.17. Transient absorption kinetics of Ru-C117 at 605 nm	68
Figure 3.18. The first Trp mutation (W118F) right next to the labeling site (S117C)	69
Figure 3.19. Transient absorption data collected at 510 nm and 555 nm (Trp radical absorption regions) for both S117C mutant and S117C-W118F mutant.....	70
Figure 3.20. Multi-wavelength transient absorption data for Ru-C117 mutant and Ru-C117-W118F mutant (overlaid)	71
Figure 3.21. The second Trp mutation (W96F) near the tri-nuclear copper cluster	72
Figure 3.22. The transient absorption signals at 510nm and 555nm for the first Trp (W118F) and the second Trp mutant (W96F).....	72
Figure 3.23. Multi-wavelength transient absorption spectroscopy using the diode array system for Ru-C117 mutant.....	74
Figure 3.24. Oxidized green laccase getting photochemically reduced	75
Figure 3.25. S102C labeling site and its distance from a nearby Trp residues (W96) and the TNC	76
Figure 3.26 Activity check with pMeODMA and ABTS after the laser experiment	77
Figure 3.27. Laccase samples with NaN ₃ upon irradiation.....	78
Figure 3.28. Electron/hole transfer kinetics of intramolecular long-range copper oxidation	79
Figure 3.29. Transient absorption kinetics for S117C-W118F and S102C with HeNe.....	82
Figure 3.30. Ru-photosensitizer labeling site (red) near the Cu _{T1} of <i>Th</i> -lac	83
Figure 3.31. Luminescence kinetics of [Ru(bpy) ₃ Cl]Cl ₂	84
Figure 3.32. Luminescence kinetics of Ru-C390 at 650 nm (log time scale).....	85

Figure 3.33. Luminescence kinetics of Ru-C390 at 650 nm	85
Figure 3.34. Transient absorption kinetics of [Ru(bpy) ₃ Cl]Cl ₂ at 450 nm	87
Figure 3.35. Transient absorption kinetics of Ru-C390 at 450 nm	88
Figure 3.36. Transient absorption kinetics of Ru-C390 at 633 nm with HeNe	89
Figure 3.37. Kinetics profile of the Ru-C390 mutant obtained with HeNe	90

Chapter IV:

Figure 4.1. Proposed steps in the catalytic cycle (red arrows). NI decay to RO is off path; upon 4-electron reduction, RO re-enters the cycle (black dashed arrows)	97
Figure 4.2. A Trp/Tyr chain (W96/Y120/W118) and a Trp/Tyr pair (W133-Y134) near the trinuclear copper cluster (PDB: 2XU9).....	98
Figure 4.3. Activity assay profiles at early times (initial 5 min) for WT and mutant enzymes (500 nM [E], 3 mM ABTS ²⁻ in MES buffer, pH 5.3, [O ₂] = 240 μM).....	100
Figure 4.4. Enzyme activity assays with ABTS ²⁻ (mM, ± 0.2 %) at pH 4.5 for WT and Trp/Tyr mutants.	100
Figure S4.1. Graphic showing the active site coppers (blue) with coordinating residues (grey) and Trp/ Tyr residues near the trinuclear copper cluster	111
Figure S4.2. LC-MS results before and after enzyme total turnovers.....	111
Figure S4.3. Graphic showing a Trp/Tyr (W133/Y134) pair near the Tth-lac TNC and the results of peptide analysis indicating the oxidation sites	112
Figure S4.4. MS/MS data for the oxidized peptide fragments in <i>Tth</i> -lac WT	113
Figure S4.5. MS/MS data for the oxidized peptide fragments in <i>Tth</i> -lac W133F mutant	114

Figure S4.6. UV-vis spectra showing the relative amount of ABTS ^{•-} formed after enzyme total turnovers.....	115
Figure S4.7. Absorbance of ABTS ^{•-} at 420 nm monitored for 1 h at 20 °C	115
Figure S4.8. Enzyme activity assays (substrate ABTS at pH 4.5) for Trp/Tyr chain mutants.	117
Figure S4.9. Enzyme activity assays (substrate ABTS at pH 5.3) for WT and Trp/Tyr chain mutants	120
Figure S4.10. Circular dichroism spectra of <i>Tth</i> -lac WT and W133F-Y134F.	121

Chapter V:

Figure 5.1. Oxidation of leucoberbelin blue (LBB) and manganese by <i>Tth</i> -lac, pH 4.5	124
Figure 5.2. Anaerobic titration of three reducing equivalents of ABTS to the <i>Tth</i> -lac WT	125
Figure 5.3. Activity assay of <i>Tth</i> -lac WT and the W118F mutant with promazine hydrochloride as a substrate, pH 4.5	128
Figure 5.4. T1D Enzyme (C445S) and its surface Trp mutant (C445S-W118F)	129
Figure 5.5. Distribution of Met (magenta), His (yellow), Asp (green) and the active-site copper-coordinating residues (blue) in <i>Tth</i> -lac	132
Figure 5.6. Activity comparison of <i>Tth</i> -lac WT and <i>Tth</i> -lac MA (all Met, Asp and His in the Met-rich loop mutated to Ala) in the presence and absence of excess copper in the activity assay media (NaOAc, pH 4.5)	134
Figure 5.7. Effects of excess copper concentration on the degree of activity enhancement for <i>Tth</i> -lac WT, <i>Tth</i> -lac MA and <i>Tth</i> -lac DH.....	135

Figure 5.8. Structural homology between <i>Thh</i> -lac and CotA-lac.....	136
--	-----

Chapter VI:

Figure 6.1. Structures of bacterial LPMOs (PDB: 4OY6, 4OY7, 6TWE) active towards cellulose and chitin: an aromatic amino acid on the cellulose/chitin binding domain (red), Trp (cyan) and Try (green) residues.....	146
Figure 6.2. A schematic diagram of ScLPMO10B gene construct	147
Figure 6.3. Sequence analysis of ScLPMO10B by trypsin digestion.....	149
Figure 6.4. MALDI-TOF analysis of LPMO purified by the Avicel purification method	150
Figure 6.5. MALDI-TOF data showing degradation products of PASC (phosphoric acid swollen cellulose) in a standard activity assay with ScLPMO10C	153
Figure 6.6. MALDI-TOF data showing degradation products of β -chitin in a standard activity assay with BILPMO10A (top) and no product formation in a photo-excitation experiment (bottom).....	155
Figure 6.7. UV-Vis absorption spectrum of Cu(II)-bound DAHK peptide	156
Figure 6.8. Transient absorption kinetics spectra of the photochemically generated DAHK-Cu(III) at around 400 nm (right).....	157
Figure 6.9. Single wavelength transient absorption spectroscopy with ScLPMO10C and Ru(NH ₃) ₆ Cl ₃ (440 nm probed).	159
Figure 6.10. Single wavelength transient absorption spectroscopy with BILPMO10A and Ru(NH ₃) ₆ Cl ₃ (440 nm probed)	159
Figure 6.11. Single wavelength transient absorption spectroscopy with BILPMO10A and methyl viologen (400 nm probed).....	160

Figure 6.12. Single wavelength transient absorption spectroscopy with BILPMO10A and methyl viologen (450 nm probed)	160
Figure 6.13. Single wavelength transient absorption spectroscopy with BILPMO10A and methyl viologen (510 nm probed)	161
Figure 6.14. Single wavelength transient absorption spectroscopy with BILPMO10A and methyl viologen (550 nm probed)	161
Figure 6.15. Single wavelength transient absorption spectroscopy with Ru(bpy) ₃ Cl ₃ and [Co(NH ₃) ₅ Cl]Cl ₂ (450 nm probed)	162
Figure 6.16. Single wavelength transient absorption spectroscopy with BILPMO10A and [Co(NH ₃) ₅ Cl]Cl ₂ (450 nm probed)	162

Chapter VII:

Figure 7.1. Distribution of proteins (left) and proteins with more than or equal to one surface vicinal Y-Y (right) in different protein classes	167
Figure 7.2. Distribution of proteins with different numbers of surface vicinal Y-Y in different protein classes	175
Figure 7.3. Distribution of different numbers of internal vicinal Y-Y in different enzyme classes	177
Figure 7.4. Distribution of X-X amino acid pairs in biological assemblies.....	179
Figure 7.5. Relative numbers in % of location specific vicinal X-X in biological assemblies	179
Figure 7.6. Percentage of amino acid X appearing as a vicinal X-X in biological assemblies	180
Figure 7.7. Percentage of amino acid X appearing as a cluster in biological assemblies	181

LIST OF TABLES

Chapter I:

Table 1.1 Different types of coppers in metalloenzymes.	2
---	---

Chapter II:

Table 2.1. Effect of temperature on the reaction between $\text{Cu}_{\text{T1}}^{2+}$ and $[\text{Ru}(\text{NH}_3)_4(\text{bpy})]^{2+}$	22
Table 2.2. Reduction entropy change ($\Delta S_{\text{rc}}^{\circ}$, $\text{J mol}^{-1} \text{K}^{-1}$) and measures of solvent accessibility of copper ligands in blue copper proteins	25
Table S2.1. Temperature dependent equilibrium concentrations of Cu^{1+} , Cu^{2+} , Ru^{2+} and Ru^{3+}	37
Table S2.2 Hydrophobicity of residues within 8 Å from Cu_{T1}	41
Table S2.3 Polarity of residues within 8 Å from Cu_{T1}	41
Table S2.4 The solvent accessible surface area (Å^2) and the relative solvent accessible surface area (RSA) of copper ligands in blue copper proteins	44

Chapter III:

Table 3.1. Biomolecular quenching rate constant.....	62
---	----

Chapter IV:

Table 4.1. Catalytic efficiencies: WT and the Trp/Tyr mutants in NaOAc, pH 4.5	101
--	-----

Table 4.2. Activity assay: double mutant under O ₂ limiting conditions	101
Table S4.1. Absorbance of ABTS ^{•-} at 420 nm in partially oxidized [1 mM] _{initial} ABTS solution containing 1 mM CuSO ₄ in 25 mM NaOAc buffer, pH 5.3	116
Table S4.2. Michaelis-Menten kinetics in NaOAc buffer, pH 4.5	116
Table S4.3. Activity assay with different batches of proteins to check reproducibility	117
Table S4.4. Activity assay results with a doubled enzyme concentration	118
Table S4.5. Uncertainties of catalytic efficiencies: WT and the Trp/Tyr mutants in NaOAc, pH 4.5	119
Table S4.6. Michaelis-Menten kinetics in MES buffer, pH 5.3	120
Table S4.7. Catalytic efficiencies: WT and the Trp/Tyr mutants in MES, pH 5.3	121

Chapter V:

Table 5.1. Various laccase substrates with a wide range of potentials	123
Table 5.2. Substrate concentration dependence of W118F activity toward Promazine hydrochloride	127
Table 5.3. Protein concentration dependence of W118F activity toward Promazine hydrochloride	127
Table 5.1. Various laccase substrates with a wide range of potentials	123

Chapter VI:

Chapter VII:

Table 7.1. Surface vicinal Y-Y.....	167
Table 7.2. Distribution of surface Y-Y containing proteins in different sub-categories	168
Table 7.3. Surface vicinal W-W	169
Table 7.4. Surface vicinal Y-W	169
Table 7.5. Surface vicinal H-H.....	170
Table 7.6. Surface vicinal F-F	170
Table 7.7. Summary.....	171
Table 7.8. Expected and observed amino acid distribution.....	171
Table 7.9. PDB frequency vs. Total frequency.....	172
Table 7.10. Percentage of aromatic residue pairs	173
Table 7.11. Distribution of different number of Y-Y pairs in proteins.....	174
Table 7.12. Identification of proteins with the highest number of surface Y-Y pairs	174
Table 7.13. Percentage of surface vicinal aromatic amino acid pairs.....	175
Table 7.14. Number of proteins having different numbers of internal Y-Y without any surface Y-Y.....	176

LIST OF ABBREVIATIONS

ABTS	2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)
Å	Angstrom
BV	Benzyl viologen
bpy	2,2'-bipyridine
CD	Circular Dichroism
Co	Cobalt
CotA-lac	CotA laccase from <i>Bacillus subtilis</i>
Cu	Copper
Cu _{T1}	Type I copper
Cu _{T2}	Type II copper
Cu _{T3}	Type III copper
CuSO ₄	Copper sulfate
Cys	Cysteine
DCC	N,N'-dicyclohexylcarbodiimide
Dpn I	Type IIM restriction enzyme
e ⁻	Electron
E	Redox potential
EDTA	Ethylenediaminetetraacetic acid
ET	Electron transfer
EPR	Electron paramagnetic resonance
eV	Electron volt
FPLC	Fast protein liquid chromatography
g	Gram
HeNe	Helium-neon laser

His	Histidine
H ₂ O	Water
H ₂ O ₂	Hydrogen peroxide
H ⁺	Proton
h	Hour
hν	Light
IAPhen	5-iodoacetamido-1,10-phenanthroline
ICP-MS	Inductively coupled plasma mass spectrometry
IPTG	Isopropyl β-D-1-thiogalactopyranoside
J	Joule
K	Kelvin
K _{eq}	Equilibrium constant
K _m	Description of the affinity of the substrate to the active site of the enzyme
k _{cat}	Turnover number
k _{cat} /K _m	Catalytic efficiency
L	Liter
LB	Lysogeny broth
lb	Lower bound
LBB	Leucoberbelin blue
LCMS	Liquid chromatography-mass spectrometry
LED	Light-emitting diodes
LPMO	Lytic polysaccharide monooxygenase
M	Molar
MALDI-TOF	Matrix-assisted laser desorption/ionization-Time of flight
Mass-spec	Mass Spectroscopy

MCO	Multicopper oxidase
MES	2-ethanesulfonic acid
Mn	Manganese
MV	Methyl viologen
Micro-ED	Microcrystal electron diffraction
min	Minute
mL	Milliliter
mM	Millimolar
mol	Mole
mV	Millivolt
NaCl	Sodium chloride
NaN ₃	Sodium azide
NaOAc	Sodium acetate
Nd:YAG	Neodymium: yttrium aluminum garnet
NHE	Normal hydrogen electrode
NH ₃	ammonia
NI	Native intermediate
NTA	Nitrilotriacetic acid
nm	Nanometer
ns	Nanosecond
OD	Optical density
OPO	Optical paramagnetic oscillator
O ₂	Dioxygen
PASC	Phosphoric acid swollen cellulose
PCR	Polymerase chain reaction

PDB	Protein data bank
PF ₆	hexafluorophosphate
PI	Peroxide intermediate
pH	Power of hydrogen
pI	Isoelectric point
pK _a	Acid dissociation constant
pMeODMa	para-methoxydimethylaniline
py	Pyridine
QE HF	Q-exactive high field
ROS	Reactive oxygen species
RSA	Relative solvent accessible surface area
Ru	Ruthenium
Ru(NH ₃) ₆ Cl ₃	Hexaammineruthenium(III) chloride
SASA	Solvent accessible surface area
SP HP	Sepharose High Performance
s	Second
T1D	Type I copper depleted enzyme
TB	Terrific broth
TCEP	Tris-(2-carboxyethyl) phosphine
T _m	Melting temperature
TNC	Tri-nuclear copper cluster
Tris	Tris(hydroxymethyl)aminomethane (2-Amino-2-(hydroxymethyl)propane-1,3-diol)
Trp	Tryptophan
<i>Tth-lac</i>	Laccase from <i>Thermus thermophilus</i> HB27

<i>Tth</i> -lac MA	All methionines (Met), an aspartic acid (Asp) and a histidine (His) on the loop of <i>Tth</i> -lac mutated to alanines (Ala)
<i>Tth</i> -lac DH	Only Asp and His residues on the loop of <i>Tth</i> -lac mutated to Ala
Tyr	Tyrosine
UV	Ultraviolet
ub	upper bound
V	Volt
Vis	Visible
WT	Wild-type
XRF	X-ray Fluorescence
°	Degree
°C	Degree Celsius
Δ	Change in
ΔH°_{rc}	Standard reduction enthalpy
ΔS°_{rc}	Standard reduction entropy
ϵ	Extinction coefficient
λ	Wavelength
λ_{max}	Wavelength of local maximum intensity
μL	Microliter
μs	Microsecond
τ	Time constant
*	Excited state
Φ	Quantum yield

Chapter 1

Bio-degrading Cu Enzymes and Reactive Intermediates

Multi-copper Oxidases (MCOs) and Lytic Polysaccharide Monooxygenases (LPMOs)

1.1. Biological Copper Centers

Copper ions are found in many of the biological systems, and there exist many enzymes, mostly oxidoreductases, with different numbers of coppers in them [1]. Although they can all be grouped into copper proteins, not all coppers in them are the same. They are classified into different types depending on the coordination environment around the active copper sites.

Table 1.1. Different types of coppers in metalloenzymes [2], [3].

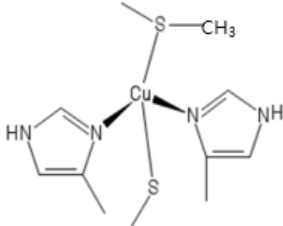
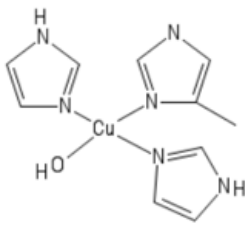
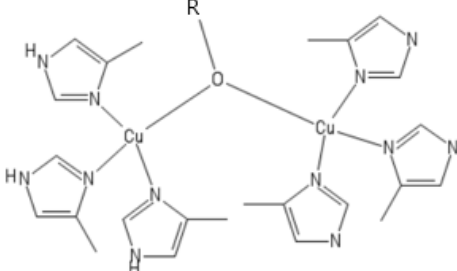
Type 1 Cu	Type 2 Cu	Type 3 Cu
		
<p>Cys-S \rightarrow Cu_{T1}^{II} charge transfer band at 610nm giving the intense blue color ($\epsilon = \sim 5000 \text{ M}^{-1}\text{cm}^{-1}$)</p>	<p>Weak absorption features</p>	<p>Oxygen-bridged dimer, absorption at $\sim 330\text{nm}$ due to the formation of hydroxo-bridge</p>

Table 1.1 shows different types of coppers existing in multicopper oxidases. Type 1 Cu (Cu_{T1}) is responsible for giving the absorption at around 600 nm due to Cys-S to Cu(II) charge transfer, whereas type 2 Cu (Cu_{T2}) does not have any distinct absorption feature in the UV-Vis spectrum. Type 3 Coppers (Cu_{T3}), which appear as an oxygen-bridged dimer have the absorption at 330 nm [2].

1.2. Bio-degrading Copper Enzymes

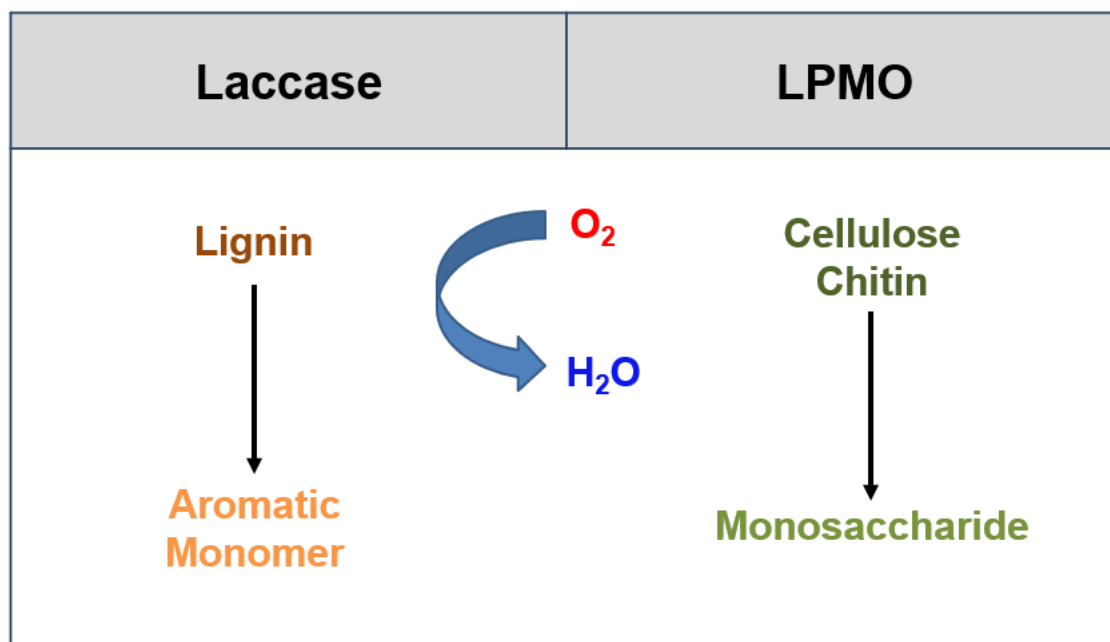


Figure 1.1. Lignin and cellulose degradation by laccases [4] and LPMOs [5], respectively.

Laccases and LPMOs have copper centers in their active sites, and they are renowned for their functions of degrading lignin and cellulose, respectively [1, 2]. Laccases have four copper centers capable of coupling the oxidation of lignin substrates to the four-electron reduction of O_2 to form H_2O . LPMOs are known to degrade insoluble and recalcitrant polysaccharides such as chitin, cellulose, or lignin. LPMOs tend to have their active site copper exposed on the surface and one single aromatic redox active residue on the putative cellulose binding surface. The reason why these enzymes are worthy of studying is that lignin and cellulose comprise a large portion of the renewable biomass on earth [6]. Substantially due to difficult course of processing, the conversion efficiency of these biomaterials to accessible biofuel is very low [7], and these enzymes could contribute to solving this biomass problem. Although these enzymes have been investigated for decades, there still remain a lot of details in their catalytic

mechanism to be elucidated, and studying these enzymes may also provide us a better insight into developing energy-efficient synthetic biocatalysts.

1.3. Multicopper Oxidases (MCOs)

Laccases are a group of enzymes classified as multi-copper oxidases (MCOs), which have four copper sites involved in electron transfer processes and enzyme turnovers. Some of the other most commonly recognized MCOs are ascorbase oxidases and ceruloplasmin. Nitrite reductase is closely related to MCOs, in terms of both sequence and structural similarity, although it physiologically functions as a reductase [8].

MCOs evolved from single domain cupredoxins which went through a series of domain duplication and active-site deletion and creation. Single domain cupredoxins evolved to two-domain MCOs, three-domain MCOs, and eventually six-domain MCOs. Three-domain MCOs exhibit the largest degree of variability in sequences and functions, and they occur in a great diversity of organisms [8], [9]. In most MCOs, there is a Cu_{T1} center and a tri-nuclear Cu cluster (TNC) which is composed of three Cu atoms: a Cu_{T2} and binuclear Cu_{T3} coppers. With these four coppers, they are capable of coupling the oxidation of substrates to the four-electron reduction of O_2 to form H_2O . According to the consensus mechanism, substrate oxidation is believed to occur near the Cu_{T1} donating an electron to the Cu_{T1} , which is followed by a long-range electron transfer of around 13\AA distance to the TNC where dioxygen reduction occurs to produce water ($\text{O}_2 + 4\text{H}^+ + 4\text{e}^- \rightarrow 2\text{H}_2\text{O}$) [10].

Two different three-domain MCOs from *Thermus thermophilus* HB27 and *Bacillus subtilis* were studied to elucidate the mechanistic details.

1.3.1. *Thermus thermophilus* HB27 Laccase (*Tth-lac*)

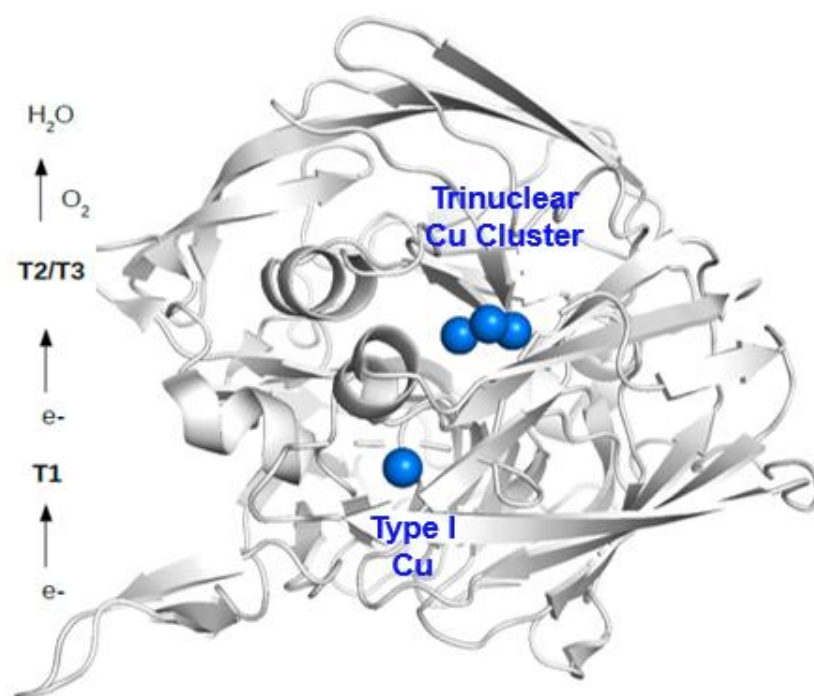


Figure 1.2. Structure of a multicopper oxidase from *Thermus thermophilus* HB27 (PDB: 2YAE [10]).

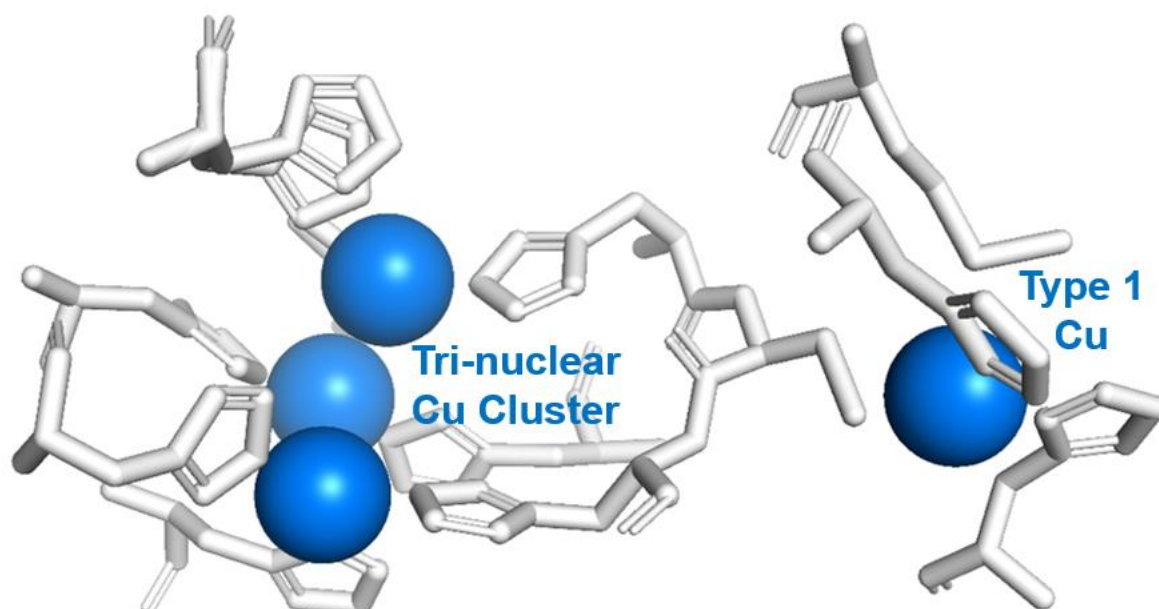


Figure 1.3. Coordinating residues of the active-site coppers in *Tth-lac* [10].

Cu_{T1}-coordinating residues are His393, His450, and Cys445 with an axial Met455 in close proximity. Cu_{T2} is coordinated by His95 and His396, and Cu_{T3} and Cu_{T3'} are connected via a hydroxo-bridge, and are coordinated by three histidines each: His137, His398, and His444 for Cu_{T3} and His97, His135, and His446 for Cu_{T3'}. These residues are critical for maintaining the active-site geometry and thus the enzyme function in *Tth*-lac [10].

1.3.2. CotA Laccase from *Bacillus Subtilis* (CotA-lac)

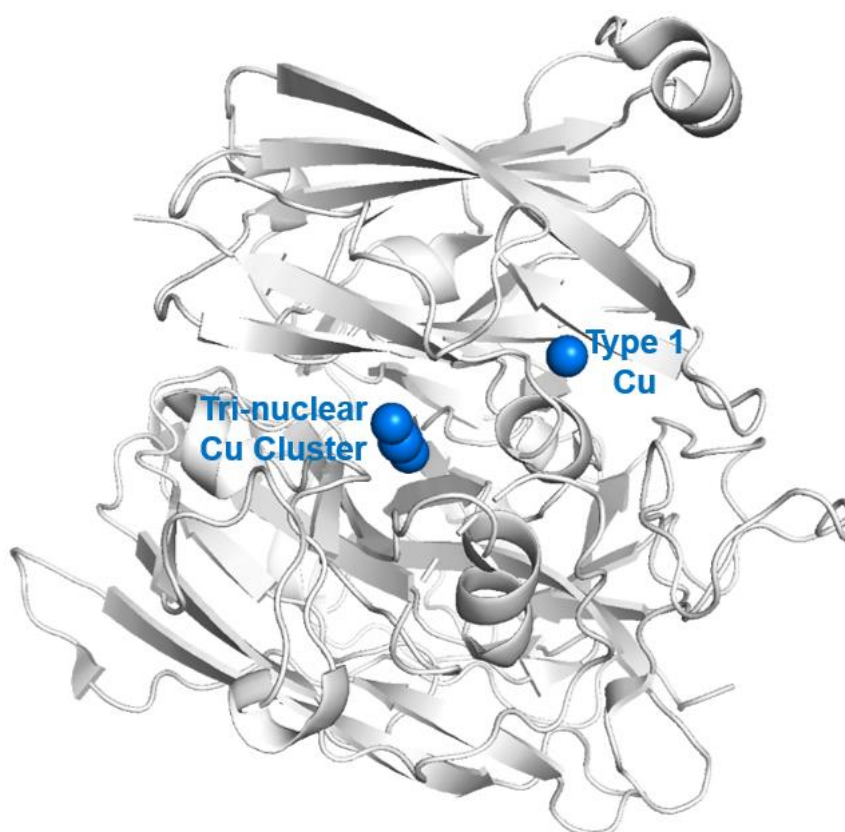


Figure 1.4. Structure of a multicopper oxidase from *Bacillus Subtilis* (PDB: 1GSK [11]).

The MCO from *Bacillus Subtilis*, also called CotA laccase (CotA-lac), is a bacterial endospore coat component. It belongs to a three-domain laccase just like *Tth*-lac. In CotA-lac, Cu_{T1}-coordinating residues are His419, His497, and Cys492 with an axial Met502 at a distance of

around 3.3Å. Cu_{T2} is coordinated by His105, His422, and a water ligand. Cu_{T3} and Cu_{T3'} are connected via a hydroxo-bridge, and are coordinated by three histidines each: His107, His153, and His493 for Cu_{T3} and His155, His424, and His491 for Cu_{T3'}. These residues are critical for maintaining the active-site geometry and thus the enzyme function in CotA-lac [10].

1.4. Lytic Polysaccharide Monooxygenase (LPMO)

Lytic polysaccharide monooxygenases (LPMOs) are a family of enzymes that comprise fungal and bacterial enzymes known to degrade insoluble and recalcitrant polysaccharides such as chitin, cellulose, or lignin. Bacterial LPMOs belong to the auxiliary activity family 10 (AA10) which is in general found in cellulolytic bacteria. There are LPMOs which target either C1 or C4 exclusively and some which can oxidize both C1 and C4 for the cleavage of glycosidic bonds [12], [13]. These independently functioning enzymes can even work synergistically with hydrolases (hydrolytic enzymes) and also with each other enhancing the catalytic activity for an efficient cellulose degradation [12].

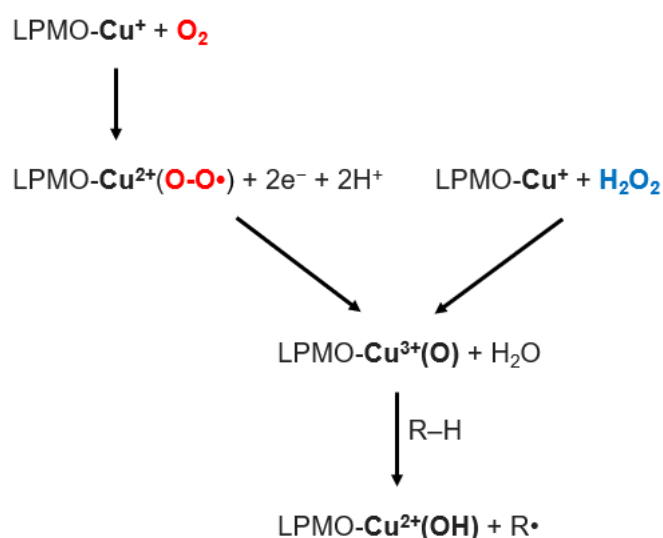


Figure 1.5. Proposed mechanisms of LPMO catalysis: O₂ and H₂O₂ mechanisms [14], [15].

According to the consensus mechanism, copper in the active site of LPMOs is believed to reduce dioxygen by obtaining two electrons from an external electron donor which tends to be a low molecular weight redox-active compound such as ascorbic acid or polyphenolic substance. Examples of phenolic reducing agents include gallic acid or the macromolecular lignin [13]. Reduced dioxygen takes hydrogen from the substrate, which ultimately leads to the cleavage of β -1,4 glycosidic linkages in polysaccharides [16]. However, this view was challenged by new data indicating that both O_2 and H_2O_2 can function as co-substrates. This is further complicated by the fact the LPMOs can generate H_2O_2 when the enzyme is not bound to substrate [15].

Interestingly, both of these postulated mechanisms include generation of Cu(III) intermediates (Cu(II)-oxyl or Cu(III)-oxo species) resulting from the oxidation of Cu(I) by O_2 or H_2O_2 , followed by O-O bond cleavage [17], but questions also arise concerning the feasibility of lignocellulosic substrate oxidation when the potential of Cu(II)(OH)/Cu(I)(OH₂) (estimated to be less than 0.3 V for LPMOs [18], [19]) is not high enough to promote hydroxyl radical rebound.

Substrate accessibility and the source of electrons are critical for enabling the efficient enzymatic degradation of insoluble polysaccharides. It has been reported that some small molecular weight phenolic compounds and lignin components existing in natural plant biomass serve as electron donors for LPMO activity. However, the specificity is presumably low, since only a certain number of polyphenolic substances have the sufficiently low redox potential to achieve the reduction of the active-site copper (+250 mV vs NHE) [20]. Monophenols, unlike some low potential compounds such as the ones with a 1,2-benzenediol or a 1,2,3-benzenetriol

moiety, were not adequate sources of electrons due to their relatively high redox potential [13]. Therefore, it can be understood that the supply of electrons at the right potential to enable the reduction of copper is a critical aspect to consider.

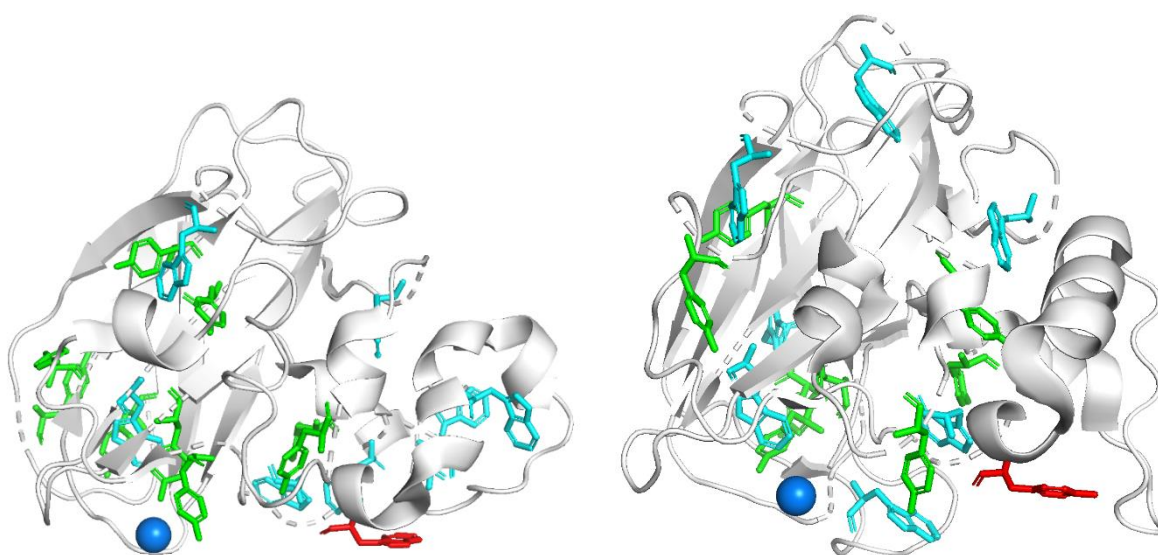


Figure 1.6. Enzyme structures of ScLPMO10B (PDB: 4OY6 [5]) and ScLPMO10C (PDB: 4OY7) with the active-site copper (blue), conserved Trp and Tyr (red) and Trp and Tyr throughout the protein (cyan and green, respectively).

Two cellulose-active LPMOs (C1-oxidizing ScLPMO10C (Cels2) and a C1/C4 oxidizing ScLPMO10B) from *Streptomyces coelicolor* and a chitin-active LPMO (BILPMO10A) were studied to investigate the details of cellulose and chitin degradation. Trp 88 in ScLPMO10B and Tyr 79 in ScLPMO10C indicated in red (Figure 1.6) are conserved aromatic residues with their ring parallel to the putative cellulose binding surface [12]. The active-site copper and these conserved aromatic residues are positioned in structurally similar locations on the surface of each protein.

LPMOs are worthy of studying, since they enable the cleavage of glycosidic bonds which are not degradable by other hydrolytic enzymes such as endoglucanases and cellobiohydrolases [5]. For this reason, many of the current commercial cellulase mixtures contain LPMOs to enhance the efficiency of lignocellulosic biomass processing.

1.5. Roles of Redox-active Trp/Tyr in Electron Transfer

It is important to study the electron transfer in metalloproteins, since it allows us to gain deeper insights into many of the biological processes or catalytic cycles. Trp and Tyr are redox active amino acids which facilitate electron transfer. Monitoring the redox chemistry between the metal centers and the intermediates such as Trp or Tyr radicals can give information about the electron transfer pathway. Critical parameters that affect the rate of electron transfer are (1) Driving force, (2) Nuclear reorientation, and (3) Electronic coupling [21], [22].

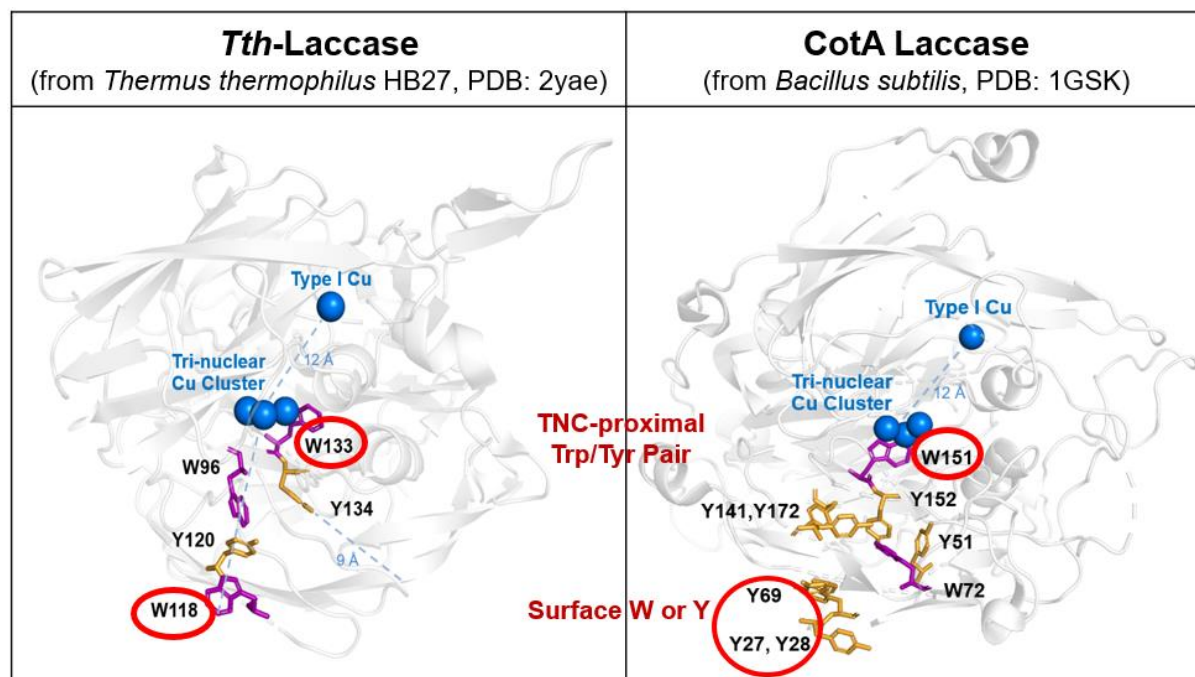


Figure 1.7. Structural homology between *Tth*-lac [10] and CotA-lac [11].

Trp and Tyr are commonly found near the active sites of the MCOs as well. One prominent feature in *Tth*-lac is a Trp/Tyr pair adjacent to the TNC within around 3.5 Å distance as well as a chain of closely spaced Trp and Tyr residues stretching from the TNC out to the surface (Figure 1.7, left). Similar Trp/Tyr chains are observed in CotA-lac as well (Figure 1.7, right). These chains of Trp and Tyr are actually notable features commonly occurring in most MCOs. When 25 X-ray crystal structures of three domain laccases existing in the protein data bank were examined, all of them except two had a Trp or a Trp/Tyr pair adjacent to the TNC [23]. We postulate that these residues are responsible for transferring the oxidizing equivalent to sites that are less vital for enzyme function. They may also contribute to providing multiple catalytic pathways supporting the substrate non-specific nature of these enzymes in nature. Tryptophan and tyrosine residues indeed occur with greater-than-average frequency in oxidoreductases: 66% and 81% have Tyr and Trp in above-average frequencies, respectively [24], and we speculate their crucial roles in enzyme function and survival.

1.6. References

- [1] K. J. Waldron, J. C. Rutherford, D. Ford, and N. J. Robinson, “Metalloproteins and metal sensing,” *Nature*, vol. 460, no. 7257, pp. 823–830, Aug. 13, 2009, doi: 10.1038/nature08300.
- [2] E. I. Solomon, U. M. Sundaram, and T. E. Machonkin, “Multicopper oxidases and oxygenases,” *Chem. Rev.*, vol. 96, no. 7, pp. 2563–2605, 1996, doi: 10.1021/cr950046o.
- [3] E. I. Solomon *et al.*, “Copper active sites in biology,” *Chemical Reviews*, vol. 114, no. 7, American Chemical Society, pp. 3659–3853, Apr. 09, 2014, doi: 10.1021/cr400327t.
- [4] Y. Zhu, X. Ouyang, Y. Zhao, L. Jiang, H. Guo, and X. Qiu, “Oxidative depolymerization of lignin improved by enzymolysis pretreatment with laccase,” *J. Energy Chem.*, vol. 27, no. 3, pp. 801–805, May 2018, doi: 10.1016/j.jechem.2017.04.018.
- [5] Z. Forsberg *et al.*, “Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 23, pp. 8446–8451, Jun. 2014, doi: 10.1073/pnas.1402771111.
- [6] A. Zoghalmi and G. Paës, “Lignocellulosic Biomass: Understanding Recalcitrance and Predicting Hydrolysis,” *Frontiers in Chemistry*, vol. 7, Frontiers Media S.A., p. 874, Dec. 18, 2019, doi: 10.3389/fchem.2019.00874.
- [7] M. N. S. Kumar, A. K. Mohanty, L. Erickson, and M. Misra, “Lignin and its applications with polymers,” *J. Biobased Mater. Bioenergy*, vol. 3, no. 1, pp. 1–24,

- Mar. 2009, doi: 10.1166/jbmb.2009.1001.
- [8] K. Nakamura and N. Go, “Function and molecular evolution of multicopper blue proteins,” *Cellular and Molecular Life Sciences*, vol. 62, no. 18. Springer, pp. 2050–2066, Sep. 09, 2005, doi: 10.1007/s00018-004-5076-x.
- [9] H. B. Gray and J. R. Winkler, “The Rise of Radicals in Bioinorganic Chemistry,” *Israel Journal of Chemistry*, vol. 56, no. 9–10. Wiley-VCH Verlag, pp. 640–648, Oct. 01, 2016, doi: 10.1002/ijch.201600069.
- [10] H. Serrano-Posada, S. Centeno-Leija, S. P. Rojas-Trejo, C. Rodríguez-Almazán, V. Stojanoff, and E. Rudiño-Piñera, “X-ray-induced catalytic active-site reduction of a multicopper oxidase: Structural insights into the proton-relay mechanism and O₂-reduction states,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 71, no. Pt 12, pp. 2396–2411, Nov. 2015, doi: 10.1107/S1399004715018714.
- [11] F. J. Enguita, L. O. Martins, A. O. Henriques, and M. A. Carrondo, “Crystal structure of a bacterial endospore coat component: A laccase with enhanced thermostability properties,” *J. Biol. Chem.*, vol. 278, no. 21, pp. 19416–19425, May 2003, doi: 10.1074/jbc.M301251200.
- [12] Z. Forsberg *et al.*, “Comparative study of two chitin-active and two cellulose-active AA10-type lytic polysaccharide monooxygenases,” *Biochemistry*, vol. 53, no. 10, pp. 1647–1656, Mar. 2014, doi: 10.1021/bi5000433.
- [13] M. Frommhagen *et al.*, “Boosting LPMO-driven lignocellulose degradation by polyphenol oxidase-activated lignin building blocks,” *Biotechnol. Biofuels*, vol. 10, no. 1, p. 121, May 2017, doi: 10.1186/s13068-017-0810-4.
- [14] K. E. H. Frandsen *et al.*, “The molecular basis of polysaccharide cleavage by lytic

- polysaccharide monooxygenases,” *Nat. Chem. Biol.*, vol. 12, no. 4, pp. 298–303, Apr. 2016, doi: 10.1038/nchembio.2029.
- [15] B. Bissaro *et al.*, “Oxidative cleavage of polysaccharides by monocopper enzymes depends on H₂O₂,” *Nat. Chem. Biol.*, vol. 13, no. 10, pp. 1123–1128, Oct. 2017, doi: 10.1038/nchembio.2470.
- [16] G. R. Hemsworth, G. J. Davies, and P. H. Walton, “Recent insights into copper-containing lytic polysaccharide mono-oxygenases,” *Current Opinion in Structural Biology*, vol. 23, no. 5. Elsevier Current Trends, pp. 660–668, Oct. 01, 2013, doi: 10.1016/j.sbi.2013.05.006.
- [17] P. H. Walton and G. J. Davies, “On the catalytic mechanisms of lytic polysaccharide monooxygenases,” *Current Opinion in Chemical Biology*, vol. 31. Elsevier Ltd, pp. 195–207, Apr. 01, 2016, doi: 10.1016/j.cbpa.2016.04.001.
- [18] G. R. Hemsworth *et al.*, “The copper active site of CBM33 polysaccharide oxygenases,” *J. Am. Chem. Soc.*, vol. 135, no. 16, pp. 6069–6077, Apr. 2013, doi: 10.1021/ja402106e.
- [19] C. H. Kjaergaard *et al.*, “Spectroscopic and computational insight into the activation of O₂ by the mononuclear Cu center in polysaccharide monooxygenases,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 24, pp. 8797–8802, 2014, doi: 10.1073/pnas.1408115111.
- [20] M. Frommhagen, A. H. Westphal, W. J. H. van Berkel, and M. A. Kabel, “Distinct substrate specificities and electron-donating systems of fungal lytic polysaccharide monooxygenases,” *Frontiers in Microbiology*, vol. 9, no. MAY. Frontiers Media S.A., p. 1080, May 29, 2018, doi: 10.3389/fmicb.2018.01080.

- [21] H. B. Gray and J. R. Winkler, “Long-range electron transfer,” *Proc. Natl. Acad. Sci.*, vol. 102, no. 10, pp. 3534–3539, Mar. 2005, doi: 10.1073/PNAS.0408029102.
- [22] J. R. Winkler and H. B. Gray, “Electron Flow through Metalloproteins,” 2013, doi: 10.1021/CR4004715.
- [23] H. B. Gray and J. R. Winkler, “Living with Oxygen,” *Acc. Chem. Res.*, vol. 51, no. 8, pp. 1850–1857, Aug. 2018, doi: 10.1021/acs.accounts.8b00245.
- [24] J. R. Winkler and H. B. Gray, “Could tyrosine and tryptophan serve multiple roles in biological redox processes?,” *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, vol. 373, no. 2037, Mar. 2015, doi: 10.1098/rsta.2014.0178.

Chapter 2

Stability/Activity Tradeoffs in *Thermus thermophilus* HB27 Laccase¹

Reprinted with permission from Springer Nature Customer Service Centre GmbH:

J. Shin, H. B. Gray, and J. R. Winkler, “Stability/activity tradeoffs in *Thermus thermophilus* HB27 laccase,” *J. Biol. Inorg. Chem.*, vol. 25, no. 2, 2020, doi: 10.1007/s00775-020-01754-7.

Copyright 2020

¹**Abbreviations:** *Thermus thermophilus* HB27 laccase, *Tth*-lac (PDB ID: 2YAE); *Cucumis sativus* cucumber basic protein (CPB, 2CBP); *Spinacea oleracea* spinach basic protein (SBP, 1F56); *Armoracia laphatifolia* umecyanin (UmCy, 1X9R); *Cucumis sativus* stellacyanin (StCy, 1JER); *Spinach oleracea* plastocyanin (PlCy, 1AG6); *Alcaligenes faecalis* azurin (AfAz, 2IAA); *Pseudomonas aeruginosa* azurin (PaAz, 5AZU); *Trametes versicolor* laccase (Trv-lac, 1GYC).

Stability/Activity Tradeoffs in *Thermus thermophilus* HB27 Laccase²

2.1. Abstract

We report the temperature dependence of the formal potential of type 1 copper (Cu_{T1}) in *Thermus thermophilus* HB27 laccase. Employing $[\text{Ru}(\text{NH}_3)_4(\text{bpy})](\text{PF}_6)_2$ (0.505 vs. NHE) as the redox titrant, we found that the $\text{Cu}_{\text{T1}}^{2+/+}$ potential decreased from approximately 480 to 420 mV (vs. NHE) as the temperature was raised from 20 to 65°C. Of importance is that the $\Delta S^\circ_{\text{rc}}$ of $-120 \text{ J mol}^{-1} \text{ K}^{-1}$ is substantially more negative than those for other blue copper proteins. We suggest that the highly unfavorable reduction entropy is attributable to Cu_{T1} inaccessibility to the aqueous medium. Although the active site residues are buried, which is critical for maintaining thermostability, the flexibility around Cu_{T1} is maintained, allowing enzyme activity at ambient temperature.

2.2. Introduction

The effects of temperature on enzyme survival and function have greatly influenced studies of enzyme evolution [1, 2]. The enzyme of interest in this study is a laccase from a thermophilic bacterium, *Thermus thermophilus* HB27 (*Tth*-lac), which grows optimally at 65°C. Laccases, members of the multicopper oxidase (MCO) family, have four copper sites involved in electron transfer and enzyme catalysis. According to the consensus mechanism [3, 4], substrate oxidation occurs near a type 1 copper (Cu_{T1}), followed by long-range electron transfer

²**Abbreviations:** *Thermus thermophilus* HB27 laccase, *Tth*-lac (PDB ID: 2YAE); *Cucumis sativus* cucumber basic protein (CPB, 2CBP); *Spinacea oleracea* spinach basic protein (SBP, 1F56); *Armoracia laphatifolia* umecyanin (UmCy, 1X9R); *Cucumis sativus* stellacyanin (StCy, 1JER); *Spinacea oleracea* plastocyanin (PlCy, 1AG6); *Alcaligenes faecalis* azurin (AfAz, 2IAA); *Pseudomonas aeruginosa* azurin (PaAz, 5AZU); *Trametes versicolor* laccase (Trv-lac, 1GYC).

to a trinuclear copper cluster, where, in a fully reduced enzyme, dioxygen is converted to water ($\text{O}_2 + 4\text{H}^+ + 4\text{e}^- \rightarrow 2\text{H}_2\text{O}$). Of special interest is that laccases are capable of degrading lignin in recalcitrant lignocellulosic substrates as their primary function [5], even though the Cu_{T1} potentials in the bacterial enzymes [4] are as much as 0.5 V lower than that required for one-electron oxidation of polyphenols [6].

Tth-lac is stable and active above 90°C for catalysis of aerobic oxidation of small polyphenolic substrates as well as ABTS (2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic) acid), a common substrate for oxidative enzymes [7, 8]. In contrast to many other thermophilic enzymes [2, 9], *Tth*-lac is active at ambient temperature, although the catalytic efficiency (defined by the turnover number divided by the Michaelis constant, $k_{\text{cat}}/K_{\text{m}}$) for ABTS oxidation increases substantially with increasing temperature (Figure 1). We anticipate that investigations of thermodynamic parameters will enhance the understanding of the remarkable stabilities and activities of extremophiles at elevated temperatures.

Surprisingly, the effects of temperature on the potentials of both the substrates and the active-site coppers have largely been neglected. As reduction entropies of $\text{M}^{(\text{ox/red})}$ redox couples are proportional to $(Z_{\text{ox}}^2 - Z_{\text{red}}^2)$, where Z_{ox} and Z_{red} are the charges of oxidized and reduced states [10], $\Delta S^\circ_{\text{rc}}$ for $\text{ABTS}^{-/2-}$ is predicted to be negative as observed for other $\text{M}^{-/2-}$ redox couples. It follows that the $\text{ABTS}^{-/2-}$ potential, which is about 670 mV (vs. NHE) at ambient temperature [11], is expected to decrease with increasing temperature. Based only on the decrease in ABTS potential, the activity of the enzyme likely would be higher at elevated temperatures. However, the charge dependence of $\Delta S^\circ_{\text{rc}}$ is only a small-molecule phenomenon, and $\Delta S^\circ_{\text{rc}}$ for proteins is unpredictable. Since the enzymatic activity also depends on the Cu_{T1} potential [12], we have investigated the temperature dependence of the Cu_{T1} potential in *Tth*-lac.

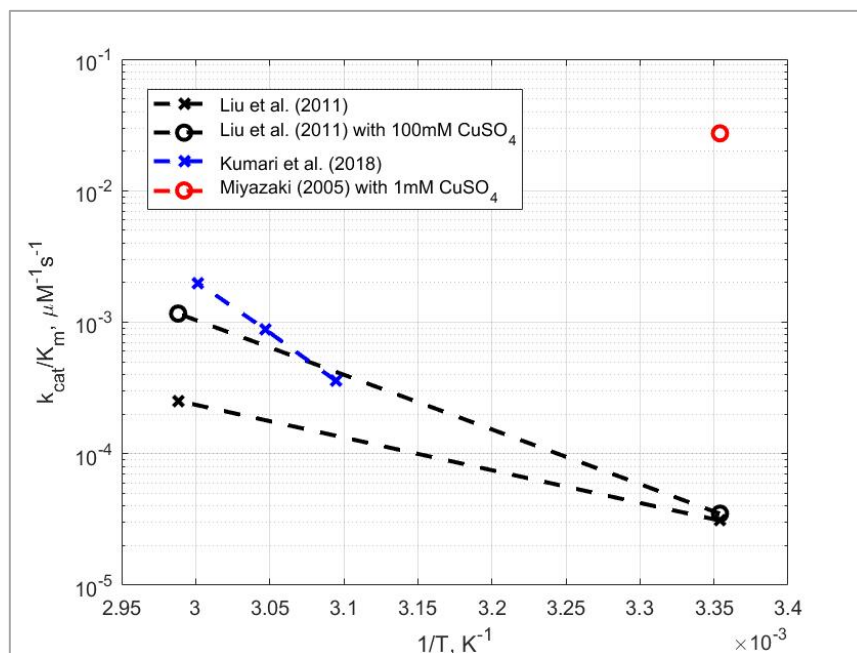


Figure 2.1. Temperature dependence of the catalytic efficiency of *Tth*-lac oxidation of substrate ABTS [7, 8, 13].

2.3. Methods

Sample Preparation

Thermus thermophilus HB27 laccase was expressed in *E. coli* and purified following published methods [7] with slight modifications. (See the Supplementary Material (SM) for details on protein expression and purification protocols.) A two-step procedure with slight modifications was employed to ensure full metalation of the enzyme [8, 14]. The enzyme was metalated in 20 mM Tris buffer, pH 8 with 1 mM CuSO₄; and it was metalated again in 25 mM sodium acetate buffer, pH 6 with 1 mM CuSO₄ for at least 48 h. The enzyme was stored with excess copper at 4 °C until use, and the excess copper was removed by gradual buffer exchange to 20 mM MES buffer at pH 5.3 with multiple rounds of centrifugal spin filtration. The presence of a type 2 copper which has the smallest binding constant in the protein was confirmed with EPR

and ICP-MS. $[\text{Ru}(\text{NH}_3)_4(\text{bpy})](\text{PF}_6)_2$ was synthesized and characterized following published protocols (with slight modifications) [15].

Circular Dichroism Spectra

Circular dichroism (CD) spectra of the protein samples under N_2 were recorded from 260 to 190 nm to monitor conformational changes and/or thermal denaturation over the temperature range 20 to 65°C. Measurements were made on 3 μM protein in 20 mM sodium phosphate buffer, pH 6. CD spectra reflecting the combined profiles of β -sheets, α -helices and random coils were obtained, and the spectra at different temperatures looked very similar (Figure S2.1.), consistent with prior observations [8].

Estimation of the Cu_{T1} Potential from Redox Equilibria

$\text{Cu}_{\text{T1}}^{2+/+}$ reduction potentials over the range 20 to 65°C were estimated by monitoring changes in UV-vis spectra of a deoxygenated sample containing the wild type protein and $[\text{Ru}(\text{NH}_3)_4(\text{bpy})](\text{PF}_6)_2$. The formal $\text{Ru}^{3+/2+}$ potential of $[\text{Ru}(\text{NH}_3)_4(\text{bpy})](\text{PF}_6)_2$ was reported to be 0.505 V [16], and $\Delta S^\circ_{\text{rc}}$ is approximately +56 $\text{J mol}^{-1} \text{K}^{-1}$ [10]. A previous report estimated the *Tth*-lac $\text{Cu}_{\text{T1}}^{2+/+}$ potential to be approximately 0.5 V vs NHE at pH 5 and 6.5 [17].

Wild type *Tth*-lac (60 μM) and four equivalents (240 μM) of $[\text{Ru}(\text{NH}_3)_4(\text{bpy})](\text{PF}_6)_2$ were deoxygenated by gentle evacuation/Ar-backfill cycles and then mixed together in a sealed quartz cuvette. UV-vis spectra of the sample were monitored at temperatures from 20 to 65°C. The spectra of $[\text{Ru}(\text{NH}_3)_4(\text{bpy})]^{2+}$ and the wild type $\text{Cu}_{\text{T1}}^{2+}$ protein were monitored separately as functions of temperature. We assume that the Cu_{T1}^+ protein and $[\text{Ru}(\text{NH}_3)_4(\text{bpy})]^{3+}$ do not make substantial contributions to spectra in the 400-700 nm range. Equilibrium concentrations

of $\text{Cu}_{\text{T1}}^{2+}$ and Ru^{2+} in the mixed sample were determined by least squares decomposition of the mixed spectrum (480-650 nm) into a linear combination of the two component spectra.

2.4. Results and Discussion

Temperature Dependence of the Cu_{T1} Formal Potential

Based on redox equilibria determined by titration with $[\text{Ru}(\text{NH}_3)_4(\text{bpy})](\text{PF}_6)_2$, the Cu_{T1} formal potential decreased by approximately 60 mV (480 to 420 mV) from 20 to 65°C. As temperature variations also affect buffer pH (the temperature coefficient for the pH of MES buffer is approximately -0.011 per °C [18]), the pH is predicted to decrease from 5.3 to 4.8 with a temperature increase from 20 to 65°C; and a pH decrease of 0.5 could produce an apparent increase in the $\text{Cu}_{\text{T1}}^{2+/+}$ potential at elevated temperature, even though solvent accessibility to the deeply buried copper site is low. It follows that the decrease in $\text{Cu}_{\text{T1}}^{2+/+}$ potential extracted from redox titration data at the higher temperature would be slightly greater if corrected for the pH change.

$\text{Cu}_{\text{T1}}^{2+/+}$ formal potentials in blue copper proteins range from approximately 200 to 800 mV and usually tend to decrease with increasing temperature (i.e., $\Delta S^\circ_{\text{rc}} < 0$) [19, 20]. The enthalpic contribution ($-\Delta H^\circ_{\text{rc}}/F$) to the potential is largely determined by ligand interactions with Cu_{T1} . A copper site with a weak axial bond has a higher formal potential, owing mainly to destabilization of the oxidized state [21]. The Cu_{T1} sites in high-potential laccases have a trigonal planar Cu_{T1} geometry with noncoordinating Phe or Leu residues in axial positions; and low-potential laccases have an axial Met residue in the inner coordination sphere. The entropic contribution ($T\Delta S^\circ_{\text{rc}}/F$) to the potential, on the other hand, is influenced by interactions with the protein scaffold and the surrounding aqueous medium.

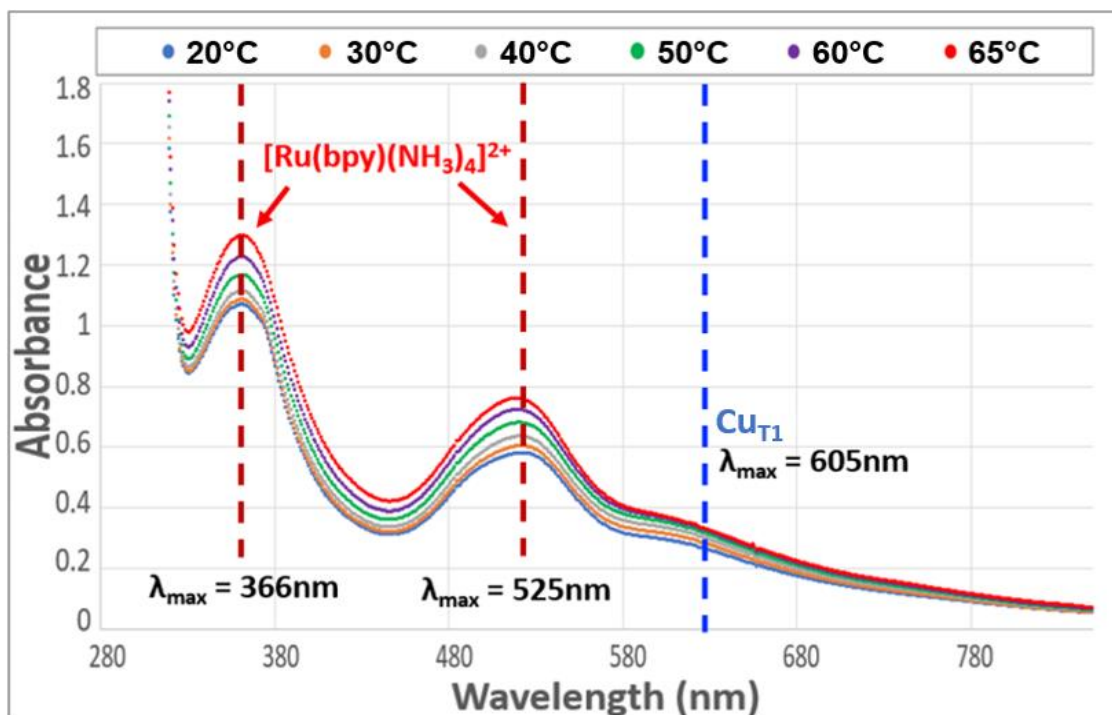
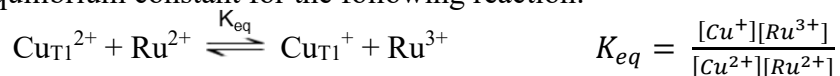


Figure 2.2. UV-vis spectra (20 to 65°C) of a deoxygenated sample containing both wild type *Tth*-lac and four equivalents of $[\text{Ru}(\text{NH}_3)_4(\text{bpy})](\text{PF}_6)_2$.

Table 2.1. Effect of temperature on the reaction between $\text{Cu}_{\text{T1}}^{2+}$ and $[\text{Ru}(\text{NH}_3)_4(\text{bpy})]^{2+}$.

T (K)	293	303	313	323	333	338
K_{eq}^a	0.33(1) ^b	0.218(7)	0.123(5)	0.061(3)	0.034(4)	0.025(7)
ΔG° (kJ mol ⁻¹)	2.72(9)	3.8(1)	5.5(2)	7.5(4)	9.4(9)	10(2)
ΔE° (mV)	-28(1)	-40(1)	-56(2)	-78(4)	-100(10)	-110(10)
$E^\circ(\text{Ru}^{3+/2+})$ (mV) <i>c,d</i>	505	511	517	522	528	531
$E^\circ(\text{Cu}^{2+/+})$ (mV) ^c	477(1)	471(1)	460(2)	445(4)	430(10)	420(10)

^a Equilibrium constant for the following reaction:



^b Estimated uncertainties in the last digit appear in parentheses. Error estimates provided by least-squares analyses of the data. [See the SM for details.]

^c Formal potential vs. NHE.

^d Reference 10.

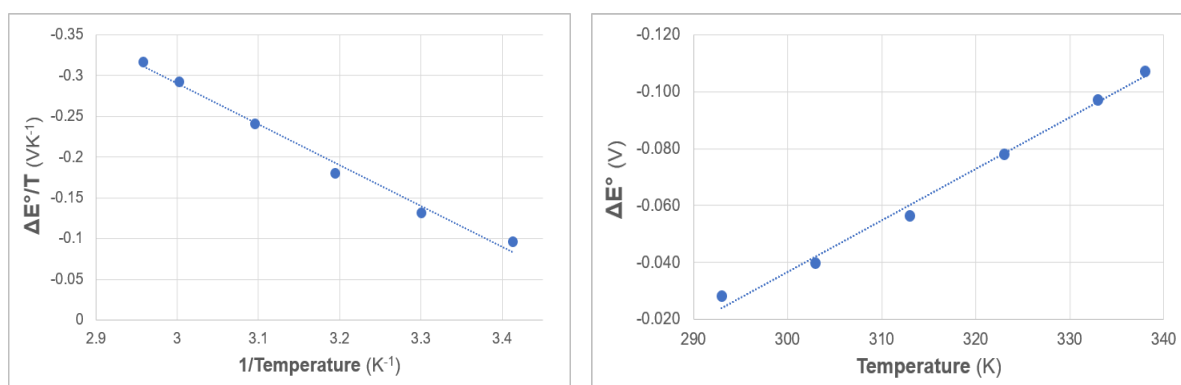


Figure 2.3. Plots of $\Delta E^\circ/T$ vs $1/T$ (left) and ΔE° vs T (right) provide values of $\Delta H^\circ = -48 \text{ kJ mol}^{-1}$ and $\Delta S^\circ = -175 \text{ J mol}^{-1} \text{ K}^{-1}$ for the reaction of $Tth\text{-lac Cu}_{\text{T1}}^{2+}$ with $[\text{Ru}(\text{NH}_3)_4(\text{bpy})]^{2+}$. The standard reduction enthalpy ($\Delta H^\circ_{\text{rc}}(\text{Cu}^{2+/+}) = -81 \text{ kJ mol}^{-1}$, NHE reference) and entropy ($\Delta S^\circ_{\text{rc}}(\text{Cu}^{2+/+}) = -120 \text{ J mol}^{-1} \text{ K}^{-1}$) changes associated with $\text{Cu}_{\text{T1}}^{2+}$ reduction are estimated by adding $\Delta H^\circ_{\text{rc}}(\text{Ru}^{3+/2+})$ and $\Delta S^\circ_{\text{rc}}(\text{Ru}^{3+/2+})$ [10] to ΔH° and ΔS° , respectively.

The $Tth\text{-lac}$ $\Delta H^\circ_{\text{rc}}$ is comparable to that of other blue copper proteins [19, 20], but the standard reduction entropy change is much more negative (Figure 2.3). It is notable that the *Polyporous versicolor* laccase $\Delta H^\circ_{\text{rc}}$ ($-73.1 \text{ kJ mol}^{-1}$) is close to that of $Tth\text{-lac}$ (-81 kJ mol^{-1}), but $\Delta S^\circ_{\text{rc}}$ for *P. versicolor* laccase ($+7.1 \text{ J mol}^{-1} \text{ K}^{-1}$) is much more positive, accounting for its high potential (780 mV vs. NHE) [19]. As a consequence of an unfavorable $\Delta S^\circ_{\text{rc}}$ ($-120 \text{ J mol}^{-1} \text{ K}^{-1}$), the potential of $Tth\text{-lac}$ is much lower (Table 2.1).

The reduction entropy change for small molecules depends primarily on the ionic charge, owing to the smaller entropy associated with solvent polarization for more highly charged ions. For redox centers buried inside proteins, the contribution to $\Delta S^\circ_{\text{rc}}$ from solvent polarization will likely be attenuated, and the response of the polypeptide matrix to a change in oxidation state will become more important. We anticipate, then, that the large negative $\Delta S^\circ_{\text{rc}}$ value for

Tth-lac may be attributable to a combination of reduced solvent exposure and a more hydrophobic environment around the Cu_{T1} active site [19]. Hydrophobicity [22] and polarity [23] indices for all residues within 8 Å of Cu_{T1} were summed to characterize redox-center environments in nine different proteins, including *Tth-lac*. Of interest is that ΔS°_{rc} is not strongly correlated with either parameter (see Tables S2.2.-S2.3., Figures S2.4.-S2.6.).

Calculations of solvent accessible surface areas (SASA) of Cu_{T1} ligands provided quantitative estimates of redox-site exposure to aqueous solvent (Table S2.4.). The relative solvent accessibility (RSA) for each residue is defined as the SASA normalized by a maximum allowed SASA [24]. Values of SASA and RSA for Cu_{T1} sites in nine proteins with known thermodynamic properties are given in Table 2.2. Notably, there is an apparent correlation between reduction entropy (ΔS°_{rc}) and SASA: copper proteins with smaller SASA tend to exhibit more negative entropy changes upon reduction (Figure 2.4.). From our analysis, we have confirmed that reduced solvent accessibility around the metal site is one of the critical parameters affecting ΔS°_{rc} . Note that in *Tth-lac*, only 0.04% of His393 and 1.4% of C445 are exposed to the aqueous medium. It is clear that the solvent accessibility of Cu_{T1} ligands in *Tth-lac* is lower than in other blue copper proteins: ligand exposures are approximately 20% in proteins with more positive reduction entropies, a group including CBP, SBP, UmCy, and StCy. RSA values exhibit a similar correlation with ΔS°_{rc} as shown in the SM (Figure S2.7). Similar trends are found in a prior investigation of cytochrome redox thermodynamics which showed the correlation between ΔS°_{rc} and solvent exposure of the heme [25].

Table 2.2. Reduction entropy change (ΔS_{rc}° , $\text{J mol}^{-1} \text{K}^{-1}$) and measures of solvent accessibility of copper ligands in blue copper proteins (values of ΔS_{rc}° except for *Tth-lac* are from [19, 20] measured in the 5-45 °C temperature range).

	ΔS_{rc}° ($\text{J mol}^{-1} \text{K}^{-1}$)	SASA (\AA^2)	RSA (%)
CPB	31	39.8	18
SBP	7	41.8	19
UmCy	-17	47.8	22
StCy	-21	52.4	24
Trv-lac	-29	18.6	9
PlCy	-36	28.9	13
AfAz	-58	2.4	1.1
PaAz	-68	1.9	0.9
Tth-lac	-120	2.1	1.4

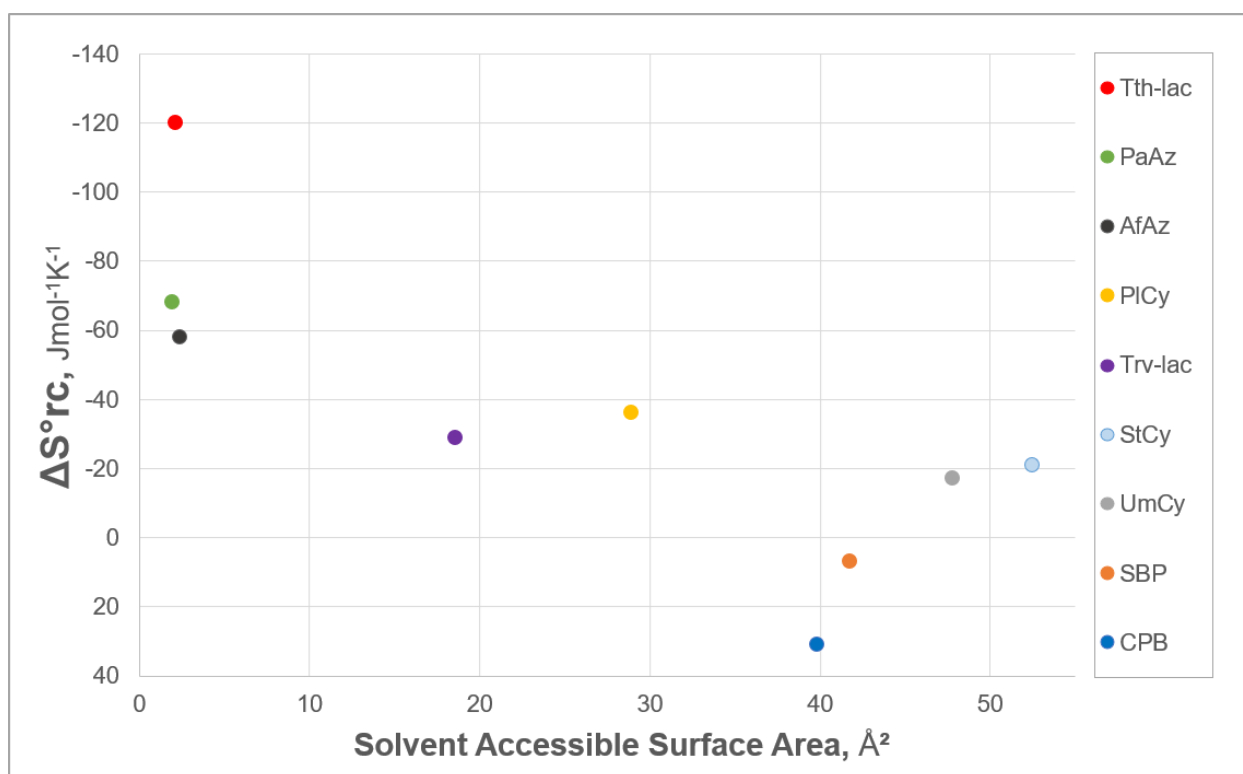


Figure 2.4. Plot of the $\text{Cu}_{\text{TI}}^{2+/+}$ reduction entropy (ΔS_{rc}° , $\text{J mol}^{-1} \text{K}^{-1}$) versus the solvent accessible surface area (\AA^2) of active-site binding residues in blue copper proteins.

The limited exposure of *Tth*-lac Cu_{T1} active site minimizes the solvent contribution to the reduction entropy change. The large negative ΔS_{rc}° observed for *Tth*-lac Cu_{T1} then likely arises from a substantial reduction in conformational entropy of the protein upon reduction. Acrylamide quenching of Trp fluorescence in oxidized *Tth*-lac (10-35 °C), an indicator of protein flexibility, is much greater than expected for a thermophilic enzyme [2, 9, 26]. We suggest that the structural flexibility of oxidized *Tth*-lac balances the molecular motions required for enzyme turnover while maintaining sufficiently low reorganization to support intra-protein electron transfer, thereby accounting for the observed oxidase activity at 25 °C [26].

If the Cu_{T1} active site ($[\text{Cu}(\text{N}_{\text{His}})_2(\text{S}_{\text{Cys}})(\text{S}_{\text{Met}})]^{2+/+}$) were free in aqueous solution, ΔS_{rc}° would be positive owing to the more positive charge on the oxidized complex [10]. When the polypeptide of *Tth*-lac folds around this active site, the value of ΔS_{rc}° drops precipitously. A thermodynamic cycle indicates that the folding entropy change for the reduced protein is substantially more negative than that of the oxidized enzyme. The two primary contributions to the folding entropy change are a polypeptide conformational component (ΔS_{conf}) and a hydration entropy change (ΔS_{hyd}) arising from encapsulation of hydrophobic residues in the interior of the protein upon folding to minimize contact with aqueous media [27]. The conformational entropy change upon folding is negative, whereas hydration makes a positive contribution to the total folding entropy change. We anticipate little difference in ΔS_{hyd} between the oxidized and reduced enzymes, suggesting that ΔS_{conf} must be substantially more unfavorable for folding around the reduced enzyme. The more negative value of ΔS_{conf} is consistent with less flexibility in the reduced enzyme.

It is notable that *Tth*-lac exhibits characteristics of both hot and cold adaptations: thermal stability (Figure S2.1.) and flexibility [26]. The robust global structure of *Tth*-lac disfavors protein unfolding and copper loss at elevated temperatures [26]. We infer from the small value of SASA that Cu_{T1} in *Tth*-lac is more buried inside the protein scaffold, and the burial of Cu_{T1} in *Tth*-lac provides protection and stability at the cost of a highly unfavorable reduction entropy. The large negative $\Delta S^{\circ}_{\text{rc}}$ value for Cu_{T1} *Tth*-lac likely will lead to highly negative ΔS° values in substrate oxidation reactions, owing to entropically disfavored formation of cationic substrates [10]. At low reaction driving forces, the activation entropy for electron-transfer (ET) reactions is [28]:

$$\Delta S^{\ddagger} = \frac{\Delta S^{\circ}}{2} - R\beta d$$

where R is the gas constant, β is the exponential distance-decay factor for electron transfer (1.1 \AA^{-1} , [29]), and d is the electron donor-acceptor separation. Negative values of ΔS° will have the same impact on ET rates as increasing the donor-acceptor distance by about 0.05 \AA per entropy unit ($\text{J mol}^{-1} \text{ K}^{-1}$), corresponding to a factor of 10 decrease in rate constant for each $-40 \text{ J mol}^{-1} \text{ K}^{-1}$ of reaction entropy change. Conversely, oxidation of *Tth*-lac Cu_{T1}⁺, as occurs during enzyme turnover, is likely to be accompanied by a favorable entropy change, which could compensate for the weak coupling associated with the long distance between Cu_{T1}⁺ and the trinuclear Cu active site.

2.5. Concluding Remarks

Thermophilic metalloenzymes tend to have active sites buried in a matrix of hydrophobic residues. We suggest that a tightly packed polypeptide scaffold limits solvent access to the active site, allowing the stability of *Tth*-lac at elevated temperatures. The reduced exposure of Cu_{T1} in *Tth*-lac, which minimizes the solvent contribution to the reduction entropy change, accounts for the large negative contribution by the polypeptide. The unfavorable entropy change upon reduction may be a consequence of unexpected flexibility in the oxidized protein. Regardless of its structural origins, the large Cu_{T1} redox entropy will have a substantial impact on electron-transfer kinetics and, hence, on enzyme activity. Our studies indicate that the remarkable capacity of thermophilic enzymes to remain active at extremely high temperatures is attributable to a subtle balance of many competing dynamic and thermodynamic factors. Understanding that interplay will be helpful in broadening the scope and utility of enzymes in industrial applications.

2.6. Acknowledgment

Research reported in this publication was supported by the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health under award number R01DK019038. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

2.7. References

1. Kumar A, Alam A, Tripathi D, Rani M, Khatoon H, Pandey S, Ehtesham NZ, Hasnain SE (2018) Protein adaptations in extremophiles: An insight into extremophilic connection of mycobacterial proteome. *Semin Cell Dev Biol* 84: 147-157. <https://doi.org/10.1016/j.semcdb.2018.01.003>
2. Feller G (2010) Protein stability and enzyme activity at extreme biological temperatures. *J Phys Condens Matter* 22(32):323101. <https://doi.org/10.1088/0953-8984/22/32/323101>
3. Galli I, Musci G, Bonaccorsi di Patti MC (2004) Sequential reconstitution of copper sites in the multicopper oxidase CueO. *J Biol Inorg Chem* 9(1):90-5. <https://doi.org/10.1007/s00775-003-0501-4>
4. Solomon EI, Sundaram UM, Machonkin TE (1996) Multicopper Oxidases and Oxygenases. *Chem. Rev* 96, 7, 2563-2605. <https://doi.org/10.1021/cr950046o>
5. Davidi L, Morais S, Artzi L, Knop D, Hadar Y, Arfi Y, Bayer EA (2016) Toward combined delignification and saccharification wheat straw by a laccase-containing designer cellulosome. *PNAS* 113(39) 10854-10859. <https://doi.org/10.1073/pnas.1608012113>
6. Lind J, Shen X, Eriksen TE, Merenyi G (1990) The one-electron reduction potential of 4-substituted phenoxy radicals in water. *J. Am. Chem. Soc* 112, 2, 479-482. <https://doi.org/10.1021/ja00158a002>
7. Miyazaki, K (2005) A hyperthermophilic laccase from *Thermus thermophilus* HB27. *Extremophiles* 9(6):415-425. <https://doi.org/10.1007/s00792-005-0458-z>
8. Liu X, Gillespie M, Ozel AD, Dikici E, Daunert S, Bachas LG (2011) Electrochemical properties and temperature dependence of a recombinant laccase from *Thermus thermophilus*. *Anal Bioanal Chem* 399(1): 361-366. <https://doi.org/10.1007/s00216-010-4345-9>

9. Vieille C, Zeikus GJ (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol Mol Biol Rev* 65(1):1–43. <https://doi.org/10.1128/MMBR.65.1.1-43.2001>
10. Hupp J, Weaver M (1984) Solvent, Ligand, and Ionic Charge Effects on Reaction Entropies for Simple Transition-Metal Redox Couples. *Inorg. Chem* 23, 22, 3639-3644. <https://doi.org/10.1021/ic00190a042>
11. Bourbonnais R, Leech D, Paice MG (1998) Electrochemical analysis of the interactions of laccase mediators with lignin model compounds. *Biochim Biophys Acta* 1379(3): 381-90. [https://doi.org/10.1016/s0304-4165\(97\)00117-7](https://doi.org/10.1016/s0304-4165(97)00117-7)
12. Xu F (1996) Oxidation of Phenols, Anilines, and Benzenethiols by Fungal Laccases: Correlation between Activity and Redox Potentials as Well as Halide Inhibition. *Biochemistry* 35, 23, 7608-7614. <https://doi.org/10.1021/bi952971a>
13. Kumari A, Kishor N, Guptasarma P (2018) Characterization of a mildly alkalophilic and thermostable recombinant *Thermus thermophilus* laccase with applications in decolourization of dyes. *Biotechnol Lett* 40(2):285-295. <https://doi.org/10.1007/s10529-017-2461-8>
14. Serrano-Posada H, Centeno-Leija S, Rojas-Trejo SP, [Rodríguez-Almazán C](#), [Stojanoff V](#), [Rudiño-Piñera E](#) (2015) X-ray-induced catalytic active-site reduction of a multicopper oxidase: structural insights into the proton-relay mechanism and O₂-reduction states. *Acta Crystallogr D Biol Crystallogr* 71(pt12): 2396-2411. <https://doi.org/10.1107/S1399004715018714>
15. Curtis JC, Sullivan BP, Meyer TJ. (1983) Hydrogen-Bonding-Induced Solvatochromism in the Charge-Transfer Transitions of Ruthenium(II) and Ruthenium(III) Ammine Complexes. *Inorg. Chem* 22, 2, 224-236. <https://doi.org/10.1021/ic00144a009>
16. Matsubara T and Ford PC (1976) Some Applications of Cyclic Voltammetry to the

- Reactions and Properties of Ruthenium Ammine Complexes. Reduction Potentials and Rate Studies. *Inorg. Chem* 1976. 15, 5: 1107-1110. <https://doi.org/10.1021/ic50159a025>
17. Agbo P, Heath JR, Gray HB (2013) Catalysis of Dioxygen Reduction by *Thermus thermophilus* Strain HB27 Laccase on Ketjen Black Electrodes. *J Phys Chem B* 117(2): 527-534. <https://doi.org/10.1021/jp309759g>
18. Good NE, Winget D, Winter W, Connolly TN, Izawa S, Singh RMM (1966) Hydrogen Ion Buffers for Biological Research. *Biochemistry* 5, 2, 467-477. <https://doi.org/10.1021/bi00866a011>
19. Battistuzzi G, Borsari M, Loschi L, Righi F, Sola M (1999) Redox Thermodynamics of Blue Copper Proteins. *J. Am. Chem. Soc.* 121, 3, 501-506. <https://doi.org/10.1021/ja982126q>
20. Battistuzzi G, Bellei M, Leonardi A, Pierattelli R, De Candia A, Vila AJ, Sola M (2005) Reduction thermodynamics of the T1 Cu site in plant and fungal laccases. *J Biol Inorg Chem* 10(8): 867-873. <https://doi.org/10.1007/s00775-005-0035-z>
21. Gray HB, Malmström BG Williams RJP (2000) Copper coordination in blue proteins. *J. Biol. Inorg. Chem.* 5(5):551-559. <https://doi.org/10.1007/s007750000146>
22. Nozaki Y and Tanford C (1971) The solubility of amino acids and to glycine peptides in aqueous ethanol and dioxane solutions. Establishment of a hydrophobicity scale. *J Biol Chem* 246(7):2211-7.
23. Zimmerman JM, Eliezer N, Simha R (1968) The characterization of amino acid sequences in proteins by statistical methods. *J Theor Biol* 21(2) 170-201. [https://doi.org/10.1016/0022-5193\(68\)90069-6](https://doi.org/10.1016/0022-5193(68)90069-6)
24. Tien MZ, Meyer AG, Sydykova DK, Spielman SJ, Wilke CO (2013) Maximum Allowed Solvent Accessibilities of Residues in Proteins. *PLoS ONE* 8(11):e80635. <https://doi.org/10.1371/journal.pone.0080635>

25. Taniguchi VT, Ellis JR WR, Cammarata V, Webb J, Anson FC, Gray HB (1982) Spectroelectrochemical Determination of the Temperature Dependence of Reduction Potentials Tris(1,10-phenanthroline) Complexes of Iron and Cobalt with c-Type Cytochromes. *Advances in Chemistry*. 201, 51-68. <https://doi.org/10.1021/ba-1982-0201.ch003>
26. Roulling F, Godin A, Cipolla A, Collins T, Miyazaki K, Feller G (2016) Activity-stability relationships revisited in blue oxidases catalyzing electron transfer at extreme temperatures. *Extremophiles*. 20(5): 621-9. <https://doi.org/10.1007/s00792-016-0851-9>
27. [Fitter J \(2003\) A Measure of Conformational Entropy Change during Thermal Protein Unfolding Using Neutron Spectroscopy. *Biophys. J.* 84\(6\):3924-3930. \[https://doi.org/10.1016/S0006-3495\\(03\\)75120-0\]\(https://doi.org/10.1016/S0006-3495\(03\)75120-0\)](https://doi.org/10.1016/S0006-3495(03)75120-0)
28. Marcus RA, Sutin N (1985) Electron transfers in chemistry and biology. *Biochimica et Biophysica Acta*. 811(3). 265-322. [https://doi.org/10.1016/0304-4173\(85\)90014-X](https://doi.org/10.1016/0304-4173(85)90014-X)
29. Winkler JR, Gray HB (2014) Long-range electron tunneling. *J. Am. Chem. Soc.* 136(8):2930-2939. <https://doi.org/10.1021/ja500215j>

Supplementary Material

Stability/Activity Tradeoffs in *Thermus thermophilus* Laccase

Reprinted with permission from Springer Nature Customer Service Centre GmbH:

J. Shin, H. B. Gray, and J. R. Winkler, “Stability/activity tradeoffs in *Thermus thermophilus* HB27 laccase,” *J. Biol. Inorg. Chem.*, vol. 25, no. 2, 2020, doi: 10.1007/s00775-020-01754-7.

S2.1. Protein Preparation

Tth-lac was expressed in *E. coli* and purified following the published method [1] with slight modifications. To sum up, the plasmid was prepared by cloning the *Tth*-lac encoding region into the pET22b expression vector at NdeI digestion site. *Tth*-lac with a N-terminal 6 x His-tag was transformed into BL21(DE3) cells on a LB/Amp agar plate. Single colonies grown on the plate were inoculated in 5mL TB/Amp at 37°C for 5-7 hours (starter culture), and the frozen stock was made for later uses. The cells were pelleted and re-inoculated in 4-6L of TB/Amp/0.4% glycerol at 37°C overnight. Protein expression was achieved after induction with IPTG at 37°C for 7-8 hours. After the induction, cells were harvested by centrifugation and were kept at -20°C until use.

The frozen cell pellets were thawed and the cells were resuspended in 20mM Tris buffer at pH 8 with the addition of protease inhibitors (Complete Mini Protease Inhibitor Cocktail Tablets, Pefabloc SC (AEBSF), and Benzamidine hydrochloride hydrate). The enzyme was released by sonicating the cells for an hour, and the sample was centrifuged to remove all the cell debris. The supernatant solution after centrifugation was further purified by heating up to around 65°C for 20 min. At this temperature, majority of other *E. coli* enzymes crash out due to thermal denaturation, and only *Tth*-lac remains in the solution. By centrifugation, all the precipitates were removed. Since the recombinant *Tth*-laccase has the N-terminal 6 x His-tag right after the start codon, the nickel immobilized metal affinity column was used to bind His-tag labeled protein to the column. All the other remnant of undesired *E. Coli* proteins can be eluted first with wash and load buffers, and the laccase gets eluted as the elution buffer is run down the column with increasing concentration of imidazole. A cation exchange column (HiPrep SP HP

16/10 column) was used to further remove impurities by selectively collecting the fraction of interest by monitoring the absorption at 280nm and 605nm after the sample was metalated.

S2.2. Circular Dichroism Spectra

Circular dichroism spectra of the protein samples under N₂ were recorded on an Aviv model 430 circular dichroism spectrometer from 260 to 190 nm to monitor conformational changes and/or thermal denaturation over the temperature range 20 to 65°C. Measurements were made on 3 μM protein in 20 mM sodium phosphate buffer, pH 6.

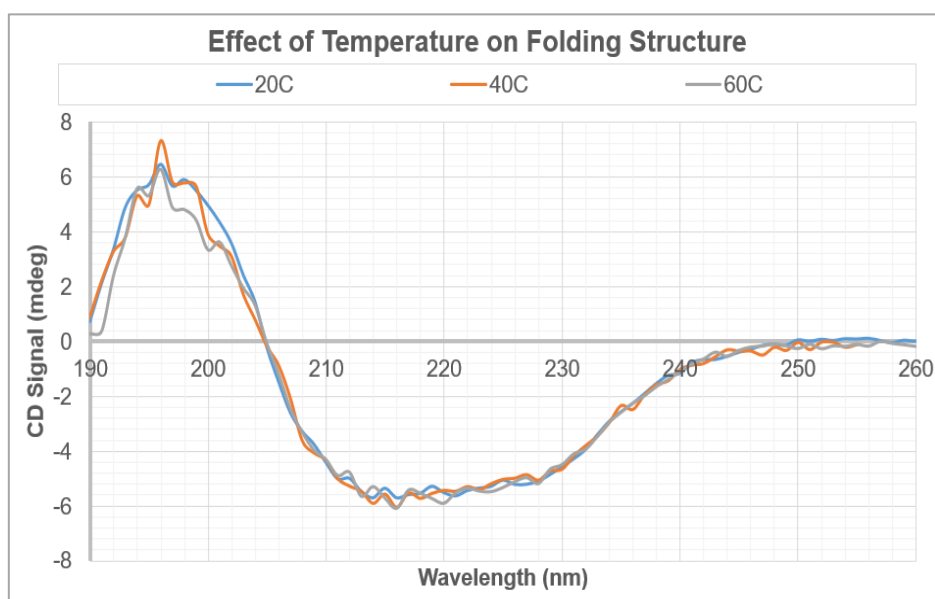


Figure S2.1. Circular dichroism spectra of *Tth*-lac from 20 to 65°C.

S2.3. Synthesis of $[\text{Ru}(\text{NH}_3)_4(\text{bpy})](\text{PF}_6)_2$

$[\text{Ru}(\text{NH}_3)_4(\text{bpy})](\text{PF}_6)_2$ was synthesized and characterized following the protocol (with slight modifications) [2]. Zinc amalgam was made with mercury chloride (HgCl_2) and mossy zinc, and $[\text{Ru}(\text{NH}_3)_5\text{Cl}]\text{Cl}_2$ was purchased from STREM chemicals INC. All the procedures were performed under argon using a Schlenk-line techniques.

S2.4. Estimation of the Cu_{T1} Potential from Redox Equilibria

60 μM wild-type *Tth*-laccase sample and the four reducing equivalents (240 μM) $[\text{Ru}(\text{NH}_3)_4(\text{bpy})](\text{PF}_6)_2$ were deoxygenated by gentle vacuum/argon pump and backfill cycles, and were mixed together in a sealed quartz cuvette. The UV-Vis spectra of the mixed sample were monitored from 20 to 65°C in the absorption regions at 366, 525, and 605 nm (the λ_{max} values of $[\text{Ru}(\text{NH}_3)_4(\text{bpy})](\text{PF}_6)_2$ and the T1 copper, respectively). Equilibrium concentrations of $\text{Cu}_{\text{T1}}^{2+}$ and Ru^{2+} in the mixed sample were determined by least squares decomposition of the mixed spectrum into a linear combination of the two component spectra. The spectra of $[\text{Ru}(\text{NH}_3)_4(\text{bpy})](\text{PF}_6)_2$ and the wild-type protein also were monitored separately at different temperatures to make sure that the temperature dependent changes in the spectra of mixed samples result from redox processes involving the two components.

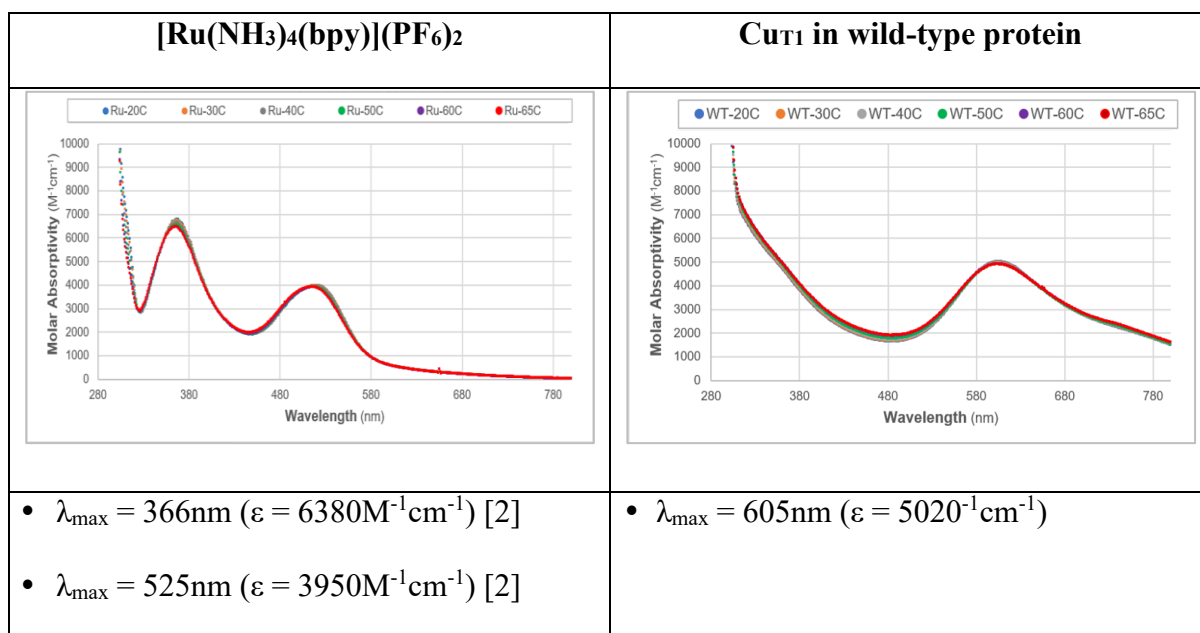


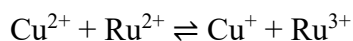
Figure S2.2. UV-Vis spectra (20 to 65 °C) for deoxygenated samples of [Ru(NH₃)₄(bpy)](PF₆)₂ and of wild type *Tth*-lac.

Table S2.1. Temperature dependent equilibrium concentrations of Cu¹⁺, Cu²⁺, Ru²⁺, and Ru³⁺

	293K	303K	313K	323K	333K	338K
[Cu ²⁺] (uM)	44.6	48.3	52.1	55.1	56.7	57.1
[Cu ¹⁺] (uM)	15.4	11.7	7.9	4.9	3.3	2.9
[Ru ²⁺] (uM)	123.1	126.5	132.6	142.7	152.0	160.1
[Ru ³⁺] (uM)	117.0	113.5	107.4	97.3	88.0	79.9

$$A_{525} = \epsilon_{\text{Ru}525}[\text{Ru}^{2+}] + \epsilon_{\text{Cu}525}[\text{Cu}^{2+}]$$

$$A_{605} = \epsilon_{\text{Ru}605}[\text{Ru}^{2+}] + \epsilon_{\text{Cu}605}[\text{Cu}^{2+}]$$



$$K_{eq} = \frac{[\text{Cu}^+][\text{Ru}^{3+}]}{[\text{Cu}^{2+}][\text{Ru}^{2+}]}$$

$$K_{eq} = \exp\left\{\frac{-\Delta G^\circ}{RT}\right\}$$

$$\Delta G^\circ = -RT \ln K$$

$$\Delta E^\circ = -\frac{\Delta G^\circ}{F} = E^\circ(\text{Cu}^{2+}/+) - E^\circ(\text{Ru}^{3+}/2+)$$

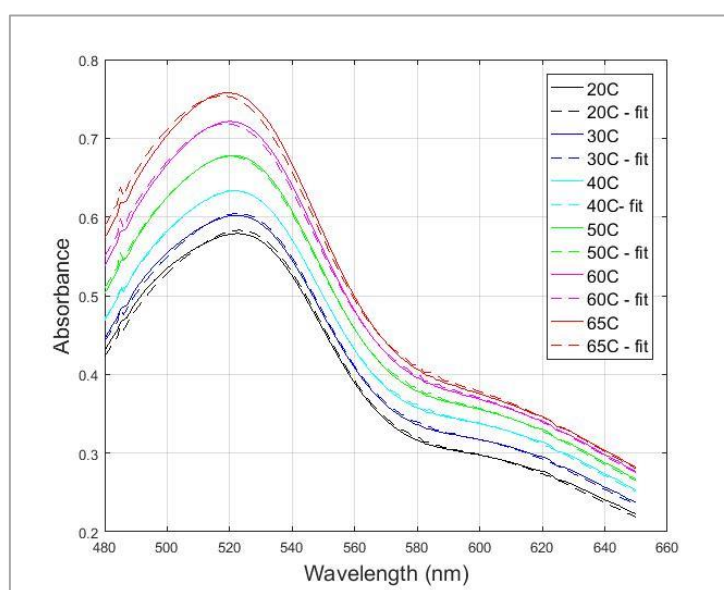


Figure S2.3. Equilibrium concentrations of $[\text{Ru}^{2+}]$, $[\text{Ru}^{3+}]$, $[\text{Cu}^{2+}]$, and $[\text{Cu}^+]$ were obtained from global fitting of the redox titration data.

Spectra of the wild-type protein and $[\text{Ru}(\text{NH}_3)_4(\text{bpy})](\text{PF}_6)_2$ obtained every 10 °C from 20 to 60 °C (also at 65 °C) monitor concentration changes at the different temperatures. As molar absorptivities at investigated wavelengths depend on temperature for both Cu_{T1} and $[\text{Ru}(\text{NH}_3)_4(\text{bpy})](\text{PF}_6)_2$, global fitting was done for mixed-sample spectra to estimate equilibrium concentrations of Ru^{2+} and Cu^{2+} .

Error Propagation

$$K_{eq} = \frac{([\text{Cu}^{2+}]_0 - [\text{Cu}^{2+}]) * ([\text{Ru}^{2+}]_0 - [\text{Ru}^{2+}])}{([\text{Cu}^{2+}] * [\text{Ru}^{2+}])}$$

- All concentrations are in μM .
- $[\text{Cu}^{2+}]_0$ = initial concentration of *Tth*-lac
- $[\text{Cu}^{2+}]$ = final concentration of *Tth*-lac
- $[\text{Ru}^{2+}]_0$ = initial concentration of $[\text{Ru}(\text{NH}_3)_4(\text{bpy})](\text{PF}_6)_2$
- $[\text{Ru}^{2+}]$ = final concentration of $[\text{Ru}(\text{NH}_3)_4(\text{bpy})](\text{PF}_6)_2$
- lb uncertainty: lower bound uncertainty (95% confidence interval)
- ub uncertainty: upper bound uncertainty (95% confidence interval)

Temperature (K)	293	303	313	323	333	338
$[\text{Cu}^{2+}]_0$	6.00E-05	6.00E-05	6.00E-05	6.00E-05	6.00E-05	6.00E-05
$[\text{Cu}^{2+}]_0$ lb uncertainty	1.00E-07	1.00E-07	1.00E-07	1.00E-07	1.00E-07	1.00E-07
$[\text{Cu}^{2+}]_0$ ub uncertainty	1.00E-07	1.00E-07	1.00E-07	1.00E-07	1.00E-07	1.00E-07
$[\text{Cu}^{2+}]$	4.46E-05	4.83E-05	5.21E-05	5.51E-05	5.67E-05	5.71E-05
$[\text{Cu}^{2+}]$ lb uncertainty	2.03E-07	1.64E-07	7.00E-08	8.40E-08	1.96E-07	3.11E-07
$[\text{Cu}^{2+}]$ ub uncertainty	1.97E-07	1.36E-07	1.30E-07	1.16E-07	2.04E-07	2.89E-07
$[\text{Ru}^{2+}]_0$	2.40E-04	2.40E-04	2.40E-04	2.40E-04	2.40E-04	2.40E-04
$[\text{Ru}^{2+}]_0$ lb uncertainty	5.00E-07	5.00E-07	5.00E-07	5.00E-07	5.00E-07	5.00E-07
$[\text{Ru}^{2+}]_0$ ub uncertainty	5.00E-07	5.00E-07	5.00E-07	5.00E-07	5.00E-07	5.00E-07
$[\text{Ru}^{2+}]$	1.23E-04	1.27E-04	1.33E-04	1.43E-04	1.52E-04	1.60E-04
$[\text{Ru}^{2+}]$ lb uncertainty	3.50E-07	2.10E-07	1.10E-07	2.10E-07	3.60E-07	5.40E-07
$[\text{Ru}^{2+}]$ ub uncertainty	3.50E-07	2.90E-07	1.90E-07	1.90E-07	3.40E-07	4.60E-07
$([\text{Cu}^{2+}]_0 - [\text{Cu}^{2+}])$	1.54E-05	1.17E-05	7.93E-06	4.92E-06	3.30E-06	2.89E-06
$([\text{Cu}^{2+}]_0 - [\text{Cu}^{2+}])$ lb uncertainty	3.03E-07	2.64E-07	1.70E-07	1.84E-07	2.96E-07	4.11E-07
$([\text{Cu}^{2+}]_0 - [\text{Cu}^{2+}])$ ub uncertainty	2.97E-07	2.36E-07	2.30E-07	2.16E-07	3.04E-07	3.89E-07
$([\text{Ru}^{2+}]_0 - [\text{Ru}^{2+}])$	1.17E-04	1.13E-04	1.07E-04	9.73E-05	8.80E-05	7.99E-05
$([\text{Ru}^{2+}]_0 - [\text{Ru}^{2+}])$ lb uncertainty	8.50E-07	7.10E-07	6.10E-07	7.10E-07	8.60E-07	1.04E-06
$([\text{Ru}^{2+}]_0 - [\text{Ru}^{2+}])$ ub uncertainty	8.50E-07	7.90E-07	6.90E-07	6.90E-07	8.40E-07	9.60E-07
$([\text{Cu}^{2+}]_0 - [\text{Cu}^{2+}]) * ([\text{Ru}^{2+}]_0 - [\text{Ru}^{2+}])$	1.80E-09	1.33E-09	8.52E-10	4.78E-10	2.91E-10	2.31E-10
$([\text{Cu}^{2+}]_0 - [\text{Cu}^{2+}]) * ([\text{Ru}^{2+}]_0 - [\text{Ru}^{2+}])$ lb uncertainty	4.85E-11	3.83E-11	2.31E-11	2.14E-11	2.89E-11	3.58E-11
$([\text{Cu}^{2+}]_0 - [\text{Cu}^{2+}]) * ([\text{Ru}^{2+}]_0 - [\text{Ru}^{2+}])$ ub uncertainty	4.78E-11	3.61E-11	3.02E-11	2.44E-11	2.95E-11	3.38E-11

Temperature (K)	293	303	313	323	333	338
$([Cu^{2+}][Ru^{2+}])$	5.49E-09	6.11E-09	6.91E-09	7.86E-09	8.62E-09	9.15E-09
$([Cu^{2+}][Ru^{2+}])$ lb uncertainty	4.06E-11	3.09E-11	1.50E-11	2.36E-11	5.02E-11	8.06E-11
$([Cu^{2+}][Ru^{2+}])$ ub uncertainty	3.99E-11	3.12E-11	2.71E-11	2.70E-11	5.03E-11	7.26E-11
K_{eq}	3.28E-01	2.18E-01	1.23E-01	6.08E-02	3.38E-02	2.52E-02
K_{eq} lb uncertainty	1.13E-02	7.37E-03	3.61E-03	2.90E-03	3.55E-03	4.14E-03
K_{eq} ub uncertainty	1.11E-02	7.02E-03	4.85E-03	3.31E-03	3.63E-03	3.90E-03
$\ln(K_{eq})$	-1.11E+00	-1.52E+00	-2.09E+00	-2.80E+00	-3.39E+00	-3.68E+00
$\ln(K_{eq})$ lb uncertainty	-3.83E-02	-5.15E-02	-6.13E-02	-1.34E-01	-3.56E-01	-6.04E-01
$\ln(K_{eq})$ ub uncertainty	-3.77E-02	-4.90E-02	-8.24E-02	-1.52E-01	-3.64E-01	-5.69E-01
ΔG° (kJmol ⁻¹)	2.71E+00	3.84E+00	5.45E+00	7.52E+00	9.38E+00	1.03E+01
ΔG° lb uncertainty	9.42E-02	1.31E-01	1.61E-01	3.61E-01	9.90E-01	1.70E+00
ΔG° ub uncertainty	9.27E-02	1.25E-01	2.16E-01	4.12E-01	1.01E+00	1.60E+00
ΔE° (V)	-2.81E-02	-3.97E-02	-5.64E-02	-7.79E-02	-9.72E-02	-1.07E-01
ΔE° lb uncertainty	-9.76E-04	-1.36E-03	-1.67E-03	-3.74E-03	-1.03E-02	-1.76E-02
ΔE° ub uncertainty	-9.61E-04	-1.29E-03	-2.24E-03	-4.27E-03	-1.05E-02	-1.66E-02

S2.5. Thermodynamic Parameters

The reaction entropy can be described by the temperature dependence of the Gibbs function.

$$\left(\frac{\delta \Delta G^\circ}{\delta T}\right)_p = -\Delta S^\circ$$

The reduction entropy can be described as a function of the temperature variance of E° .

$$nF \left(\frac{\delta E^\circ}{\delta T}\right)_p = \Delta S^\circ$$

ΔS° : determined from the slope of the plot of ΔE° versus temperature

ΔH° : determined from the slope of the plot of $\Delta E^\circ/T$ versus the inverse of temperature ($1/T$)

$$\Delta S^\circ_{rc}(Cu^{2+/+}) = \Delta S^\circ + \Delta S^\circ_{rc}(Ru^{3+/2+})$$

S2.6. Hydrophobicity, Polarity and Solvent Accessible Surface Area (SASA)

Table S2.2. Hydrophobicity of residues within 8 Å from Cu_{T1}.

(Hydrophobicity index for each amino acid residue was adopted from reference [3].)

	Hydrophobicity of Residues within 8Å from Cu	Hydrophobicity Normalized by Total Number of Residues	ΔS°_{rc} (J mol ⁻¹ K ⁻¹)	E ^o (mV vs. NHE)
CPB	35.0	1.52	31	306
SBP	26.9	1.49	7	345
UmCy	33.0	1.50	-17	290
StCy	21.1	1.11	-21	265
Trv-lac	36.2	1.72	-29	738
PICy	34.4	1.38	-36	366
AfAz	33.9	1.47	-58	266
PaAz	35.9	1.44	-68	307
Tth-lac	37.3	1.43	-120	480

Table S2.3. Polarity of residues within 8Å from Cu_{T1}.

(Polarity index for each amino acid residue was adopted from reference [4].)

	Polarity of Residues within 8Å from Cu	Polarity Normalized by Total Number of Residues	ΔS°_{rc} (J mol ⁻¹ K ⁻¹)	E ^o (mV vs. NHE)
CPB	182	7.90	31	306
SBP	124	6.89	7	345
UmCy	369	16.8	-17	290
StCy	176	9.28	-21	265
Trv-lac	316	15.0	-29	738
PICy	126	5.04	-36	366
AfAz	178	7.73	-58	266
PaAz	229	9.16	-68	307
Tth-lac	477	18.3	-120	480

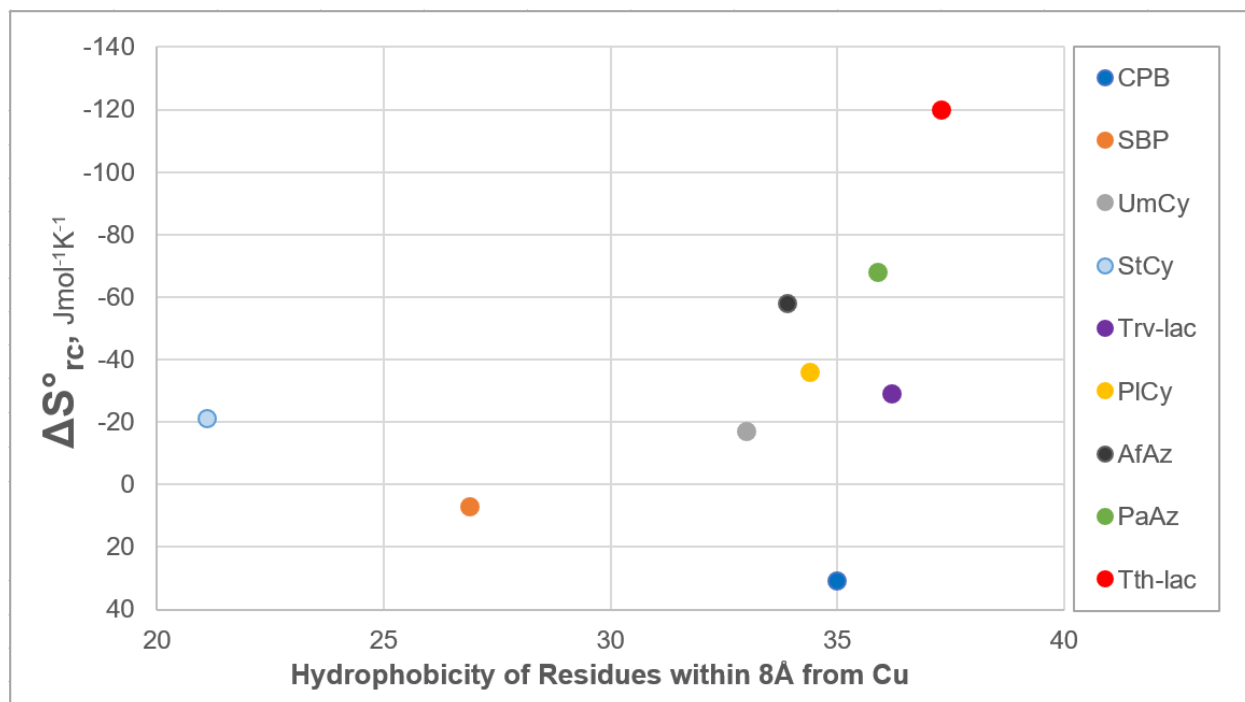


Figure S2.4. ΔS°_{rc} vs. Hydrophobicity of residues within 8 Å from Cu_{T1} .

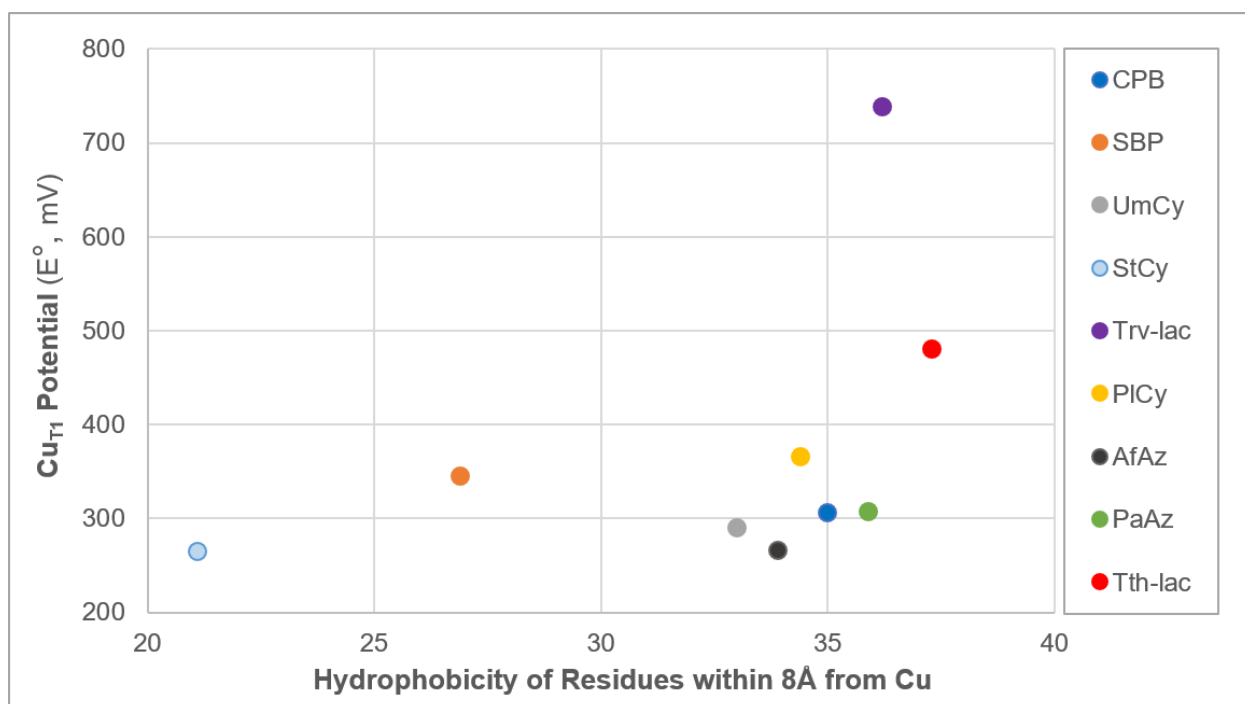


Figure S2.5. E° of Cu_{T1} vs. Hydrophobicity of residues within 8 Å from Cu_{T1} .

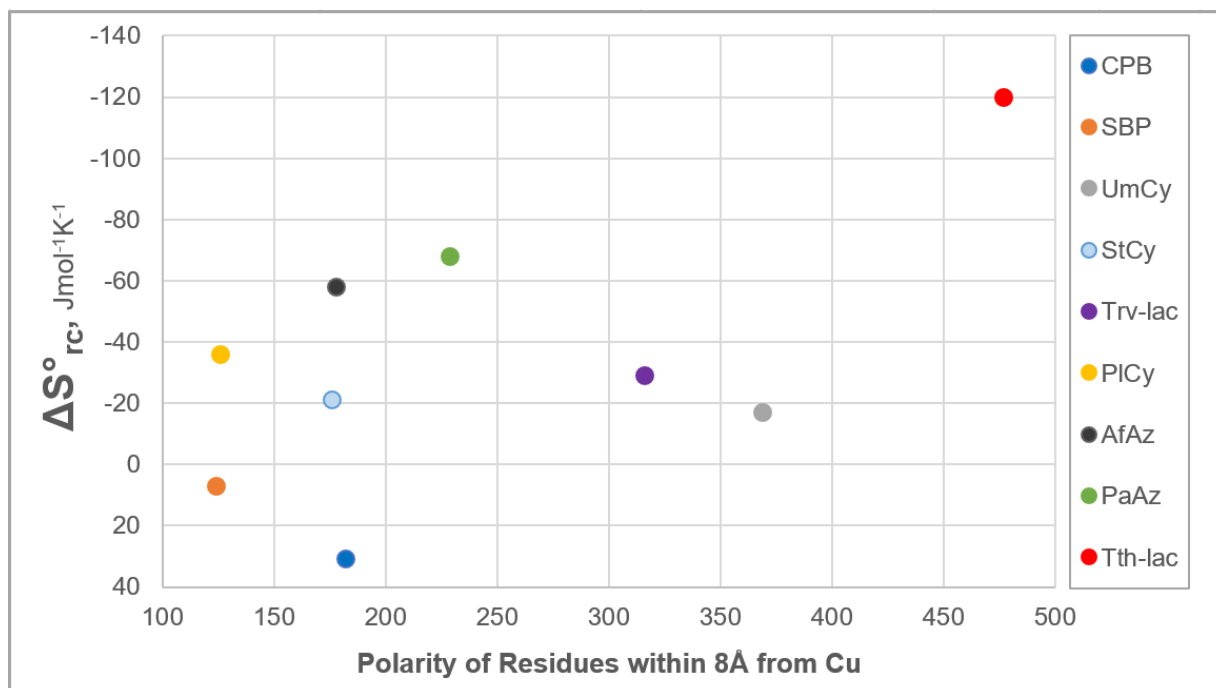


Figure S2.6. ΔS°_{rc} vs Polarity of residues within 8Å from Cu_{T1}

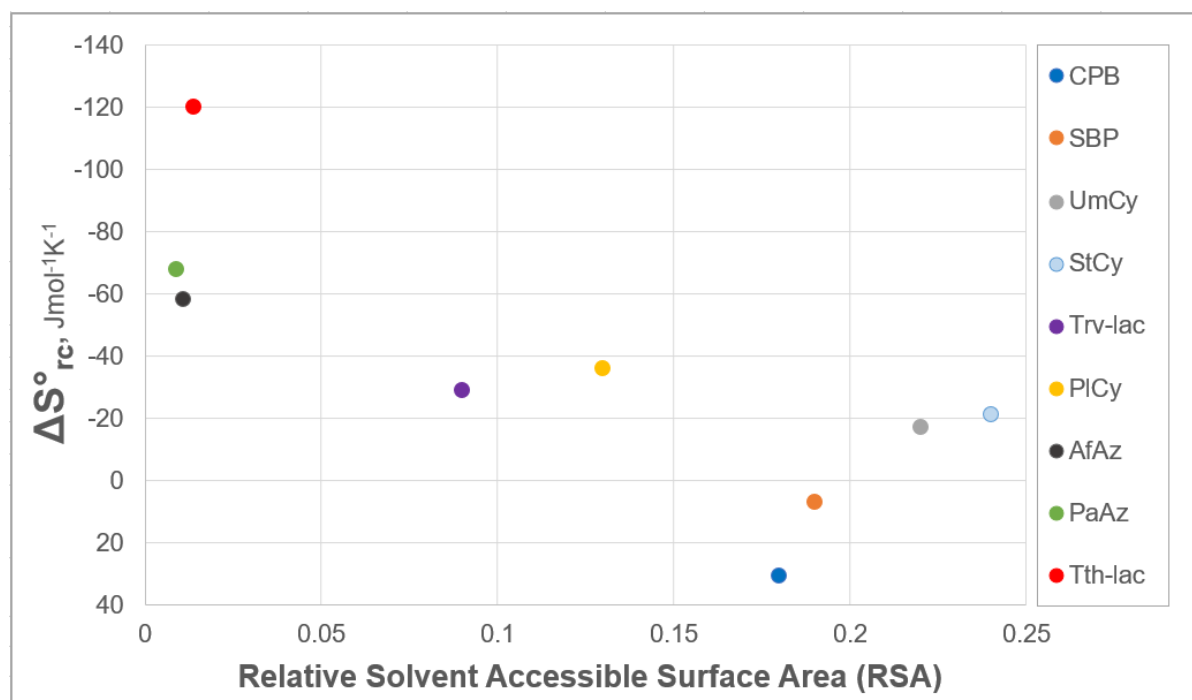


Figure S2.7. The correlation between reduction entropy (ΔS°_{rc} , J mol⁻¹K⁻¹) and relative solvent accessible surface area (RSA)

Table S2.4. The solvent accessible surface area (\AA^2) and the relative solvent accessible surface area (RSA) of copper ligands in blue copper proteins (calculated with FreeSASA software [5]).

(The maximum SASA values for blue copper ligands are: His = 216.0 \AA^2 , Cys = 148 \AA^2 , Met = 203.0 \AA^2 [6].)

CBP (<i>Cucumis sativus</i>), [PDB: 2CBP]			SBP (<i>Spinacea Oleracea</i>), [PDB: 1F56]			Umecyanin (<i>Armoracia laphatifolia</i>), [PDB: 1X9R]		
Residue	Solvent Accessible Surface Area (\AA^2)	RSA	Residue	Solvent Accessible Surface Area (\AA^2)	RSA	Residue	Solvent Accessible Surface Area (\AA^2)	RSA
H39	13.31	0.06	H34	15.03	0.07	H44	22.65	0.10
H84	26.54	0.12	H79	26.73	0.12	H90	25.04	0.12
C79	0	0.00	C74	0	0	C85	0.1	0.001
M89	0	0.00	M84	0	0	Q95	0	0.00
Total	39.85	0.18	Total	41.76	0.19	Total	47.79	0.22
Stellacyanin (<i>Cucumis sativus</i>), [PDB: 1JER]			Laccase (<i>Trametes Versicolor</i>), [PDB: 1GYC]			Plastocyanin (<i>Spinacea Oleracea</i>), [PDB: 1AG6]		
Residue	Solvent Accessible Surface Area (\AA^2)	RSA	Residue	Solvent Accessible Surface Area (\AA^2)	RSA	Residue	Solvent Accessible Surface Area (\AA^2)	RSA
H46	15.19	0.07	H395	5.32	0.02	H37	28.87	0.13
H94	37.11	0.17	H458	13.11	0.06	H87	0	0
C89	0.15	0.001	C453	0.12	0.001	C84	0	0
Q99	0	0	F463	0	0	M92	0.07	0.0003
Total	52.45	0.24	Total	18.55	0.09	Total	28.94	0.13
Azurin (<i>Alcaligenes faecalis</i>), [PDB: 2IAA]			Azurin (<i>Pseudomonas aeruginosa</i>), [PDB: 5AZU]			Laccase (<i>Thermus thermophilus HB27</i>), [PDB: 1JER]		
Residue	Solvent Accessible Surface Area (\AA^2)	RSA	Residue	Solvent Accessible Surface Area (\AA^2)	RSA	Residue	Solvent Accessible Surface Area (\AA^2)	RSA
H46	0.06	0.0003	H46	0.65	0.003	H393	0.08	0.0004
H117	1.82	0.0084	H117	1.08	0.005	H450	0	0
C112	0	0	C112	0.02	0.0001	C445	2.05	0.014
M121	0.54	0.0027	M121	0.16	0.0008	M455	0	0
Total	2.42	0.011	Total	1.91	0.009	Total	2.13	0.014

S2.7. References

1. Miyazaki, K (2005) A hyperthermophilic laccase from *Thermus thermophilus* HB27. *Extremophiles* 9(6):415-425. <https://doi.org/10.1007/s00792-005-0458-z>
2. Curtis JC, Sullivan BP, Meyer TJ. (1983) Hydrogen-Bonding-Induced Solvatochromism in the Charge-Transfer Transitions of Ruthenium(II) and Ruthenium(III) Ammine Complexes. *Inorg. Chem* 22, 2, 224-236. <https://doi.org/10.1021/ic00144a009>
3. Nozaki Y and Tanford C (1971) The solubility of amino acids and to glycine peptides in aqueous ethanol and dioxane solutions. Establishment of a hydrophobicity scale. *J Biol Chem* 246(7):2211-7.
4. Zimmerman JM, Eliezer N, Simha R (1968) The characterization of amino acid sequences in proteins by statistical methods. *J Theor Biol* 21(2) 170-201. [https://doi.org/10.1016/0022-5193\(68\)90069-6](https://doi.org/10.1016/0022-5193(68)90069-6)
5. Simon Mitternacht (2016) [FreeSASA: An open source C library for solvent accessible surface area calculation.](https://doi.org/10.12688/f1000research.7931.1) *F1000Research* 5:189 <https://doi.org/10.12688/f1000research.7931.1>
6. Tien MZ, Meyer AG, Sydykova DK, Spielman SJ, Wilke CO (2013) Maximum Allowed Solvent Accessibilities of Residues in Proteins. *PLoS ONE* 8(11):e80635. <https://doi.org/10.1371/journal.pone.0080635>

Chapter 3

Electron Transfer through Laccase from *Thermus thermophilus* HB27

3.1. Introduction

Since the enzyme turnover cycles happen within ns to μ s time scale, it is quite difficult to study what intermediates are being formed in the catalytic cycles. It would be great if we could monitor those intermediate species during turnover conditions, but even with stopped flow experiments and the freeze quench EPR, the fastest time scale we can monitor is still in the milli second range. Especially in the presence of closely spaced redox active Trp and Tyr along the chain, it is likely that the intermediates are not going to last that long. So, we needed a technique with which we could monitor species in the microsecond time scale, and transient absorption spectroscopy was one possibility.

Therefore, one of the main pillars of my thesis work centered on examining the hole/electron transfer pathways through laccase using laser flash/quench techniques. The oxidative quenching of Ru-photosensitizer-labeled *Tth*-lac enzymes was of primary interest to inject holes into the system and to see if the active site coppers can be oxidized by hole hopping through Trp and Tyr residues.

3.2. Laser Sample Preparation

Laccase of interest in this study is from a thermophilic strain of bacterium *Thermus thermophilus* HB27 (*Tth*-lac), and thus is stable and active even above 90°C [1], which opens a possibility for its application to fuel cell development. In order to study the electron transfer pathways through *Tth*-lac, two surface-exposed labeling sites (S117 and S102) near the trinuclear Cu cluster (TNC) and a labeling site (D390) near the Cu_{T1} was chosen (Figure 3.1). Ser117 is thought to be the best labeling site, since it is positioned next to the surface-exposed tryptophan which is at the end of a Trp/Tyr chain stretching all the way to the TNC.

Protein Expression

- Site-directed Mutagenesis

The mutations to the initial plasmid were made following the conventional protocol known as site-directed mutagenesis [2]. The primers (forward and reverse) which contain a codon specific to a point mutation of interest were designed to achieve mutations on the desired position. The polymerase chain reaction (PCR) was run for around 3.5 hours with the conditions optimized based on the T_m of each primer. After digesting the template with Dpn I at 37°C for 1 hour, the desired plasmid was transformed into NovaBlue cells. Cells were grown in the SOC medium, and were plated on the LB/Amp plates to allow inoculation of single colonies. The mutant plasmids were prepared by following the standard mini-prep procedures, and were sequenced to check the success of mutagenesis.

- BL21(DE3) cell

The desired plasmid was transformed to BL21(DE3) cells which are more adequate for protein expression. From the plate, a single colony was inoculated at 37°C for 4-6 hours to make a starter culture, and the frozen stock was made for later uses. The cells from the starter culture were inoculated in 1L TB/Amp at 37°C, and were grown for 16 hrs. Protein expression can be achieved by IPTG induction typically at 35°C for 7 hours. After induction, cells were harvested by centrifugation.

Protein Purification

- Sonication

The frozen cell pellets were thawed in an ice basket, and were resuspended in 20 mM Tris buffer at pH 8 with the addition of protease inhibitors (Complete mini protease inhibitor

cocktail tablets). Enzymes were released from the cells with sonication for 30 min (3 cycles of 10 min with pulsing), and the cell crude extract was centrifuged to remove all the broken cell debris.

- Heating Up

Since *Tth*-lac is a thermophilic enzyme, the supernatant solution from the above sonication step after centrifugation was further heated up to around 65 °C for 20 min. At this temperature, most other *E. coli* proteins crash out, and only the *Tth*-lac presumably remains in the solution. By centrifugation, all the precipitates were removed.

- Nickel Column

Since the expressed *Tth*-lac has a 6-His tag, the nickel immobilized metal affinity column was used for purification. All the other *E. coli* proteins remaining even after the heating step can be eluted first with wash and load buffers, and *Tth*-lac gets eluted as the elution buffer with 250 mM imidazole is run down the column.

- Wash buffer: 20mM Tris, pH 8
- Load buffer: 25mM Tris + 20mM Imidazole
- Elution buffer: 25mM Tris + 250mM Imidazole

Fractions were collected right after the buffer was switched to the elution buffer, and the activity of each fraction was checked with ABTS with the addition of excess copper to check for the presence of *Tth*-lac.

Surface labeling with a $[\text{Ru}(\text{bpy})_2(\text{Phen})\text{IA}]^{2+}$ complex

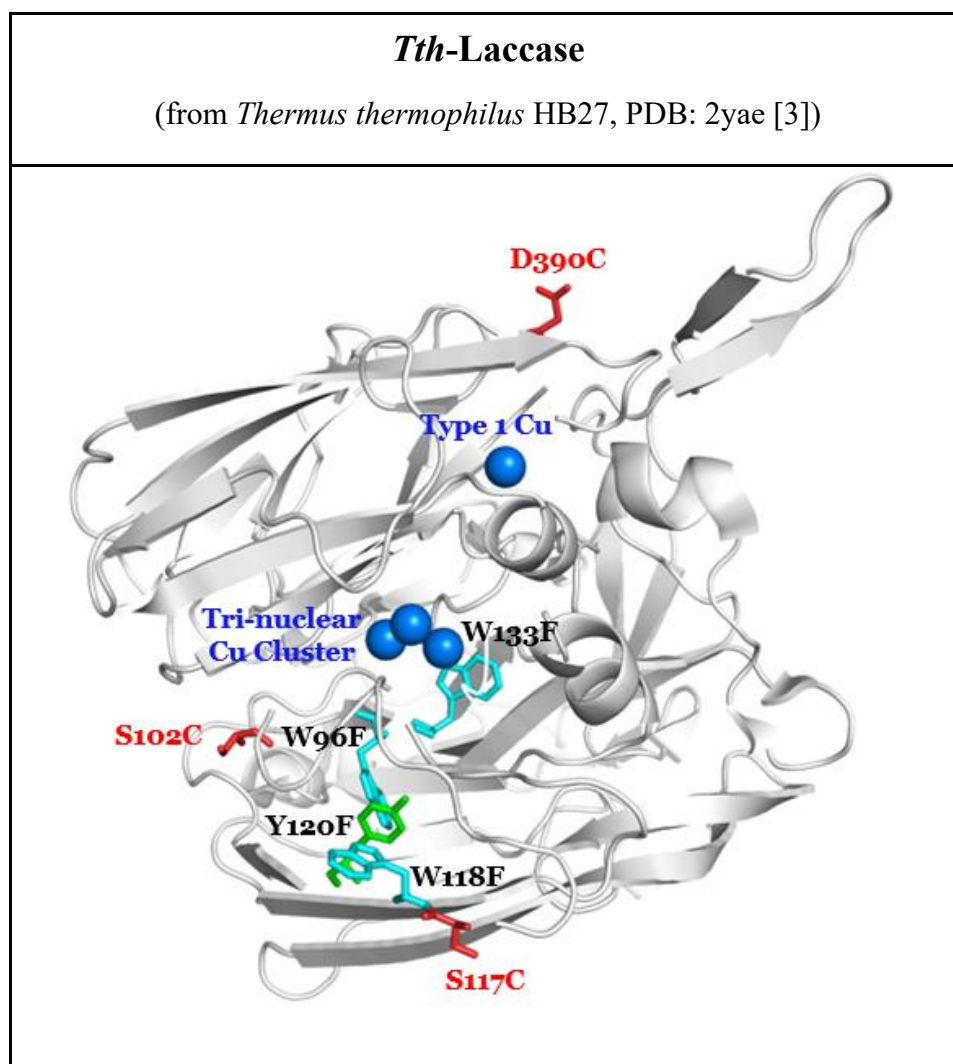


Figure 3.1. Surface labeling sites on *Tth*-lac (red): Ser 117, Ser 102 and Asp 390.

Synthesis of the Ru-photosensitizer



$[\text{Ru}(\text{bpy})_2(\text{IAphen})](\text{PF}_6)_2$ was synthesized following well-established protocols [4]. To briefly explain, $\text{Ru}(\text{bpy})_2\text{Cl}_2$ was prepared by refluxing $\text{RuCl}_3 \cdot 3\text{H}_2\text{O}$, bipyridine and LiCl in dimethylformamide (DMF) [5]. To make 5-Iodoacetamido-1,10-phenanthroline (phen-IA),

dicyclohexylurea was made by adding N,N'-dicyclohexylcarbodiimide (DCC) to iodoacetic acid in ethyl acetate. After the removal of urea from it, 5-amino-1,10-phenanthroline is obtained, dissolved in CH₃CN and mixed with iodoacetic anhydride which was also pre-dissolved in CH₃CN. Ru(bpy)₂Cl₂ and phen-IA are refluxed in methanol (MeOH), and the product ([Ru(bpy)₂(IAphen)](PF₆)₂, m.w. 1066.52 gmol⁻¹) is precipitated by ammonium hexafluorophosphate (NH₄PF₆) [4]. The resulting precipitates were analyzed with NMR and LC-MS to confirm the product formation.

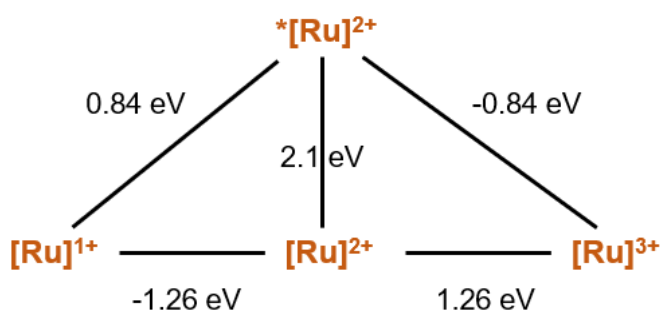


Figure 3.2. Potential of Ru-species (E° vs. NHE in water) generated by the photosensitizer [6].

Ru-labeling mechanism

A covalent thioether bond is formed upon cysteine-specific modification with an iodoacetamido linker of the Ru complex [4]. Therefore, to ensure site-specific labeling, there should be only one single surface exposed cysteine. Since wild-type *Tth*-lac does not have any Cys residue on the surface, a surface residue at a specific location of interest was mutated to Cys for Ru-labeling. Three different Ru-labeled enzymes were produced with S117C, S102C and D390C mutants following the procedure described below.

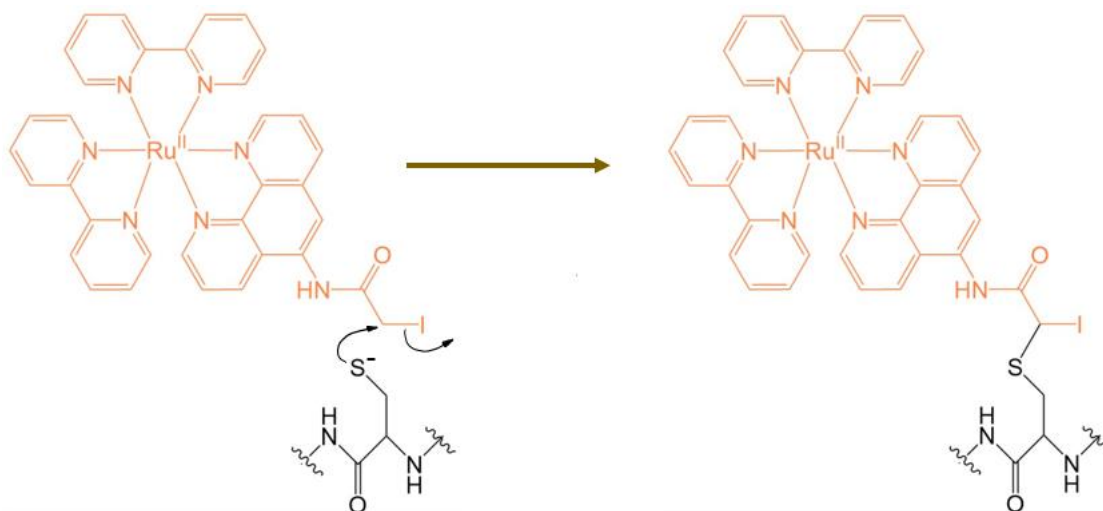


Figure 3.3. Ru-photosensitizer labeling to a surface Cys [6].

1. A protein sample ($\sim 20 \mu\text{M}$) with a surface exposed cysteine is incubated with 10 mM TCEP (tris-(2-carboxyethyl) phosphine) for 5+ hours, and TCEP is washed out with multiple rounds of spin filtration with 20mM Tris buffer, pH 8. (TCEP works as a reducing agent for cleaving disulfide bonds which get easily formed by surface-exposed cysteines. This step is to ensure that free cysteine is available for thioether bond formation.
2. Approximately 4 to 5-fold excess of $[\text{Ru}(\text{bpy})_2(\text{I}A\text{phen})](\text{PF}_6)_2$ is dissolved in 45 mL 20mM Tris buffer, pH 8 (the final concentration to be around 80 - 100 μM). At this low concentration, it can easily be dissolved directly in the buffer with vigorous shaking for a few minutes without adding DMSO. The pK_a of cysteine is close to 8 [7], so the labeling efficiency increases at this high pH owing to deprotonation of the cysteine thiol group.
3. A small volume ($\sim 1 \text{ mL}$) of concentrated enzyme is added to the $[\text{Ru}(\text{bpy})_2(\text{I}A\text{phen})](\text{PF}_6)_2$ solution made in step 2 to make the final concentration of enzyme around 10-20 μM .
4. The reaction proceeds for 5+ hours in the dark at 4 $^\circ\text{C}$. (The reaction solution was usually just left in the cold room overnight with gentle shaking at 100 rpm).

5. Excess Ru complex is washed out by multiple rounds of spin filtration with a fresh Tris buffer, pH 8 (or with other suitable buffers of interest).
6. After labeling, the protein absorption band at 280 nm will be obscured due to the absorption of the Ru-complex in the UV region. Therefore, the labeling efficiency can be confirmed by mass spec analysis in which the peak for the unlabeled enzyme should disappear (mass increase by 650 Da for the labeling enzyme).

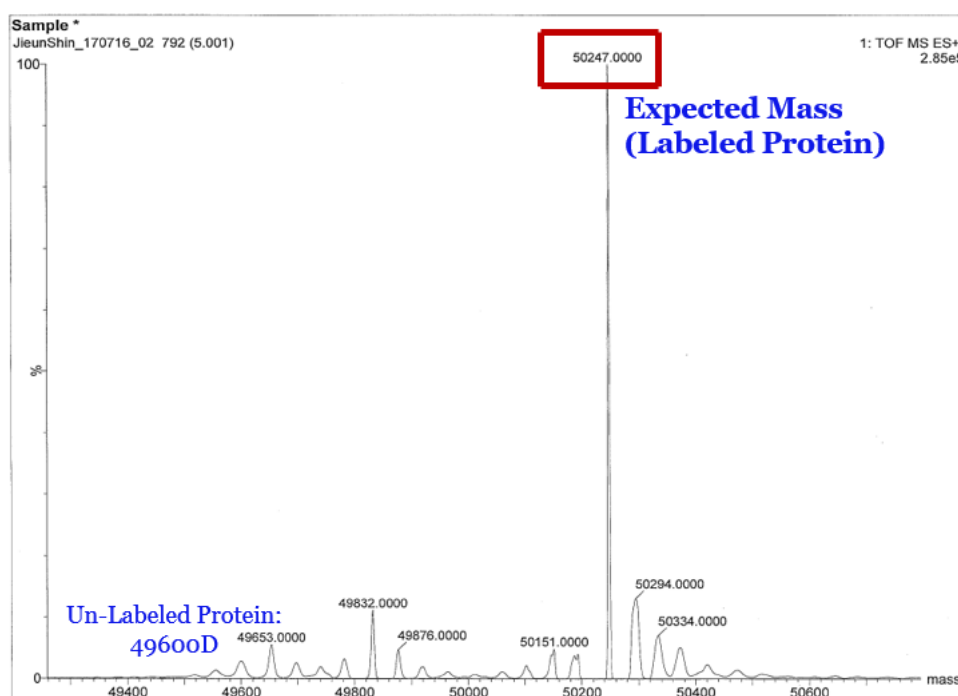


Figure 3.4. MALDI-TOF mass spectrum of the purified protein

Metalation with CuSO₄

After labeling, metalation is achieved with copper sulfate, which makes the protein turn green eventually. Around 20 μ M protein samples were put into a 12 mL or a 30 mL dialysis cassette or a 100 mL dialysis membrane tubing, and were metalated in 20 mM Tris buffer at pH 8 containing up to around 1 mM CuSO₄.

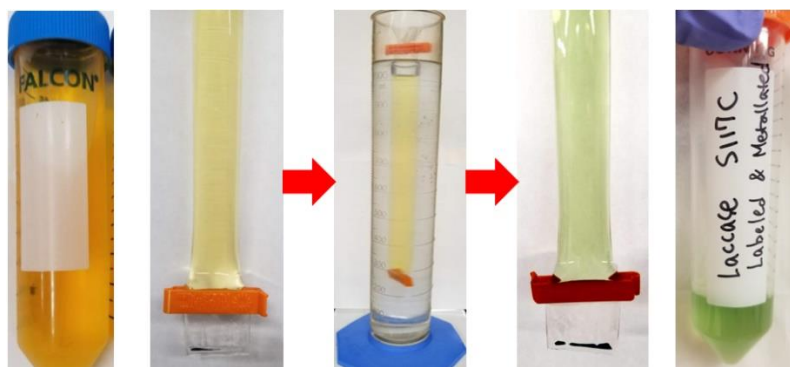


Figure 3.5. *Tth*-lac after labeling (orange) and metalation (green).

Further Purification with FPLC

After the photosensitizer labeling and metalation, a cation exchange column (HiPrep SP HP 16/10 column) was used to further remove impurities by selectively collecting the fractions of interest by monitoring the absorption at 280nm, 450nm (λ_{\max} of the Ru-photosensitizer), and 605 nm (λ_{\max} of Cu_{T1}). Moreover, although close to 100% efficiencies of labeling and metalation are expected to be achieved, this further purification step with ion exchange chromatography ensures that any possible unlabeled or unmetalated proteins are separated to achieve better homogeneity of the sample. For the cation exchange chromatography, the pH of the buffer should be at least one pH unit lower than the pI of the protein. Since the isoelectric point (pI) of *Tth*-lac is around 9.5 [8], Tris buffer at pH 7.8 was used as a equilibration/wash buffer. The same buffer with 500 mM NaCl was used as an elution buffer, and a slow gradient of 0-100% was applied for 2 hours. With the increment of NaCl concentration, the loaded protein sample gets separated into different components within the column depending on their charges. Since the cation exchange resin of the column has sulfonate (SO_3^-) functional groups, more positively charged species bind to the column resin more strongly and get eluted more slowly. *Tth*-lac is cationic at pH 7.8, and is eluted when the concentration of 500 mM NaCl flowing through the column reaches around 40%.

Protein quantification

The concentration of apo-protein is determined by examining the absorption peak at 280 nm. The type 1 copper (Cu_{T1}) exhibits an absorption feature with an extinction coefficient (ϵ) of around $5000 \text{ M}^{-1}\text{cm}^{-1}$ at 605 nm due to the $\text{S}(\text{Cys})_{\pi} \rightarrow \text{Cu}(\text{II})$ charge transfer (CT) [9]. $[\text{Ru}(\text{bpy})_2(\text{Phen})\text{IA}](\text{PF}_6)_2$ complex has a strong absorption feature at around 450 nm with an extinction coefficient of $16600 \text{ M}^{-1}\text{cm}^{-1}$ [4]. However, the photosensitizer ($[\text{Ru}(\text{bpy})_2(\text{Phen})\text{IA}](\text{PF}_6)_2$) absorbs in the 200-300 nm region as well, and its absorption peak at λ_{max} (=450 nm) broadly tails off down to the 600-700nm region. Since the photosensitizer absorption affects both the Cu_{T1} band at 605 nm and the protein band at 280 nm, the absorption band at 450 nm was primarily used to estimate the concentration of the labeled/metalated protein.

Copper quantification

Tth-lac has to be fully metalated with four coppers for its activity. However, the Cu_{T2} site in *Tth*-lac does not bind copper as tightly as other copper sites due to its weaker coordination environment which consists of only two histidines. Therefore, after preparing the samples with different metalation conditions, the copper content of each enzyme sample was quantified using four different analytical techniques: XRF, ICP-MS, EPR, and BCA assay. The concentration of CuSO_4 used for metalating 20 μM protein sample was varied from 200 μM , 300 μM , 500 μM , 1 mM to 10 mM, and non-coordinating buffers such as 25 mM MES buffer at pH 6 and 25 mM sodium phosphate buffer at pH 6 were also tried as well as 20 mM tris buffer at pH 8 and 25 mM sodium acetate buffer at pH 6 and pH 4.5.

- X-ray fluorescence (XRF)

The advantage of using the XRF is that it is a non-destructive method of quantifying copper. Since the protein samples are labeled with the Ru-photosensitizer on a surface labeling site, the ratio of Ru: Cu for one enzyme unit should be 4:1. Therefore, coppers can be quantified by examining the ratio between Ru and Cu measured by the XRF as an internal standard. However, unfortunately, parameters such as the X-ray penetration depth within a sample container, location of the sample holder on the detector, concentrations of protein samples all seem to affect the XRF measurements, causing differences in values every time the measurements are made even for the same sample. Besides, the detection limit for ruthenium is much higher than that for copper, so ruthenium could be detected only when the enzyme concentration was higher than 150 μM , which was problematic, since the amount of protein required to fill the sample container (10 mL) was too much to prepare. However, for copper quantification, standards of known copper concentrations ranging from around 10 μM to 100 μM could be used to make a calibration curve from which the copper concentration in the actual protein sample of a known concentration can be estimated.

- ICP-MS

The ICP-MS is a destructive method of quantifying the copper content within the enzyme, since the enzyme has to be fully digested to spit out all the metals into the solution that is being analyzed. The enzyme samples were digested in 10 % nitric acid with heating up to 70 $^{\circ}\text{C}$, and the resulting solutions were diluted to bring down the nitric acid concentration to around 2 % which is the concentration tolerable by the instrument.

- EPR

The Cu_{T1} has a small parallel hyperfine coupling constant ($A_{\parallel} = 40-90 \times 10^{-4} \text{ cm}^{-1}$) due to the S(Cys) to Cu(II) charge transfer. The Cu_{T2}, on the other hand, exhibits a parallel hyperfine coupling constant ($A_{\parallel} = 140-200 \times 10^{-4} \text{ cm}^{-1}$). Binuclear Cu_{T3} coppers are EPR silent due to antiferromagnetic coupling resulting from a hydroxo bridge linking the two coppers [9]. To ensure that the Cu_{T2} (with the lowest binding constant) is present in the metalated sample, EPR measurements were made on the samples prepared at 150 μM protein concentration with 20% glycerol as a cryo-protectant.

Laser samples

For the regular single wavelength or multi-wavelength transient absorption spectroscopy, the laser sample contained 60 μM Ru-labeled protein and 10 mM Ru(NH₃)₆Cl₃ as a reversible oxidative quencher in a sealed quartz cuvette. The sample was deoxygenated by multiple pump-argon backfill cycles on a Schlenk line. 60 μM protein concentration was the highest possible without obscuring the signal due to its photosensitizer absorbance of around 1 at 450 nm (λ_{max}). The concentration of Ru(NH₃)₆Cl₃ was carefully adjusted, since Ru^{III}(NH₃)₆³⁺ disproportionates to Ru^{IV}(NH₃)₆⁴⁺ shortening the lifetime of Ru^{III}(bpy)₂(phen)³⁺ due to its reduction back to Ru^{II}(bpy)₂(phen)²⁺ by Ru^{IV}(NH₃)₆⁴⁺ species. This back-reaction significantly reduces the size of the signal compared to the *Ru^{II}(bpy)₂(phen)²⁺ excited state bleach, since the time for the holes from Ru^{III}(bpy)₂(phen)³⁺ to move on and oxidize type 1 copper is shortened (Figure 3.6).

Ru(NH₃)₆Cl₃ Disproportionation:

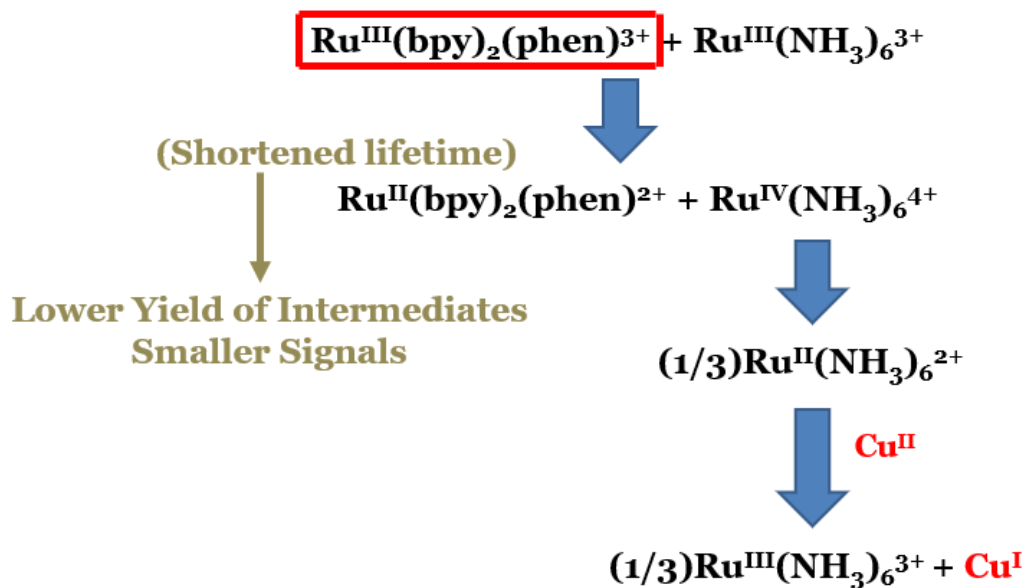


Figure 3.6. Limitations with [Ru(NH₃)₆]Cl₃ as a reversible quencher.

To enhance the signal as much as possible despite the Ru(NH₃)₆Cl₃ disproportionation issue, the quantum yield calculation was done so that the concentrations of the protein and the quencher can be optimized. The lower the concentration of quencher and the higher the concentration of protein is, the better it is for the signal size. From the quantum yield calculation (Figure 3.7), it was estimated that the $\Phi_{\text{Ru}^{3+}}$ for 10 mM Ru(NH₃)₆Cl₆ is still approximately 70% which is sufficient to promote enough quenching for the flash-quench laser experiments. The condition with 60 μM protein and 10 mM quencher was determined to be optimal for achieving the big enough signal size and high enough degree of quenching.

Quantum yield calculation

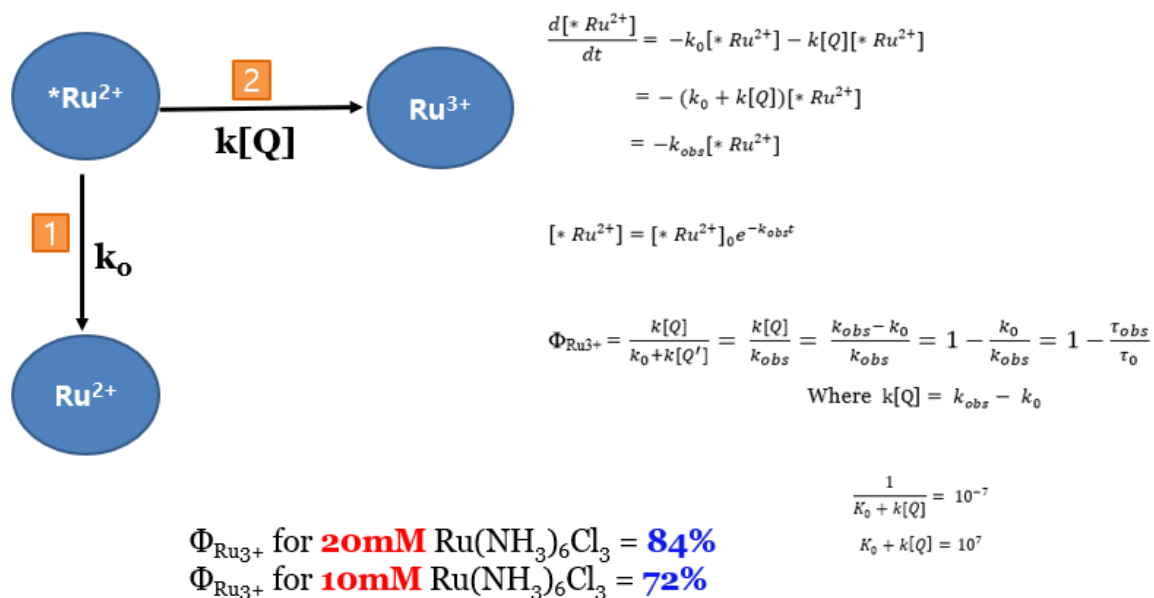


Figure 3.7. Quantum yield calculation for the $Ru(NH_4)_6Cl_3$ quencher

When the oxidative quencher is used to inject holes into the system, the enzyme sample needs to be reduced prior to the laser experiment to monitor the hole transfer pathway. The reduction of coppers in *Tth*-lac was achieved by anaerobically injecting four-reducing equivalents of deoxygenated $Ru^{2+}(NH_3)_6Cl_2$ ($E^\circ'(Ru^{3+}/Ru^{2+}) = +0.1$ V [10]). As coppers get reduced, the absorbance of Cu_{T1} at 605 nm decreases to a steady value close to zero.

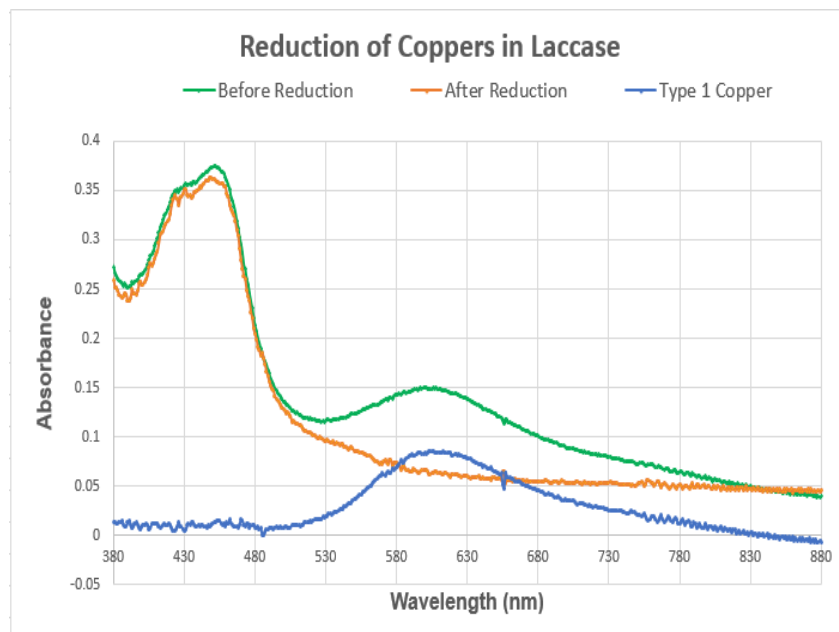


Figure 3.8. Spectra showing the reduction of coppers in Ru-labeled *Tth*-lac: oxidized (green), reduced (orange) and [oxidized – reduced] (blue)

Other reduction methods

Commonly used oxidative quenchers such as $\text{Ru}^{3+}(\text{NH}_3)_6\text{Cl}_3$ ($E^{\circ'}(\text{Ru}^{3+}/\text{Ru}^{2+}) = +0.1 \text{ V}$ [10]), methyl viologen (MV^{2+} , $E^{\circ'}(\text{MV}^{2+}/\text{MV}^{+\bullet}) = -446 \text{ mV}$ [11]) and benzyl viologen (BV^{2+} , $E^{\circ'}(\text{BV}^{2+}/\text{BV}^{+\bullet}) = -359 \text{ mV}$ [11]) can be reduced with zinc powder or zinc amalgam to $\text{Ru}^{2+}(\text{NH}_3)_6\text{Cl}_2$, methyl viologen radicals ($\text{MV}^{+\bullet}$) and benzyl viologen radicals ($\text{BV}^{+\bullet}$), respectively. A single amino acid cysteine ($E_0'(\text{cysteine}/\text{cystine}) = -220 \text{ mV}$ [12]) can also be used to reduce coppers in *Tth*-lac. Ferrocyanide ($\text{Fe}(\text{CN})_6^{3-}(\text{aq})/\text{Fe}(\text{CN})_6^{4-}(\text{aq}) = +361 \text{ mV}$ [13]) has a suitable potential as a reductant for *Tth*-lac, but it was incompatible with the protein, causing protein precipitation. Titanium (III) citrate also made the enzyme crash out even with high ionic strength provided by a high concentration of sodium citrate in the buffer.

3.3. Flash-Quench Laser Experiment

After labeling and metalation, the enzyme is ready for the laser experiment. With a laser pulse, the photosensitizer excited state ($*\text{Ru}^{2+}$) can be generated. Then, the photosensitizer excited state is quenched by either a reductive or an oxidative quencher, which can produce either the more reducing $\text{Ru}(\text{bpy})_2(\text{phen})^{1+}$ or the more oxidizing $\text{Ru}(\text{bpy})_2(\text{phen})^{3+}$, respectively. With these newly generated redox species, the electron/hole transfer processes can be investigated. The oxidative quenching pathway was of a primary focus to inject holes into the system and monitor hole hopping pathways which facilitate long range copper oxidation.

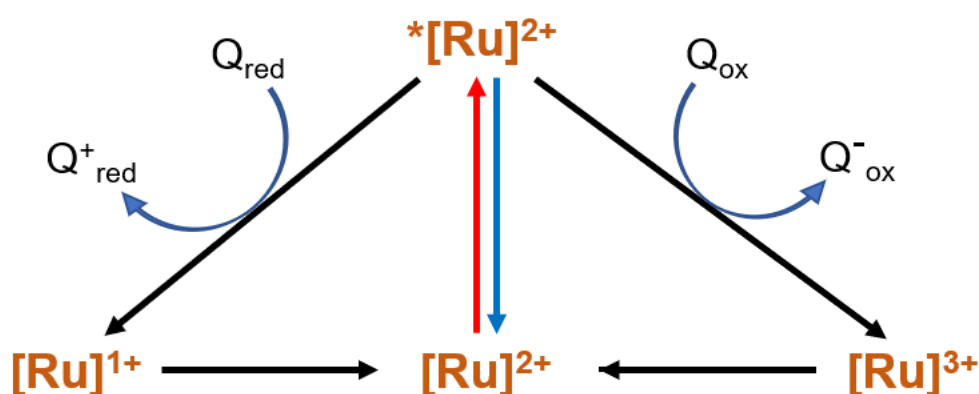


Figure 3.9. Flash-quench cycle of the Ru-photosensitizer [6].

Reductive quenching pathway

In the reductive quenching pathway, the photosensitizer excited state ($*\text{Ru}^{2+}$) is driven to Ru^{1+} . As these Ru^{1+} species get oxidized back to Ru^{2+} species, electrons get injected into the system. Although para-methoxydimethylaniline ($E^\circ(\text{MeODMA}^{+/0}) = + 0.66 \text{ V}$ [14]) can be used as a reductive quencher, the kinetics measurements made with MeODMA were not reliable as it can also act as a substrate for *Th*-lac.

Oxidative quenching pathway

Upon oxidative quenching, the photosensitizer redox state (Ru^{2+}) is driven to Ru^{3+} . As these Ru^{3+} species get reduced back to Ru^{2+} species, holes get injected into the system. The oxidative quenching pathway was the primary focus of the studies which will be explained in the later parts of this chapter.

Quencher studies

Water soluble and photostable small molecule quenchers with adequate reduction potentials are suitable for laser flash-quench studies. Redox reversibility is also desired to achieve a high signal-to-noise ratio by averaging kinetics profiles from hundreds and thousands of laser shots. Moreover, the absorbance of quenchers must not interfere significantly in the wavelengths being monitored. $\text{Ru}(\text{NH}_3)_6\text{Cl}_3$ satisfies these requirements as a reversible oxidative quencher, but the kinetics obtained with it are contaminated by many back reactions in the *Tth*-lac system. Although irreversible quenchers such as $[\text{Co}(\text{NH}_3)_5\text{Cl}]\text{Cl}_2$ could be a solution to prevent back reactions, the most prominent downside with using irreversible quenchers is that the number of laser shots that can be applied to the sample is limited. To resolve this problem, it was necessary to look for some alternatives for a better reversible quencher.

Table 3.1. Biomolecular quenching rate constant [15].

Quencher	Bimolecular Quenching Rate Constant ($k_q, \text{M}^{-1}\text{s}^{-1}$)
Methyl Viologen	1.03×10^9
Benzyl Viologen	1.37×10^9
$\text{Ru}(\text{NH}_3)_6\text{Cl}_3$	1.4×10^9
1,4-Benzoquinone	1.1×10^{10}
Fe(NTA)	4.3×10^9

To find a small molecule that can qualify for a suitable quencher, the quenching rate constants for different compounds were examined (Table 3.1). These values can be obtained either from the literature or from quick luminescence lifetime measurements of $\text{Ru}(\text{bpy})_3^{2+}$ in the presence and absence of the compound of interest. When doing so, another aspect that needs to be considered is the cage escape yield. Cage escape efficiency implies how efficiently the primary quenching products escape back electron transfer within the solvent cage and diffuse freely into the bulk solution [15]. This can be estimated by examining the transient absorption kinetics at 440 nm which shows the lifetime of Ru^{2+*} and Ru^{3+} species generated by laser flash quench. In the case of 1, 4-benzoquinone shown in Figure 3.10, the luminescence decays away in about 40 ns whereas the signal for Ru^{3+} species (from the transient absorption kinetics at 440 nm) disappear in about 200 ns (Figure 3.10). Therefore, we have a much wider additional window which allows the hole to escape back electron transfer within the solvent cage and to move on to oxidize other species.

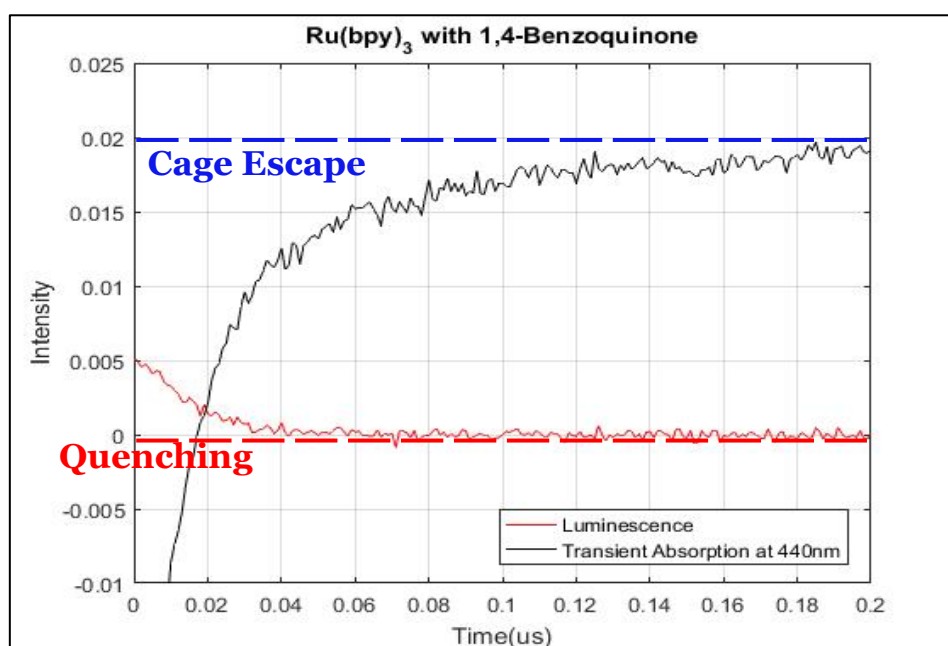


Figure 3.10. Cage escape yield of 1, 4-benzoquinone.

As mentioned earlier, another pre-requisite is the photostability of the quenchers. Although the quenching efficiency of 1,4-benzoquinone is suitable for our study, it is not photochemically stable (Figure 3.11). The degree of photo-decay could be alleviated with the addition of NaCl and the usage of a longer excitation wavelength (Figure 3.12). However, these aspects certainly cause limitations to the experimental conditions.

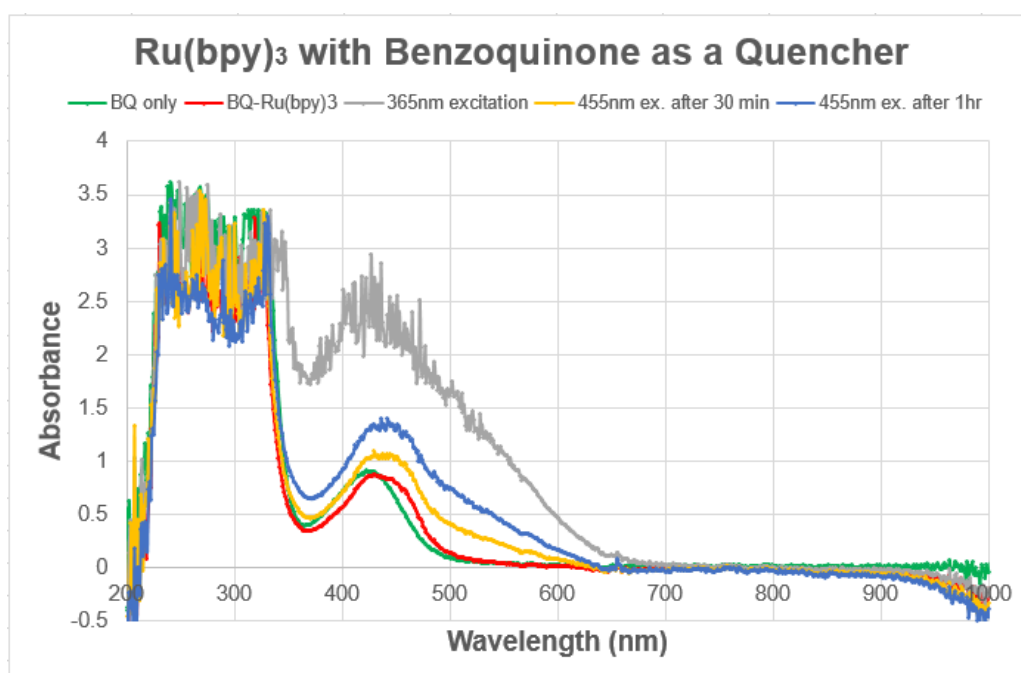


Figure 3.11. Photochemical instability of 1,4-benzoquinone.

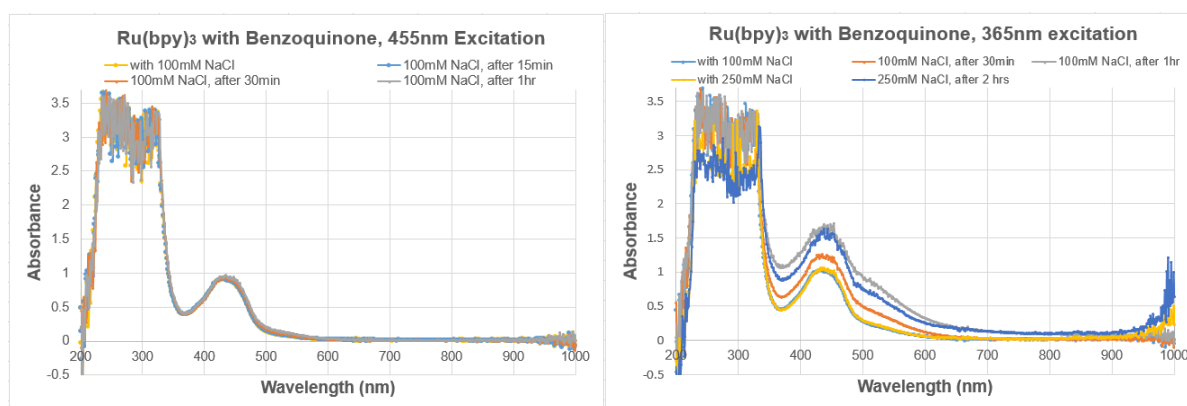


Figure 3.12. Photo-stability of benzoquinone with 455 nm (left) and 365 nm (right) excitation.

Moreover, suitable quenchers should be highly water soluble in regular buffer conditions, and should not absorb a lot on its own in the wavelength region of interest, since the quencher absorbance may obscure all the signals. Methyl viologen and benzyl viologen were also tried as quenchers, but the radicals of both have a broad absorption band at around 500 nm [16], which could interfere with the absorption features of Ru^{3+} , Trp radicals and Cu^{2+} arising at different time scales during the laser flash-quench experiments.

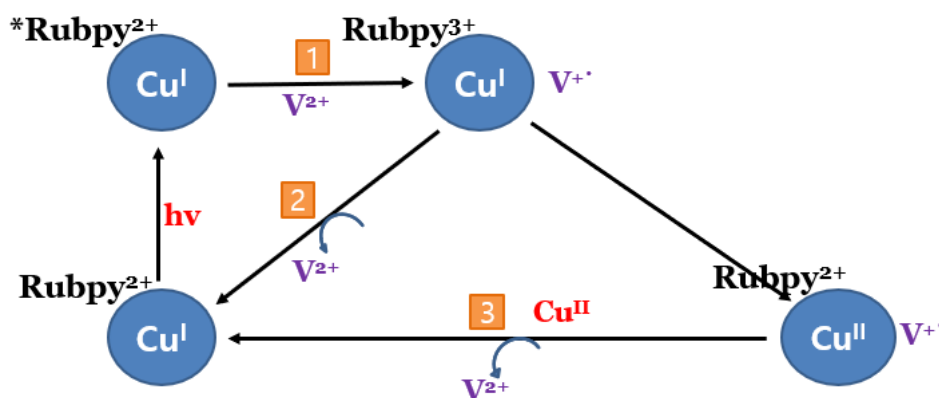


Figure 3.13. Quencher mechanism of Methyl Viologen (MV) and Benzyl Viologen (BV).

The Laser Setup

- Excitation light provided by a tunable optical parametric oscillator (OPO)
- Pumped by the 3rd harmonic from a Q-switched ND:YAG laser (8 ns pulse width)
- Adjustment of operating frequency (10 Hz or 1 Hz)
- Probe lamp: 75 W Arc lamp, continuous or pulsed
- Excite and probe collinearly
- Rejection of undesired fluorescence using various optical filters
- Probe light selection by a double monochromator, detection by a photomultiplier tube (PMT), and the signal amplification with an amplifier

3.4. Transient Absorption Kinetics

Ru-C117 Enzyme

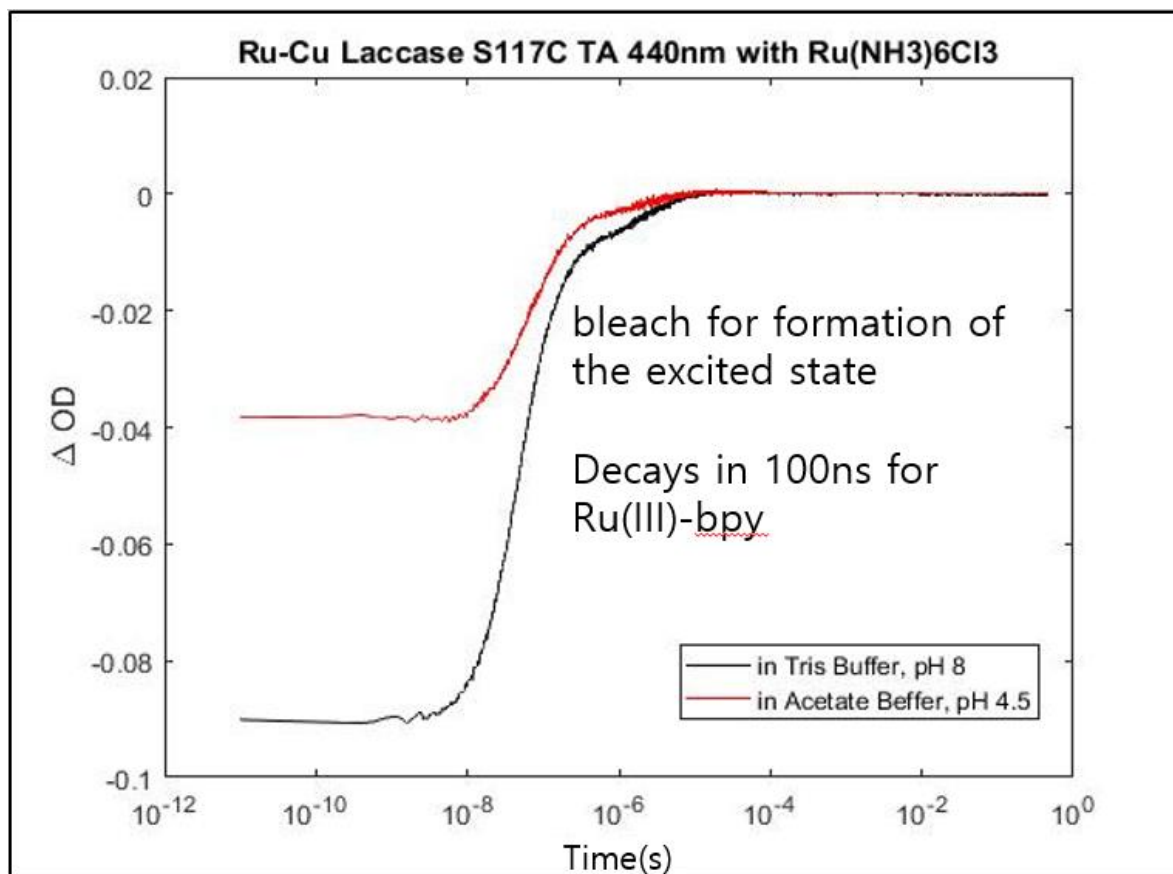


Figure 3.14. Transient absorption kinetics of Ru-C117 at 440 nm.

The laser-triggered flash-quench electron/hole transfer process can be probed by transient absorption spectroscopy. 440 nm is the Ru²⁺-bpy ground state absorption region, and a bleach is observed for the formation of Ru-bpy excited state. The bleach decays in about 100 ns to give another species which are presumably Ru³⁺-bpy. The Ru³⁺-bpy signal seems to disappear by 10 μs (Figure 3.14).

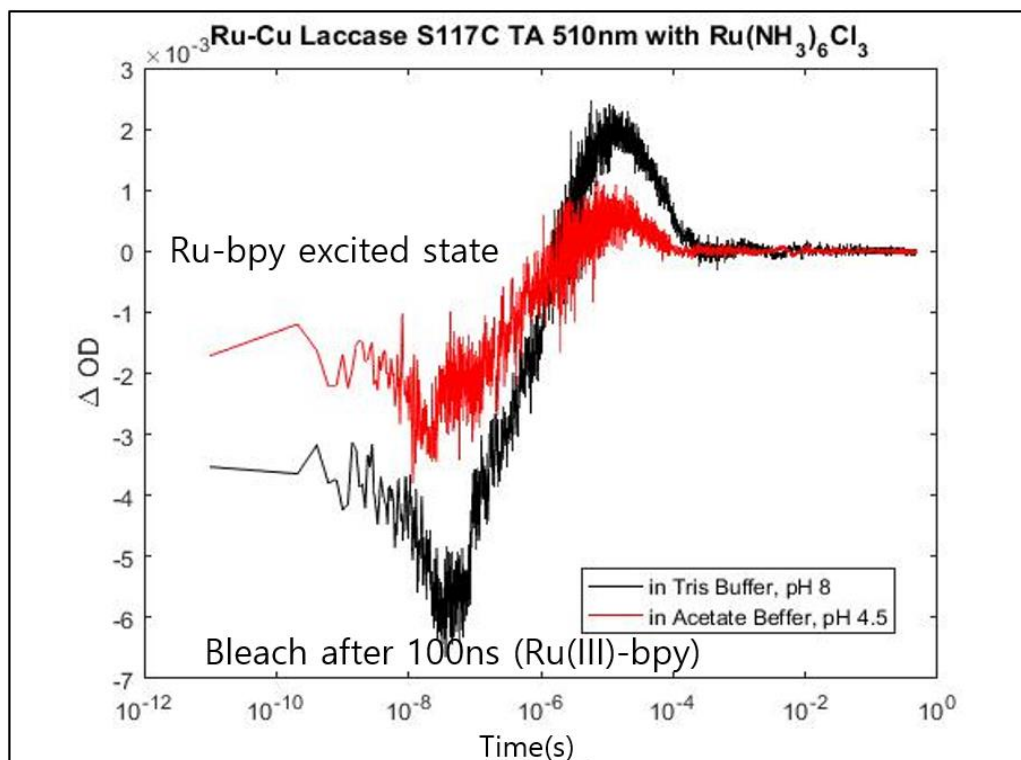


Figure 3.15. Transient absorption kinetics of Ru-C117 at 510 nm.

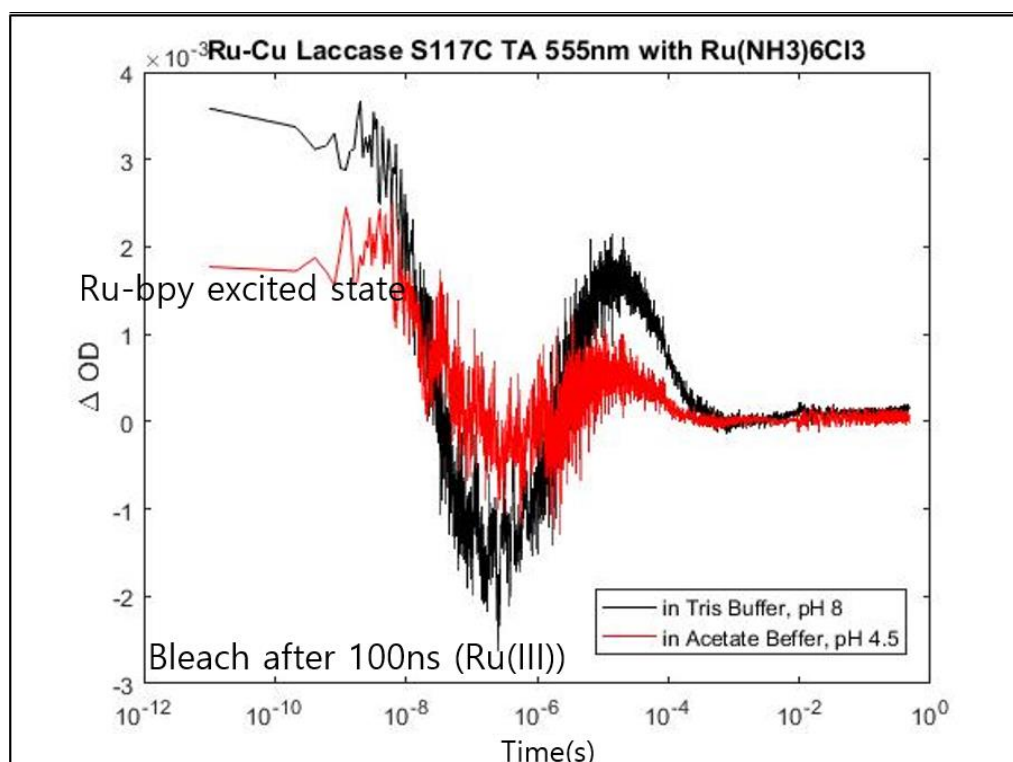


Figure 3.16. Transient absorption kinetics of Ru-C117 at 555 nm.

Since there is a tryptophan residue (W118) right next to the S117C labeling site, in search for Trp radical intermediate species, 510nm and 555nm regions were also examined, since Trp neutral radicals and Trp radical cations absorb in these visible regions, respectively. In 555 nm data (Figure 3.16), a small positive transient absorbance at early time may be due to the Ru-bpy excited state slightly absorbing at 555 nm. This positive signal decays to a bleach after around 100 ns, which may correspond to the bleach of a small amount of Ru²⁺-bpy at 440nm due to the formation of Ru³⁺-bpy. A positive signal appears at around 10 μ s due to another species and decays back to zero in about 1 ms. The transient absorption feature obtained at 510 nm looks very similar to that observed at 555nm. However, a bleach is observed at early time possibly implying that the Ru-bpy excited state gives a bleach at this wavelength (Figure 3.15).

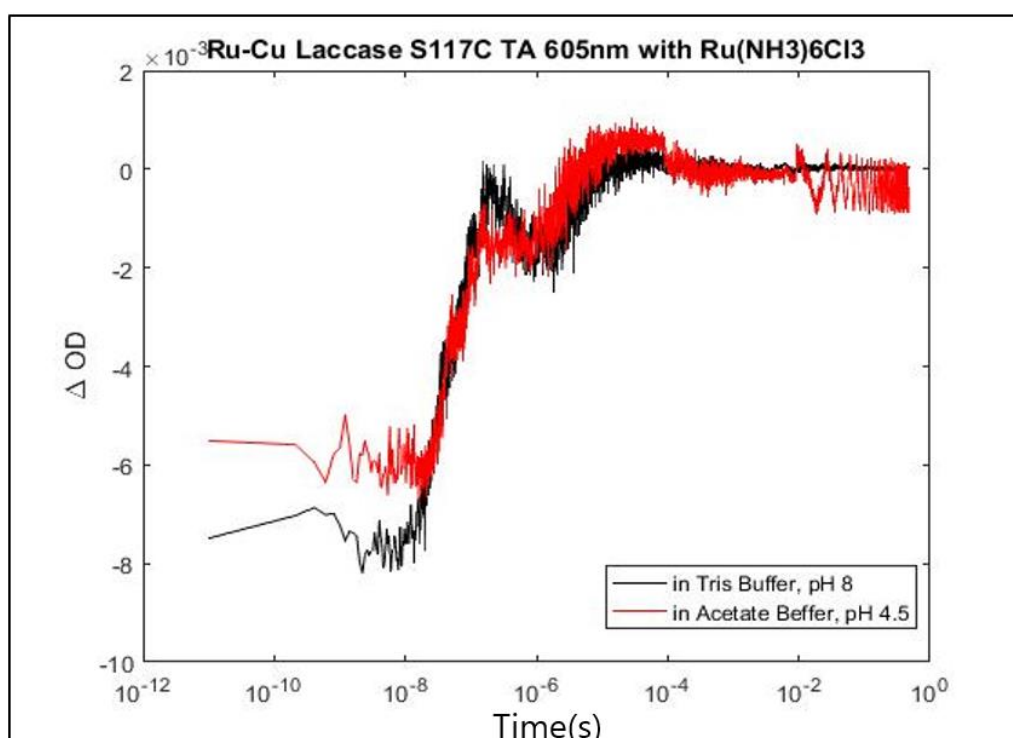


Figure 3.17. Transient absorption kinetics of Ru-C117 at 605 nm.

When the 605 nm region is probed, a bleach is observed at early times, since the signals are compromised by Ru-bpy luminescence (Figure 3.17).

Since it is very likely that the new species being generated around 10 μ s at 510 nm and 555 nm are Trp radicals (Figure 3.15 and Figure 3.16), an additional mutation to the S117C mutant was made to change the nearest redox-active tryptophan(W) at site 118 to phenylalanine(F) and to compare transient absorption signals.

Ru-C117-W118F Enzyme

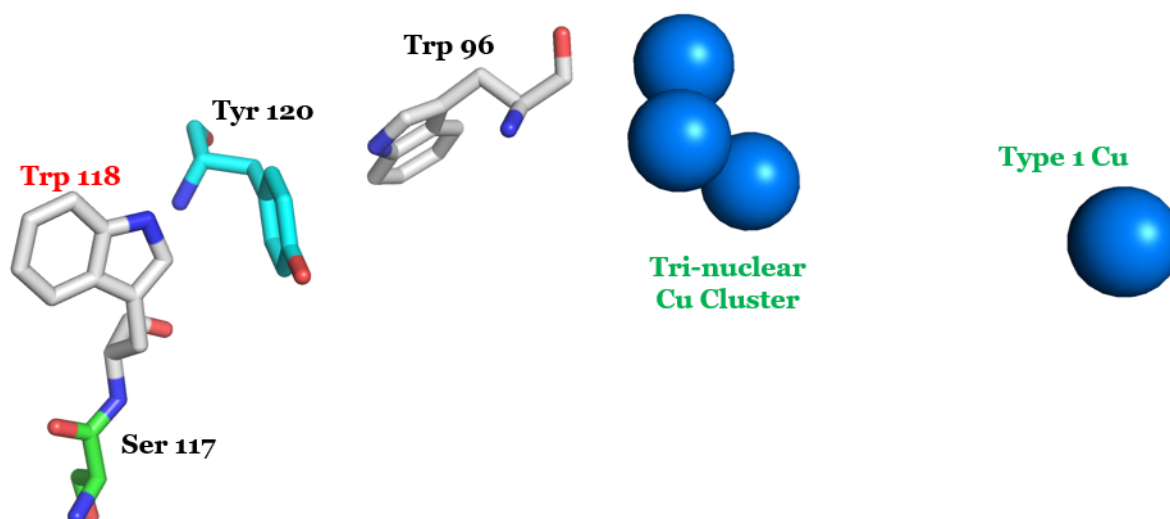


Figure 3.18. The first Trp mutation (W118F) right next to the labeling site (S117C).

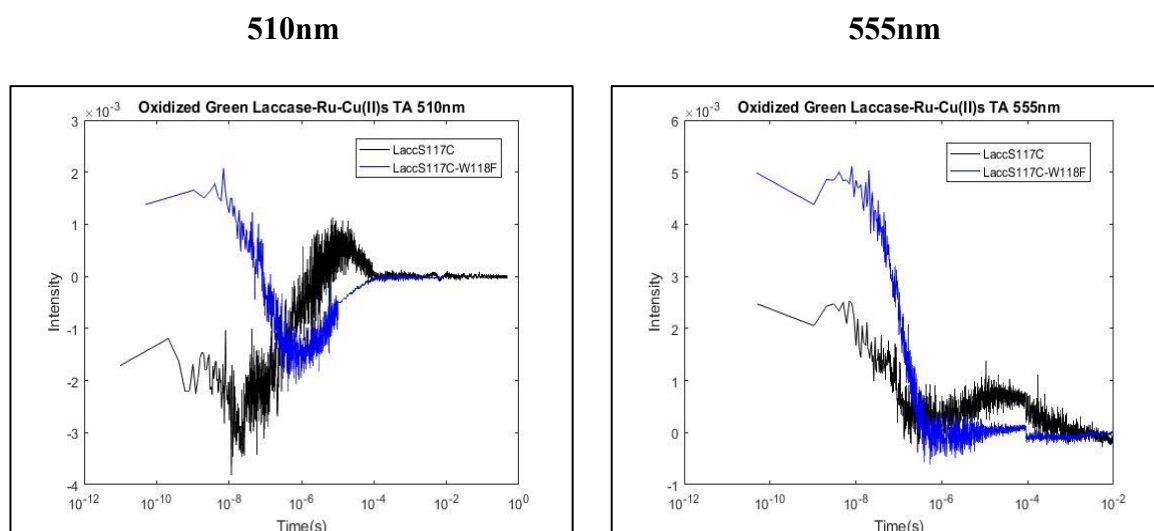


Figure 3.19. Transient absorption data collected at 510 nm and 555 nm (Trp radical absorption regions) for both S117C mutant and S117C-W118F mutant.

The positive signal at 1 μ s does not appear with the W118F mutant, and the signal flattens out to 0 (Figure 3.19). Therefore, it is clear that the positive signals at these wavelengths are due to the formation of Trp radicals, when the holes are injected into the system from the S117C surface labeling site.

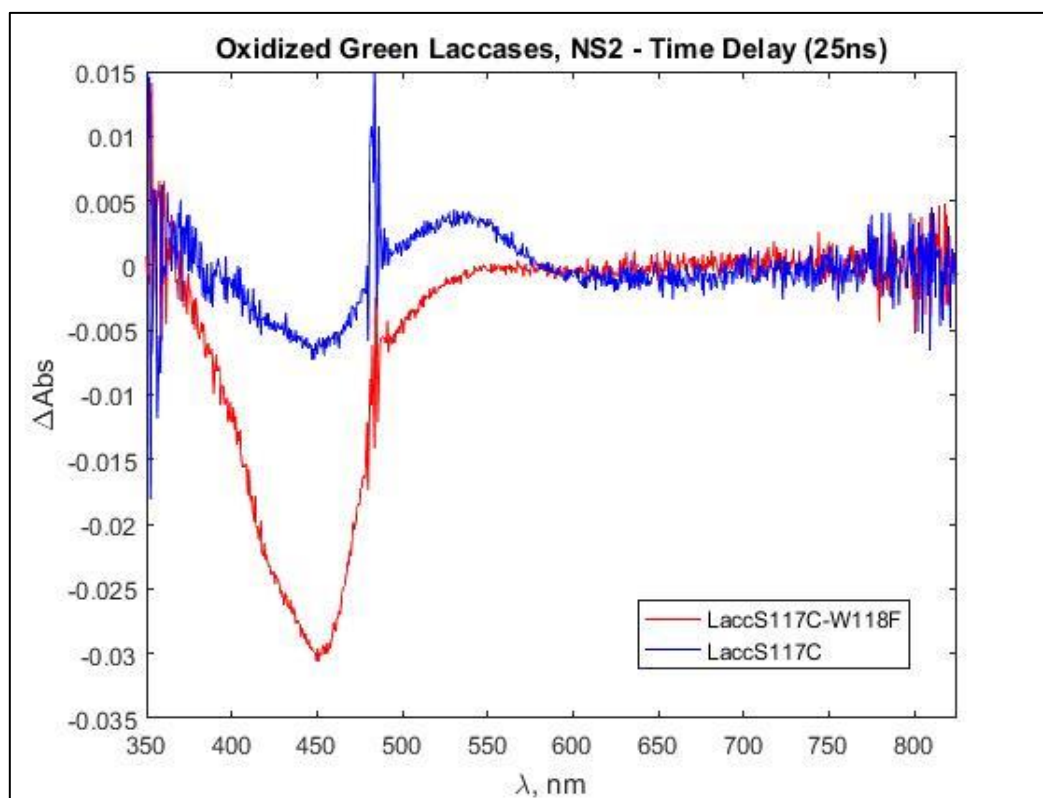


Figure 3.20. Multi-wavelength transient absorption data for Ru-C117 mutant and Ru-C117-W118F mutant (overlaid).

In order to fully monitor the intermediates being generated, the whole spectrum was examined using a diode array system. Here again, it is confirmed that the positive signal at around 500-550 nm region disappears for the W118F mutant, which indicates that we are indeed generating Trp radicals. For the Ru-C117-W118F mutant, a larger bleach is observed around the Ru-bpy ground state absorption region (450 nm), which might be due to the fact that the Ru^{3+} lives longer for the Trp mutant, since it does not have a tryptophan next to the labeling site to oxidize.

Ru-C117-W96F Enzyme

However, since positive signals are observed at both 510nm and 555nm absorption regions, there is a possibility that the next tryptophan along the chain might also be getting oxidized. In order to verify what the actual case is, the second Trp mutant (Ru-C117-W96F) was prepared.

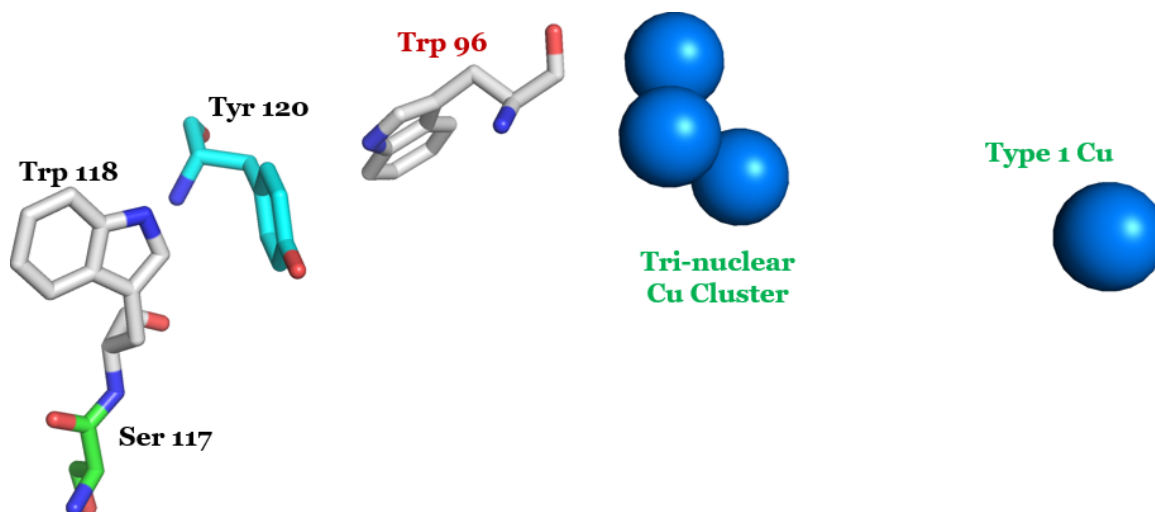


Figure 3.21. The second Trp mutation (W96F) near the tri-nuclear copper cluster.

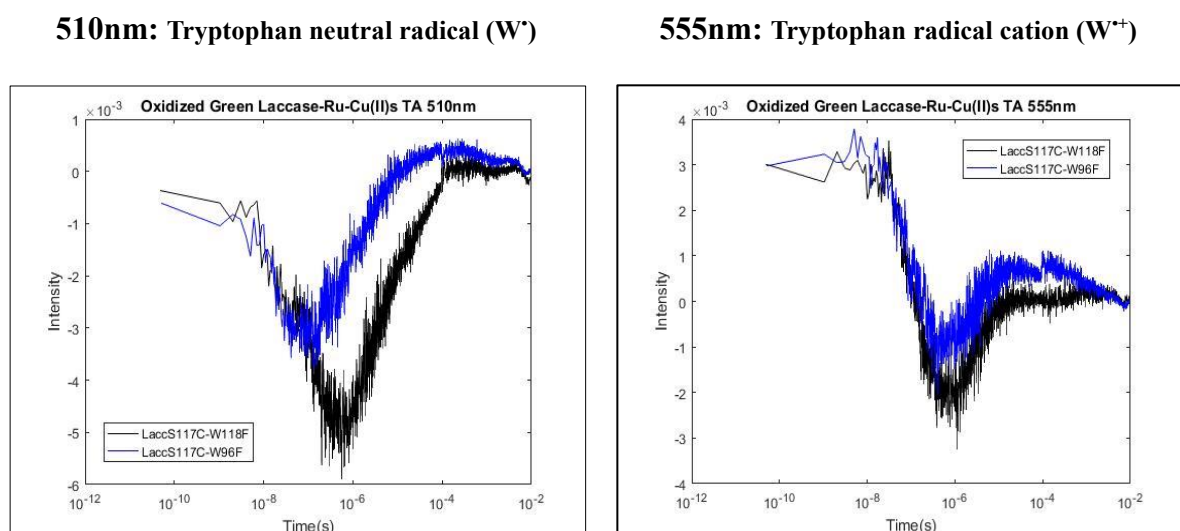


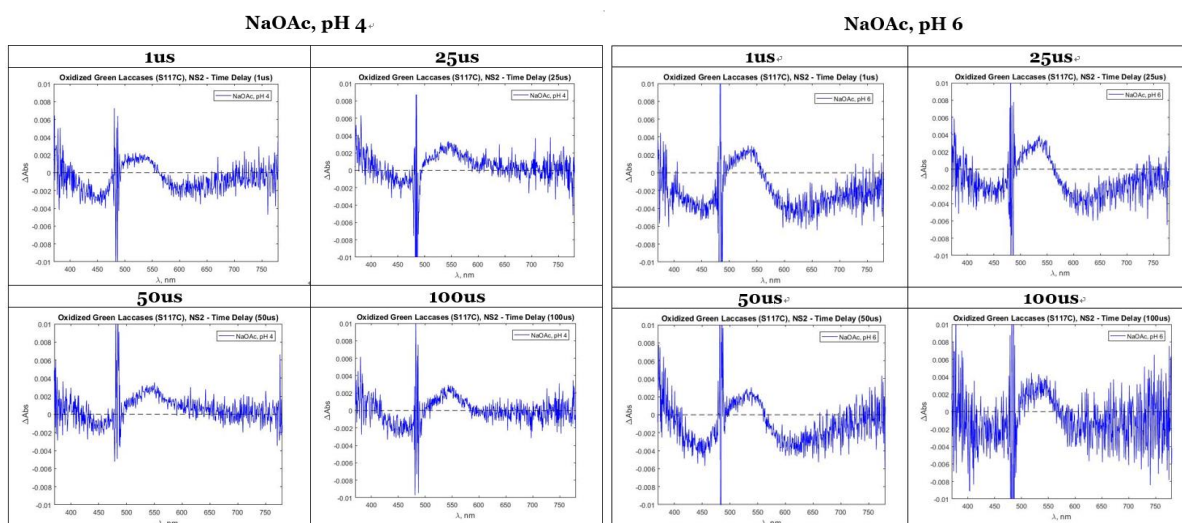
Figure 3.22. The transient absorption signals at 510 nm and 555 nm for the first Trp (W118F) and the second Trp mutant (W96F) overlaid.

Surprisingly, the positive transient absorption features at 500-550 nm disappear for the second Trp mutant (Ru-C117-W96F) indicating that the oxidation of only the first tryptophan (W118) is responsible for producing those features even though both Trp neutral radicals and Trp radical cations seem to be getting generated.

Investigating the pH effects

Although the absorption features for both tryptophan neutral radicals and tryptophan radical cations are observed, only W118 seems to be getting oxidized giving these intermediate signals. The presence of positive absorption features at both 500 nm and 550 nm could be attributed to the pK_a of tryptophan cation, since Trp radicals have a pK_a of around 4.5 in water [17]. Except the first few sets of experiments, most of the laser experiments were carried out in sodium acetate buffer at pH 4.5 to mimic the kinetics of electron transfer during enzyme catalysis, since the activity of the enzyme toward ABTS is the highest at pH 4.5 [1]. The protonation state of Trp radical (TrpH⁺ to Trp[•]) is dependent on the pH, and when the sample pH is sufficiently close to the pK_a of Trp radical, a mixture of both protonated (W^{•+}, λ_{max} = 510 nm and ε = 2300 M⁻¹cm⁻¹) and deprotonated (W[•], λ_{max} = 560 nm and ε = 3000 M⁻¹cm⁻¹) species can be observed [17].

To test whether this concurrent appearance of signals corresponding to both Trp neutral radicals and Trp radical cations is due to the pH effect, samples were prepared in buffers of varying pH and were tested on the diode array system to see if there are any changes in absorption spectra.



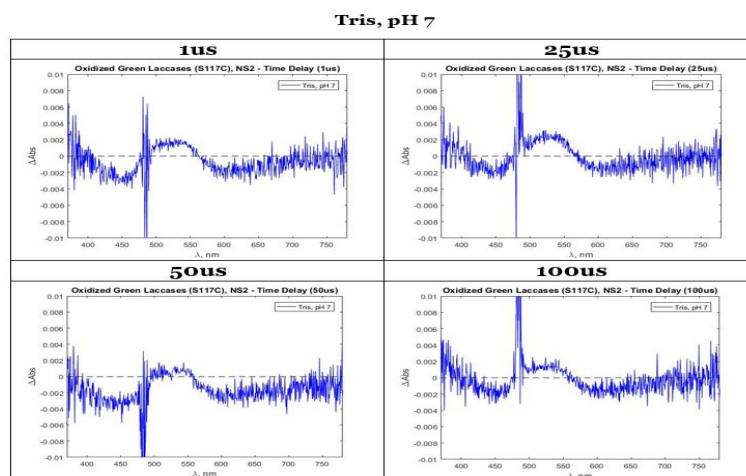


Figure 3.23. Multi-wavelength transient absorption spectroscopy using the diode array system for Ru-C117 mutant in NaOAc (pH 4), NaOAc (pH 6) and Tris (pH 7).

It was expected that more of neutral radical species would be observed as the pH increases. However, deviating from the conventional norm, the peak intensity in the radical cation absorption region increased with the increase in pH. However, as can be seen most prominently in acetate buffer at pH 4, the bleach at around 600 nm region disappears with increasing time delay after each round of 1000 laser shots. It is obvious that the oxidized enzyme is being reduced during multiple flash-quench cycles (with the excitation wavelength of 355 nm) and the sample speciation is constantly changing for the entire duration of the laser experiment. Although the photo-reduction of the enzyme seems to be slower at a higher pH (since the bleach at 600 nm lasts longer at pH 6 and 7), this photochemical reduction phenomenon seems to occur consistently in all laser samples (initially oxidized), even though the source these electrons is still not entirely clear.

Photo-reduction of laser samples

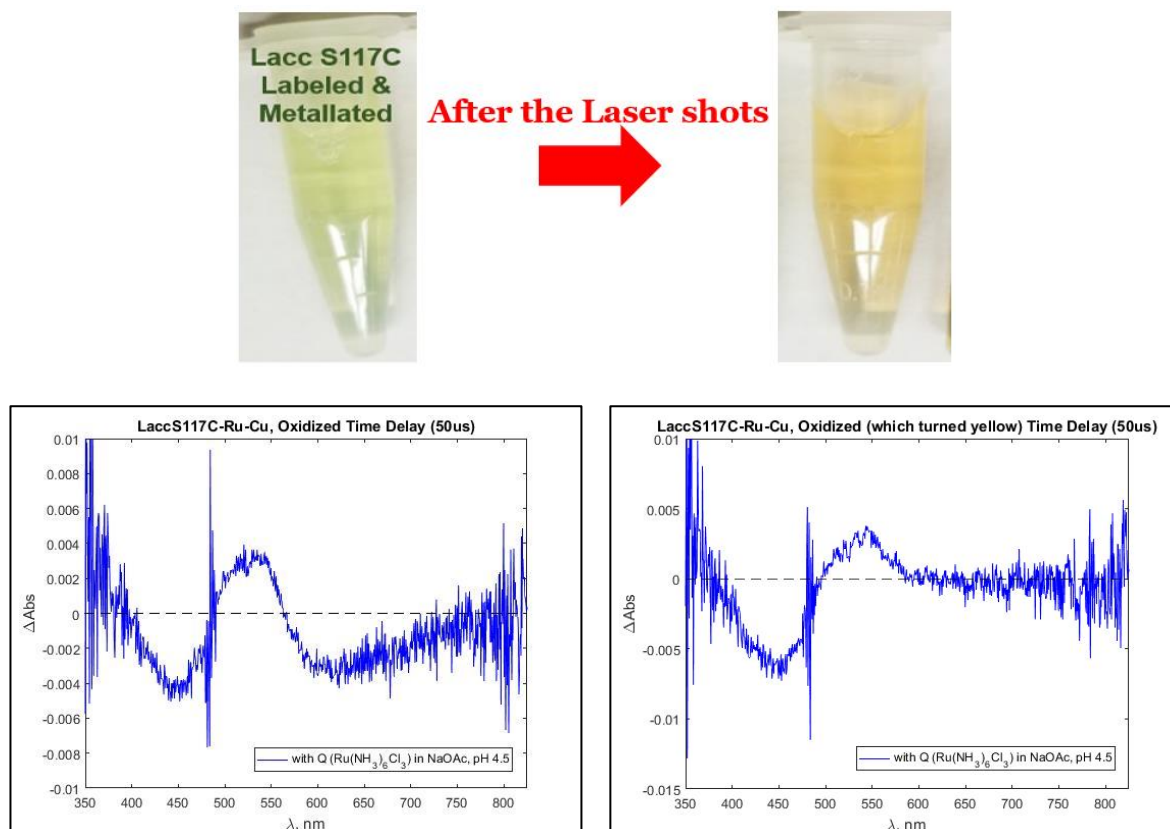


Figure 3.24. Oxidized green laccase getting photochemically reduced (turning yellow).

It is indeed interesting that even the oxidized green laccase samples turn yellow after a few photo-excitation cycles just like the chemically reduced enzyme samples. The bleach due to the Cu_{T1} at around 600nm region goes away after a few thousands of laser shots (Figure 3.24). When a small volume of aerated buffer was added to these photo-reduced samples, the enzyme color turned back to green confirming that it was the photo-reduction which turned the enzyme color to yellow. It is likely that the reduction of Cu_{T1} occurs due to the Ru^{2+} species being generated over the courses of flash quench cycles (Figure 3.6).

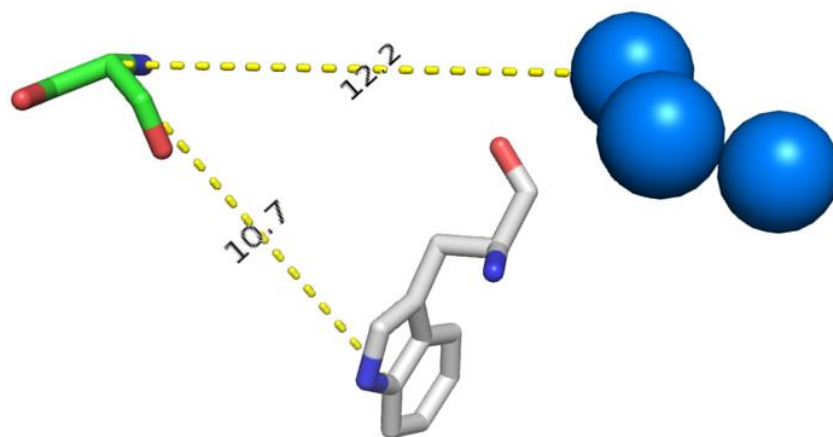
Ru-C102 Enzyme

Figure 3.25. S102C labeling site and its distance from a nearby Trp residues (W96) and the TNC.

S102 residue is only around 12 Å away from the $\text{Cu}_{\text{T}2}$ and only around 11 Å away from a neighboring Trp (W96), so a single step electron hopping is possible to both Trp 96 and the trinuclear Cu cluster. To be able to better observe the kinetics of potential hole transfer to these sites without the contamination from back reactions, an irreversible cobalt quencher ($\text{Co}(\text{NH}_3)_5\text{Cl})\text{Cl}_2$ was used to examine the kinetics with a HeNe laser (Figure 3.29).

Activity check after the laser experiments

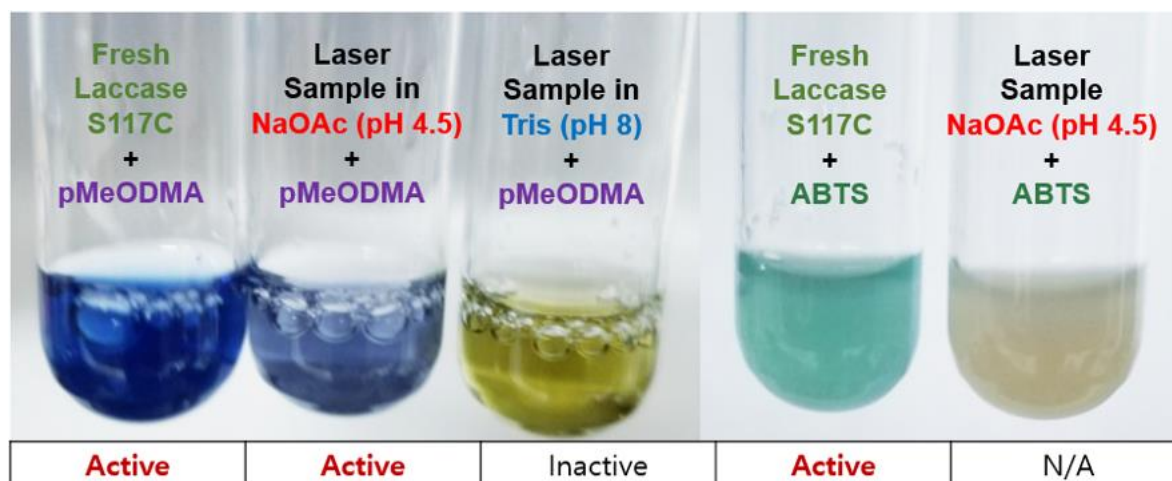


Figure 3.26. Activity check with pMeODMA and ABTS after the laser experiment.

After the laser experiments, the activity of the laser samples was checked especially for the photo-reduced samples to see if other factors such as de-metalation or enzyme degradation also contributed to the color change to yellow. The conventional activity check method using ABTS could not be used, because $[\text{Ru}(\text{NH}_3)_6]\text{Cl}_3$ forms precipitates with ABTS due to their charge interactions. When another substrate (MeODMA) was tested, the laser samples still actively oxidized MeODMA even after thousands of laser shots when the samples were prepared in NaOAc at pH 4.5. However, the enzyme became inactive when the samples prepared in Tris, pH 8 were left overnight, since $[\text{Ru}(\text{NH}_3)_6]\text{Cl}_3$ decomposes faster at a higher pH leading to protein degradation. The activity check with MeODMA confirmed that the photo-reduced laser samples remained active after the laser experiment.

Sodium azide inhibition effects

To better understand why the photo-reduction of the enzyme occurs, an azide inhibition experiment was tried. Sodium azide (NaN_3), as a metal chelator, inhibits the activity of laccases by binding to the tri-nuclear copper cluster [18]. If the photo-reduction of enzyme involves the electron transfer through the tri-nuclear copper cluster, the color change to yellow should not occur when NaN_3 is also added to the samples containing Ru-labeled *Tth*-lac and $[\text{Ru}(\text{NH}_3)_6]\text{Cl}_3$.

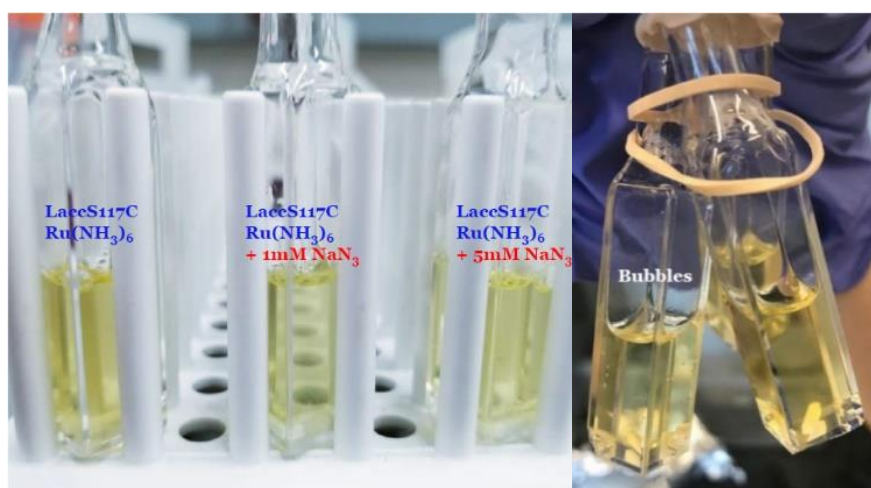


Figure 3.27. *Tth*-lac laser samples with NaN_3 upon irradiation with 455 nm blue LED lamps.

To test the azide inhibition effects, 1 mM and 5 mM NaN_3 were added to separate samples both containing 20 μM Ru-labeled *Tth*-lac and 20 mM $[\text{Ru}(\text{NH}_3)_6]\text{Cl}_3$. Although the samples containing NaN_3 were expected to stay green, they all turned yellow upon irradiation with the blue 455 nm LED lamps for an hour. From the UV-Vis spectrum, it was confirmed that the Cu_{T1} band around 605 nm was down to near 0. Nitrogen bubbles formed from NaN_3 photolysis were observed in the cuvette containing 5mM NaN_3 . Based on these observations, it is possible that the reduction of the enzyme is rather directly through the Cu_{T1} .

3.5. Intramolecular Long-range Copper Oxidation

Since the signals from the transient absorption spectroscopy using the arc lamp are compromised by the $\text{Ru}(\text{bpy})_2(\text{phen})^{2+}$ luminescence, a HeNe laser probe (632.8nm) was used to limit the collection of excess fluorescence and monitor the oxidation of Cu_{T1} . Hole transfer from Ru^{3+} species oxidizes Cu_{T1} , and the rise of Cu^{2+} signal is monitored by the transient absorption spectroscopy using a HeNe laser probe (632.8nm). Moreover, with the use of irreversible quenchers, back-reactions can be prevented and the electron/hole transfer processes occurring at longer time scales (ms to s) can be monitored. $[\text{Co}(\text{NH}_3)_5\text{Cl}]\text{Cl}_2$ was used as an irreversible quencher to monitor the long-range electron transfer through *Tth*-lac.

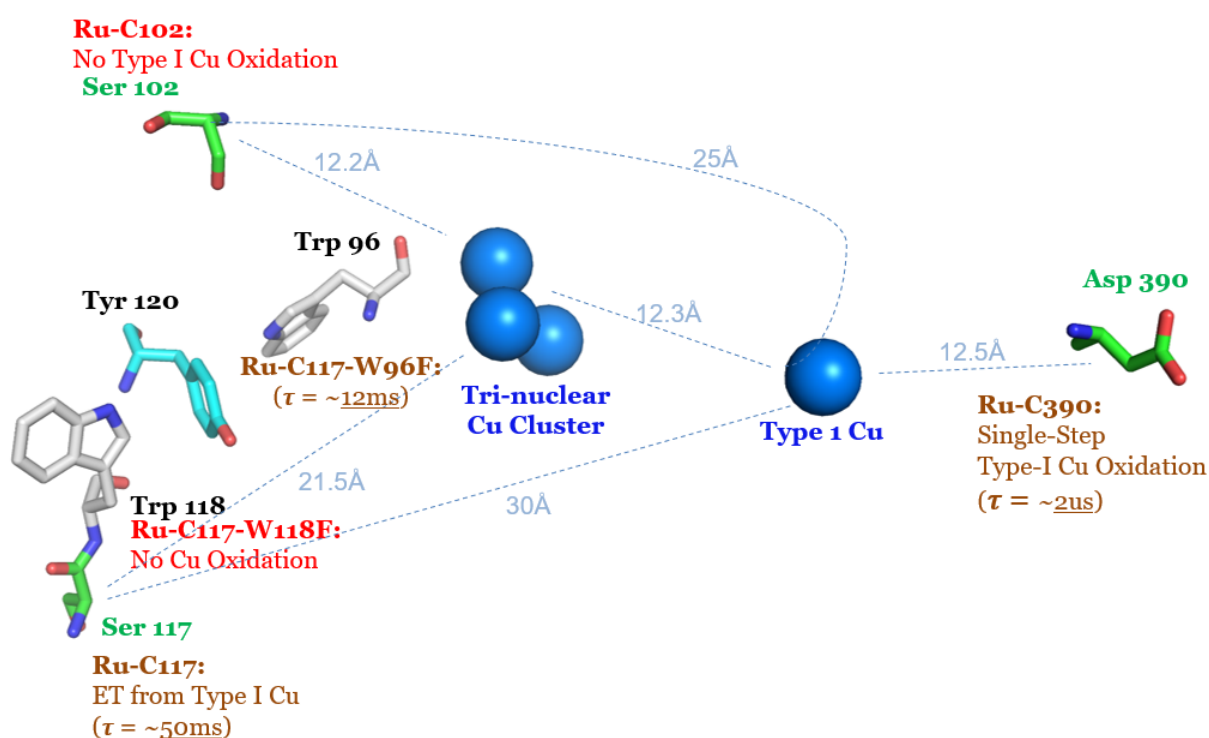


Figure 3.28. Electron/hole transfer kinetics of intramolecular long-range copper oxidation.

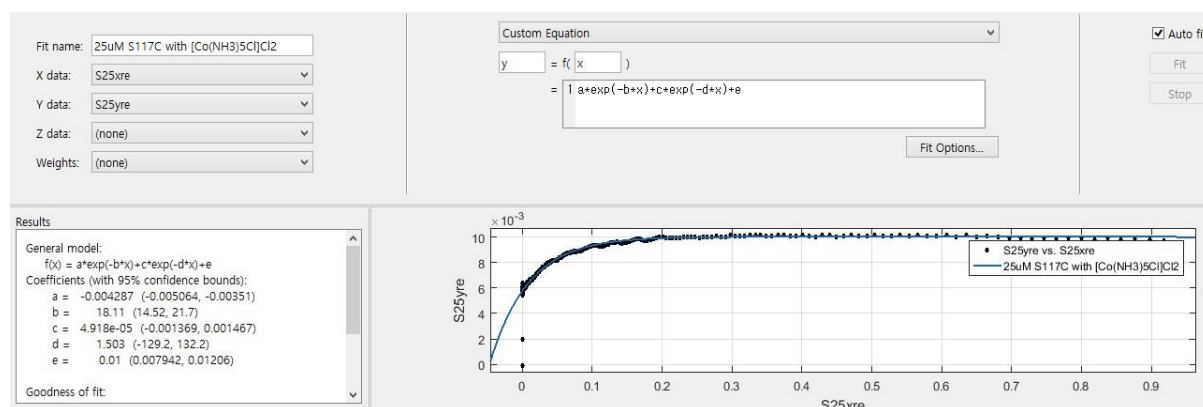
Figure 3.28 shows the overview of the electron/hole transfer kinetics data obtained so far. It can be seen that a long-range electron/hole transfer over a distance of 30 Å is possible in *Tth*-lac by hole hopping through Trp residues. When we inject holes into the system from a surface exposed Ru-C117 right next to the surface exposed W118, a signal for Trp radical and Trp radical cation appears in 50 μs time scale, and the hole goes on to oxidize Cu_{T1}. This clearly indicates that these chains of closely spaced Trp and Tyr can participate in electron/hole transfer, which could potentially have a role in enzyme catalysis itself as well.

3.6. Appendices

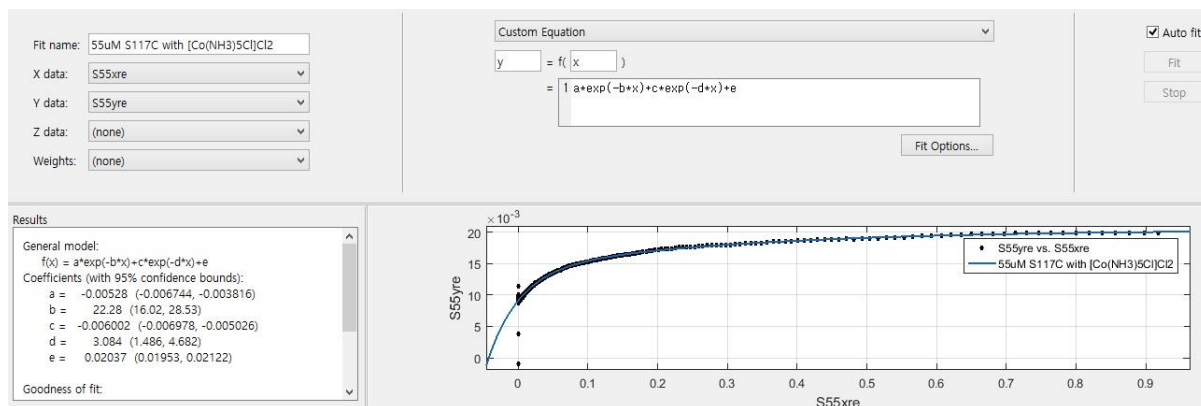
Kinetics Fitting Data: Testing the Concentration Dependence

- **Ru-C117 Enzyme**

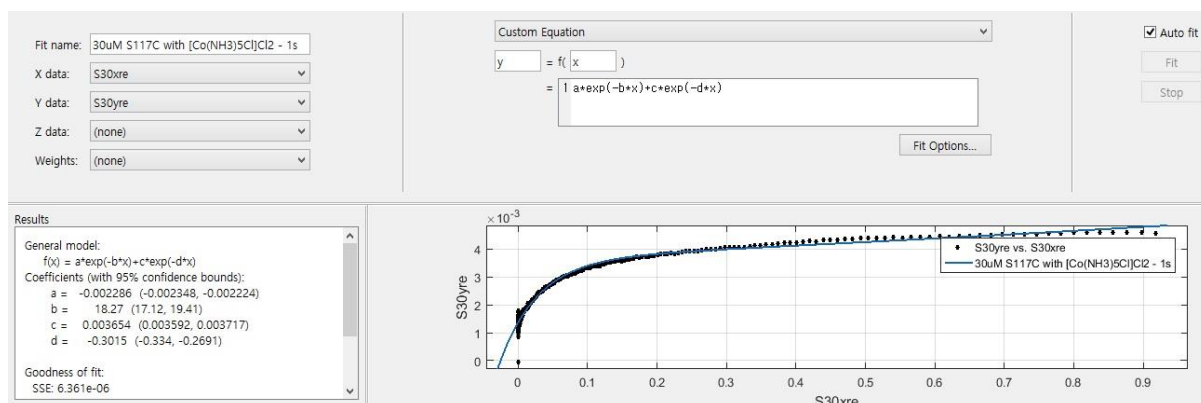
25uM – double path



55uM – double path



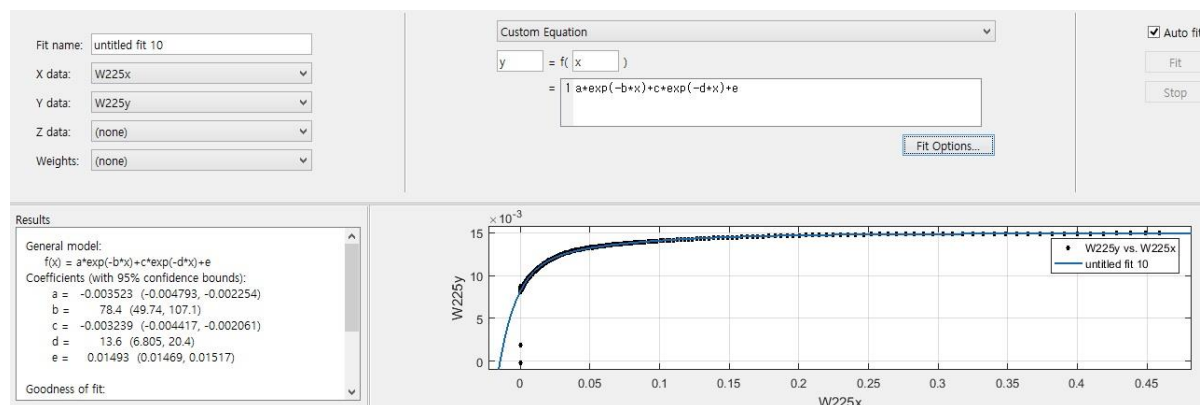
30uM – Single Path



The amplitude of the signal is proportional to the protein concentration. (The amplitude of the initial start point at the y-intercept may be attributed to the generation of Ru^{3+} and Trp radicals. As the concentration doubles, the signal amplitude for both the y-intercept and the asymptote doubles.) There seems no significant concentration dependence on rate constants for the rise of copper signal.

- **Ru-C117-W96F Enzyme**

25uM – double path



- **No copper oxidation observed with S117C-W118F and S102C Mutants**

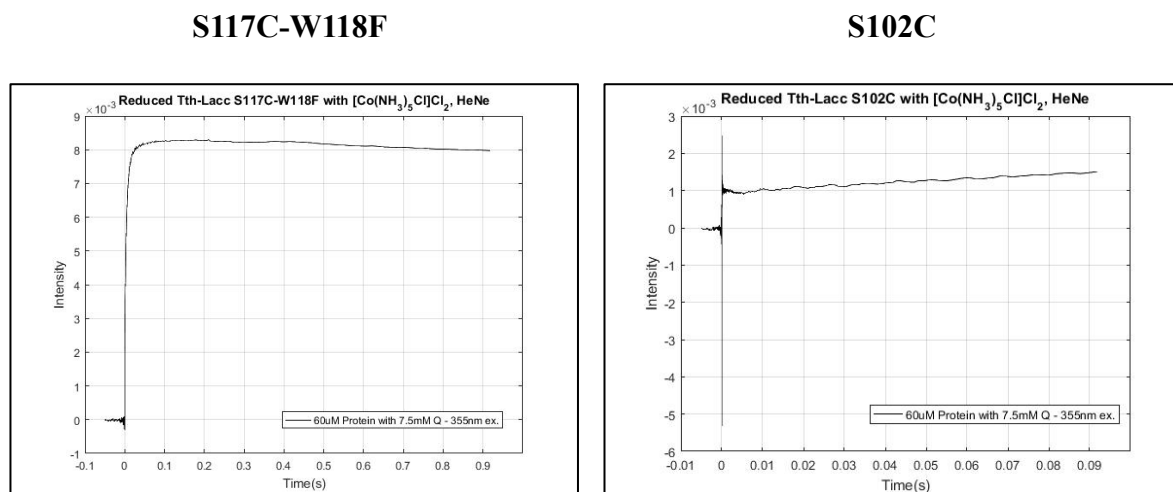


Figure 3.29. Transient absorption kinetics for S117C-W118F and S102C with HeNe.

Temperature effects on the single-step electron transfer through *Tth*-laccase

The effects of temperature on the intramolecular electron transfer through *Tth*-laccase was investigated by examining the kinetics of the photo-triggered single step oxidation of Cu_{T1} from 20°C to 65°C with a mutant labeled with a Ru-photosensitizer on the surface near Cu_{T1} over the distance of around 17 Å.

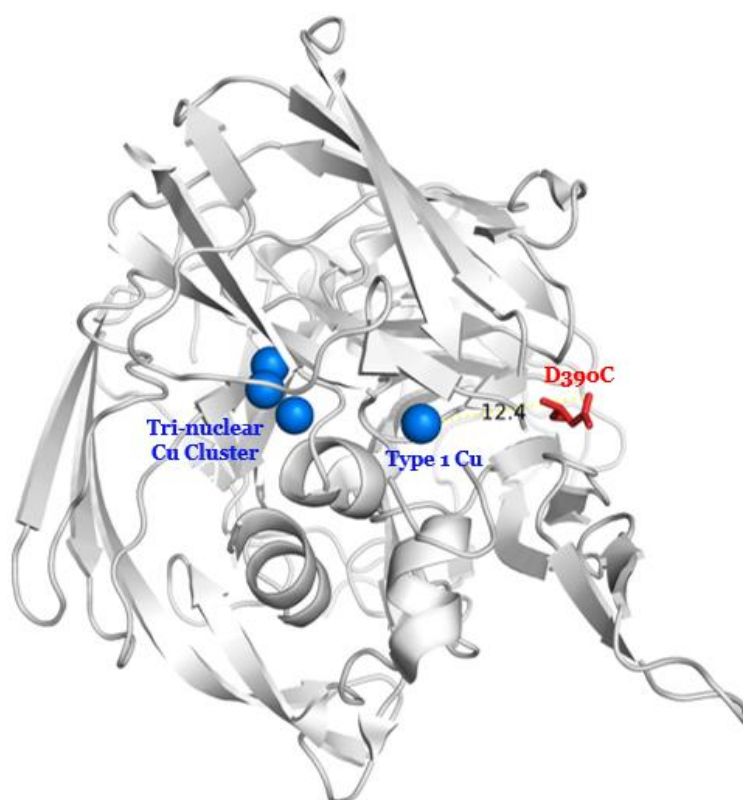


Figure 3.30. Ru-photosensitizer labeling site (red) near Cu_{T1} of *Tth*-lac.

In order to observe the photo-triggered oxidation of Cu_{T1} by injecting holes into the system in the laser flash-quench cycles, coppers in the Ru-C390 mutant were reduced using four-reducing equivalents of a deoxygenated cysteine solution (E_0' (cysteine/cystine) = -220 mV [12]). Since the protein is labeled with a photosensitizer, the photosensitizer excited state (*Ru²⁺) can be generated from Ru^{II}(bpy)₂(phen)²⁺ with a laser pulse, and the *Ru²⁺ species are quenched by

the oxidative quencher $\text{Ru}(\text{NH}_3)_6\text{Cl}_3$ to produce $\text{Ru}^{\text{III}}(\text{bpy})_2(\text{phen})^{3+}$. A single step hole transfer from these Ru^{3+} species oxidizes Cu_{T1} , and the rise of Cu^{2+} signal is monitored by the transient absorption spectroscopy using a HeNe laser probe (632.8 nm).

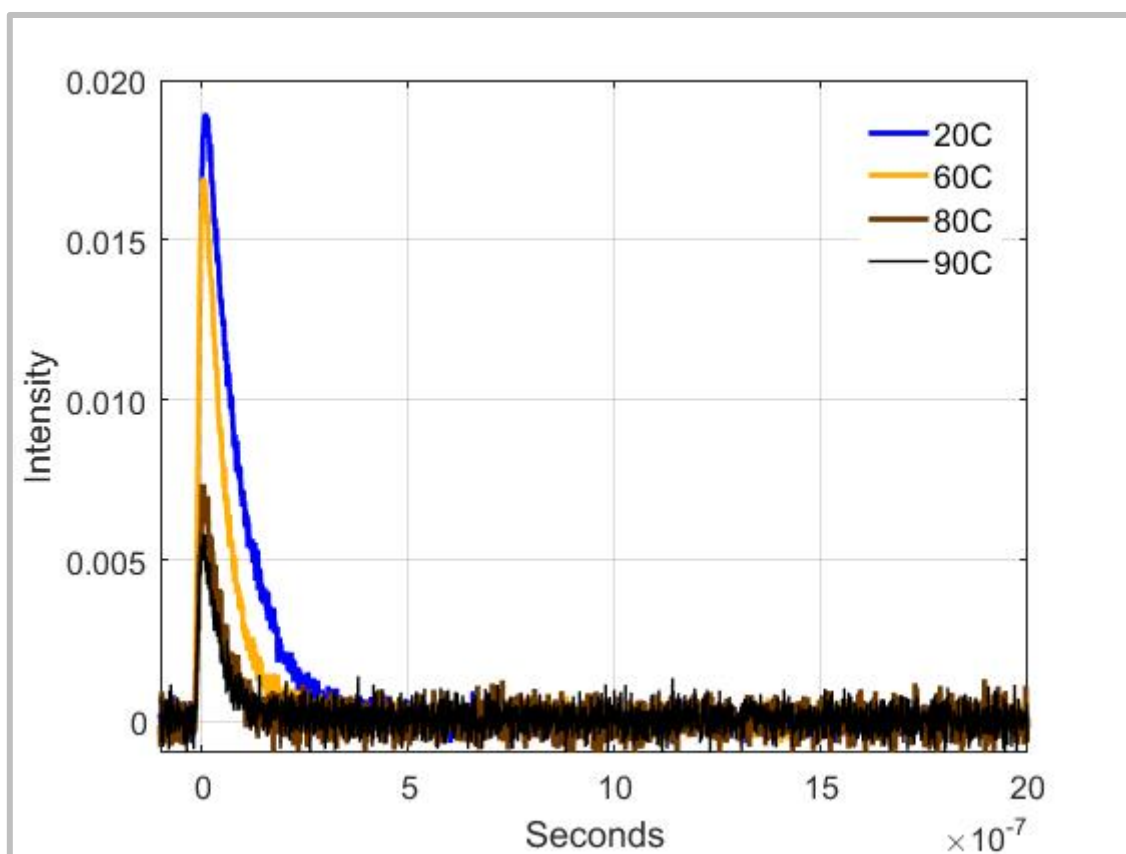


Figure 3.31. Luminescence kinetics of $[\text{Ru}(\text{bpy})_3\text{Cl}]\text{Cl}_2$

The luminescence kinetics indicate the decay of $^*\text{Ru}^{2+}$ to produce Ru^{3+} species by the quenching reaction with $\text{Ru}(\text{NH}_3)_6\text{Cl}_3$.

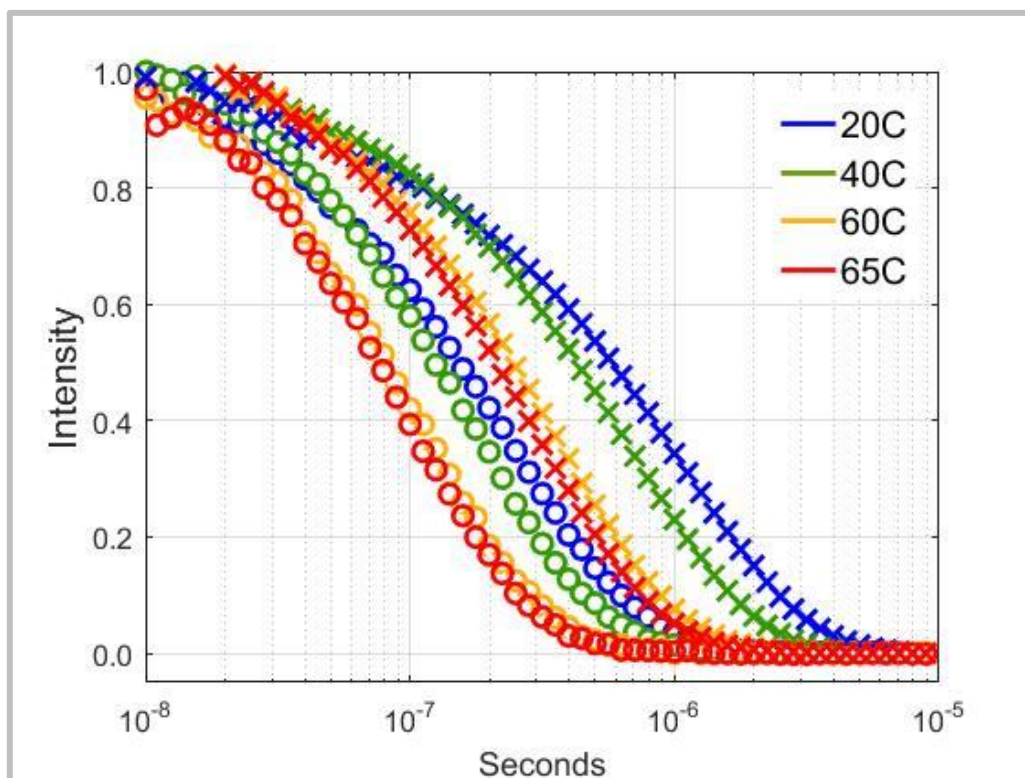


Figure 3.32. Luminescence kinetics of Ru-C390 at 650 nm.

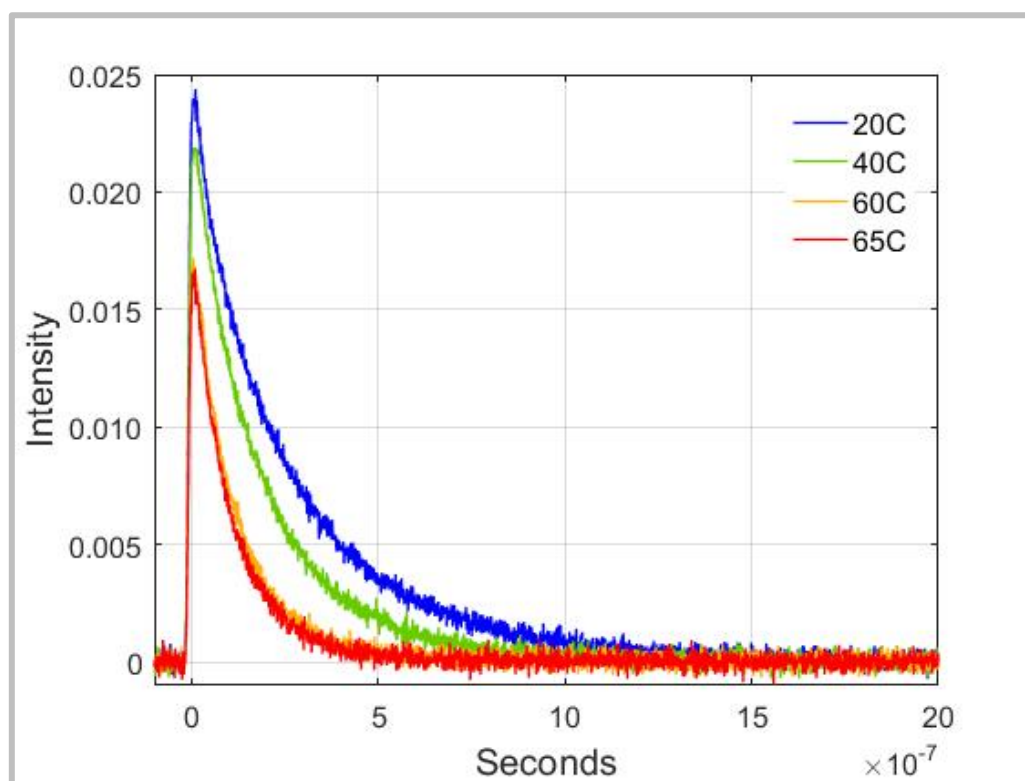


Figure 3.33. Luminescence kinetics of Ru-C390 at 650 nm.

The kinetics fitting**20C:**

$$f(x) = a \cdot \exp(-b \cdot x)$$

$$a = 0.02253$$

$$b = 3.5e+06$$

40C:

$$f(x) = a \cdot \exp(-b \cdot x)$$

$$a = 0.02146$$

$$b = 5e+06$$

60C:

$$f(x) = a \cdot \exp(-b \cdot x)$$

$$a = 0.01707$$

$$b = 8e+06$$

65C:

$$f(x) = a \cdot \exp(-b \cdot x)$$

$$a = 0.01685$$

$$b = 9e+06$$

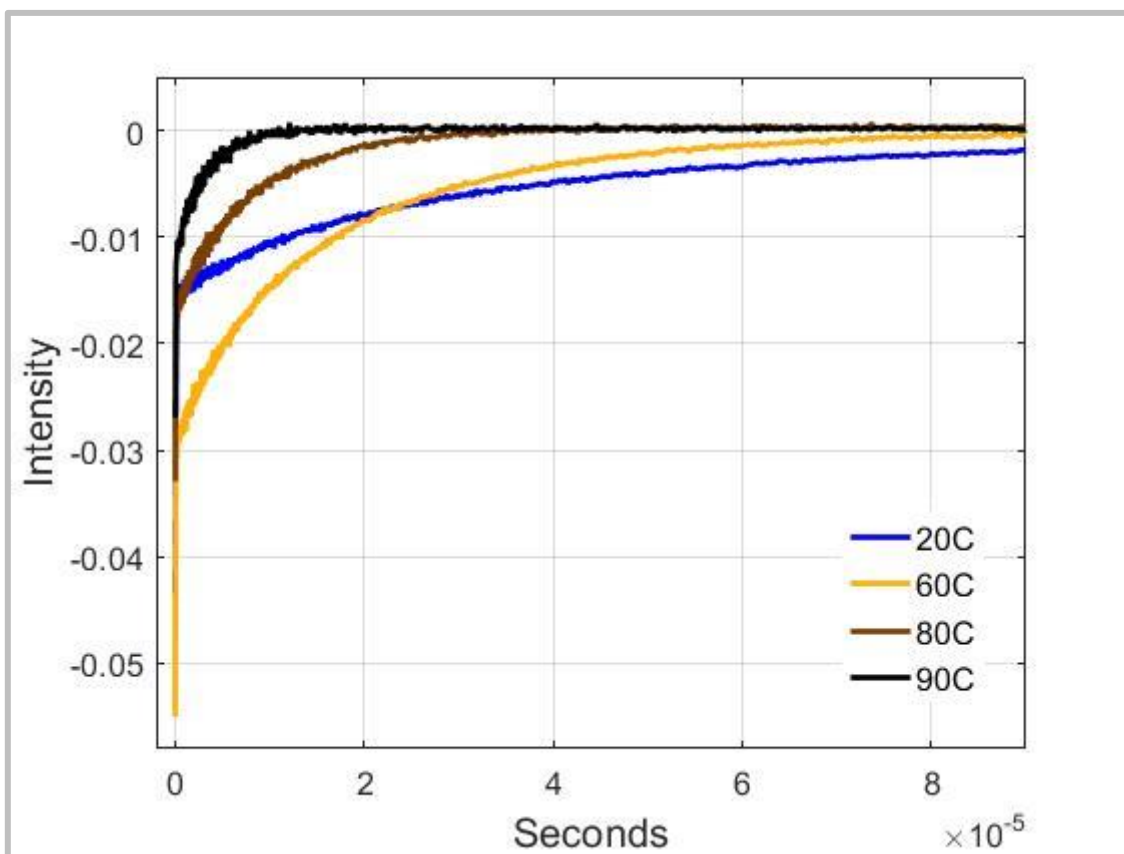


Figure 3.34. Transient absorption kinetics of $[\text{Ru}(\text{bpy})_3\text{Cl}]\text{Cl}_2$ at 450 nm.

450 nm is the Ru^{2+} -photosensitizer ground state absorption region, and a bleach is observed for the formation of Ru^{2+} excited state ($^*\text{Ru}^{2+}$). The bleach decays in about 100 ns to yield the species which are presumably Ru^{3+} -photosensitizer. The Ru^{3+} signals disappear as the hole gets transferred to oxidize Cu_{T1} at different rates at different temperature conditions. From the control experiment, it was concluded that temperatures above 80 °C are not suitable for studying the enzyme kinetics, since the thermal decay of the Ru^{3+} species happens faster than the single-step oxidation of Cu_{T1} kinetics.

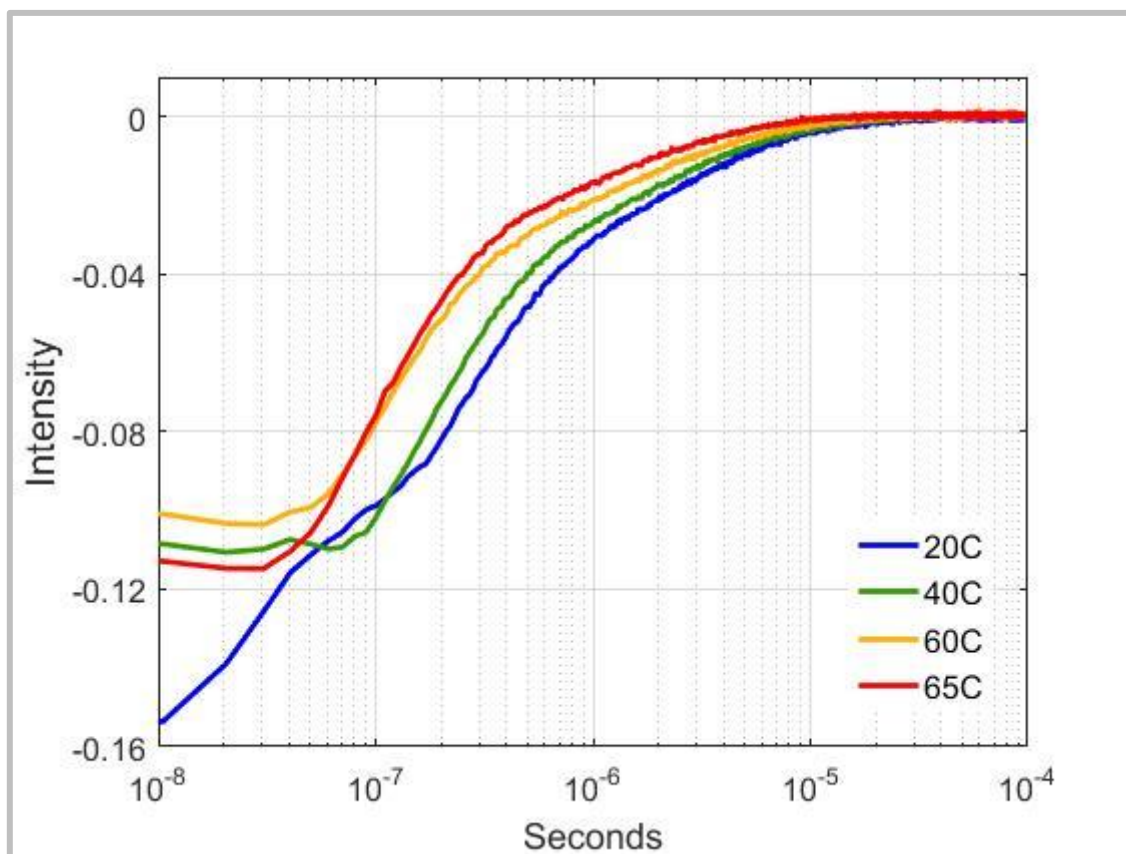


Figure 3.35. Transient absorption kinetics of Ru-C390 at 450 nm.

20C:

$$f(x) = a \cdot \exp(-3.5e6 \cdot x) + c \cdot \exp(-d \cdot x)$$

$$a = -0.09493$$

$$c = -0.03671$$

$$d = 2.544e+05$$

40C:

$$f(x) = a \cdot \exp(-5e6 \cdot x) + c \cdot \exp(-d \cdot x)$$

$$a = -0.1092$$

$$c = -0.03544$$

$$d = 3.199e+05$$

60C:

$$f(x) = a \cdot \exp(-8e6 \cdot x) + c \cdot \exp(-d \cdot x)$$

$$a = -0.1094$$

$$c = -0.03084$$

$$d = 3.6e+05$$

65C:

$$f(x) = a \cdot \exp(-9e6 \cdot x) + c \cdot \exp(-d \cdot x)$$

$$a = -0.1229$$

$$c = -0.02818$$

$$d = 4.73e+05$$

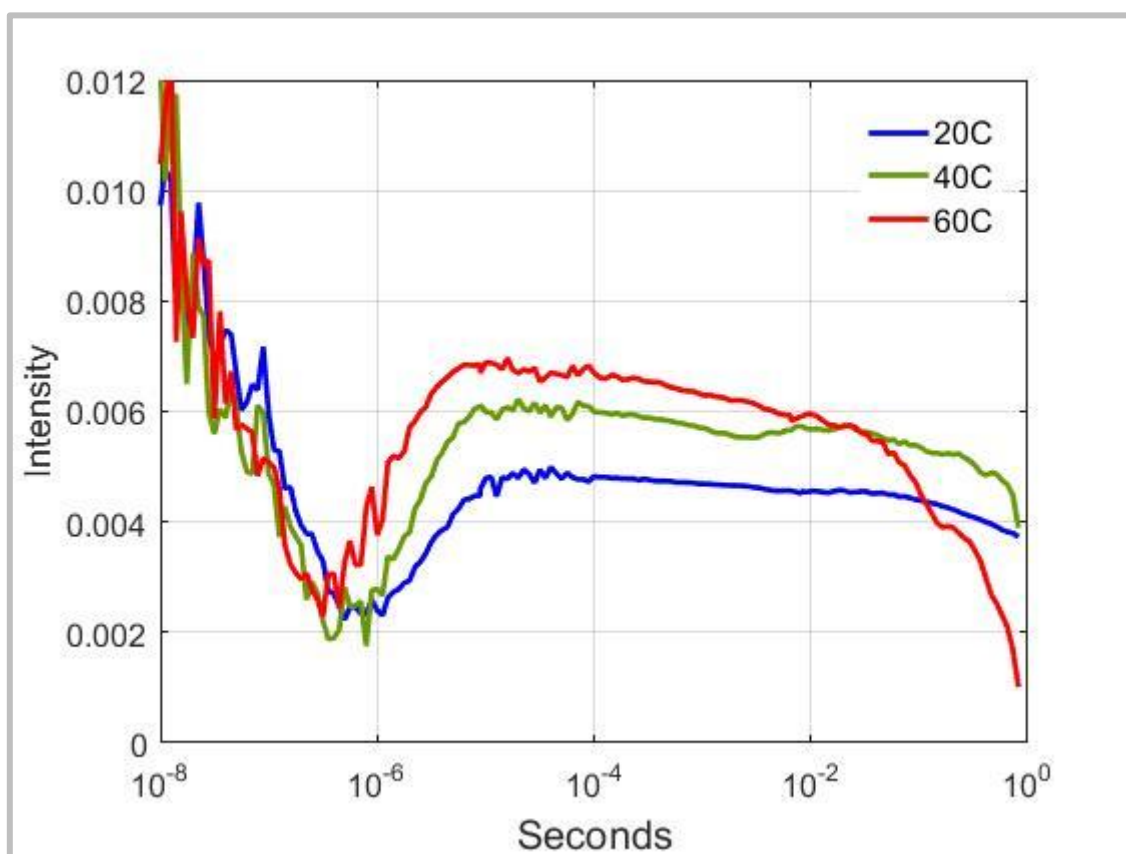


Figure 3.36. Transient absorption kinetics of Ru-C390 at 633 nm with HeNe.

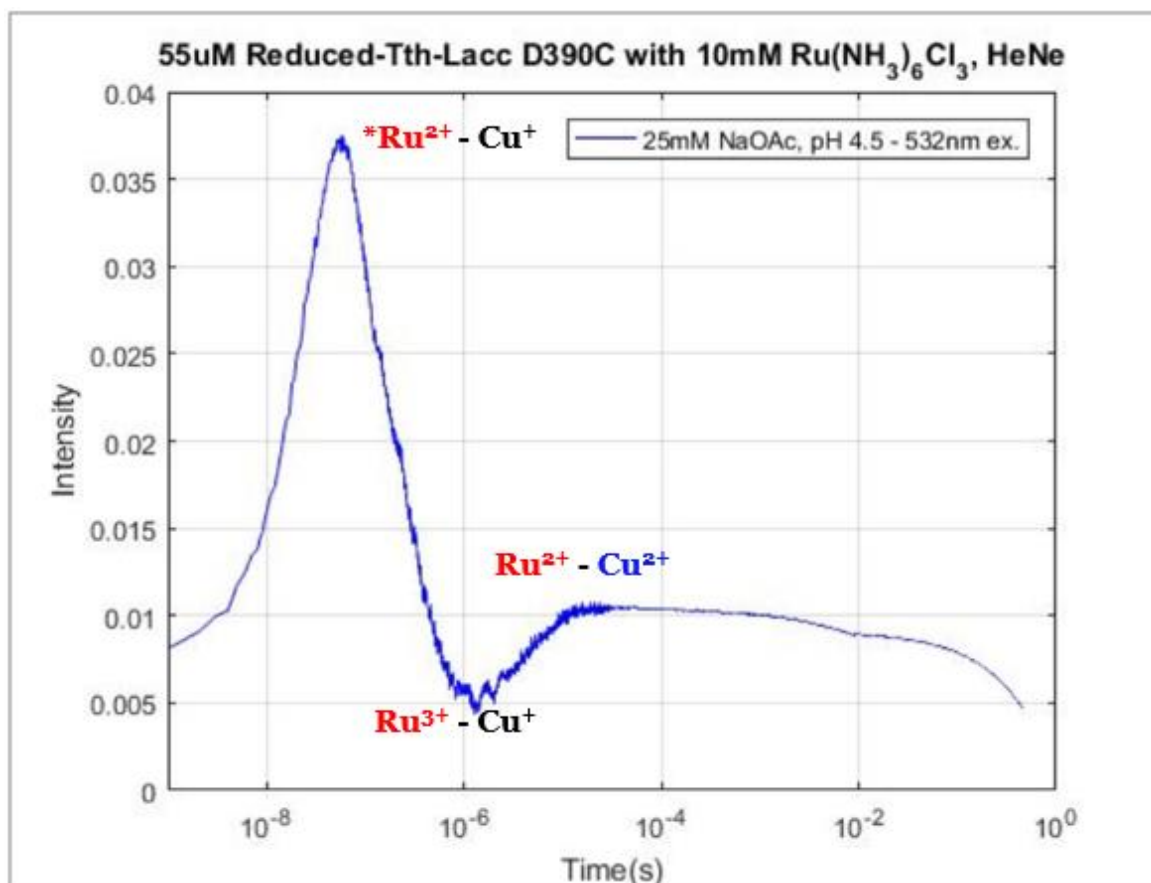


Figure 3.37. Kinetics profile of the Ru-C390 mutant obtained using a HeNe laser.

The absorption at 633nm at early times is due to the generation of $*\text{Ru}^{2+}$ species. The $*\text{Ru}^{2+}$ absorption decays as $*\text{Ru}^{2+}$ converts to Ru^{3+} in the oxidative quenching reaction with $\text{Ru}(\text{NH}_3)_6\text{Cl}_3$. A long-lasting positive signal is generated as Ru^{3+} oxidizes Cu^{1+} to Cu^{2+} . This rise of positive signal occurs at different rates at different temperatures. The curve fitting was done by fixing the first two exponents to the luminescence decay and the 450 nm transient absorption kinetics. The rest of the exponents reflects the recombination kinetics. Although the temperature effects were attempted to be investigated, the kinetics of copper oxidation obtained with $\text{Ru}(\text{NH}_3)_6\text{Cl}_3$ as a quencher seemed to have been contaminated by the back reaction kinetics. Further analysis might be required with irreversible quenchers to prevent the back reactions.

20 °C:

$$f(x) = a \cdot \exp(-3.5e6 \cdot x) + c \cdot \exp(-2.5e5 \cdot x) + e \cdot \exp(-f \cdot x) + g \cdot \exp(-h \cdot x)$$

$$a = 0.01497$$

$$c = -0.009613$$

$$e = -0.0035$$

$$f = 2.638e+05$$

$$g = 0.0002797$$

$$h = 669$$

$$k = 0.004787$$

$$p = 0.6307$$

40 °C:

$$f(x) = a \cdot \exp(-5e6 \cdot x) + c \cdot \exp(-3.2e5 \cdot x) + e \cdot \exp(-f \cdot x) + g \cdot \exp(-h \cdot x)$$

$$a = 0.01611$$

$$c = -0.008906$$

$$e = -0.004615$$

$$f = 3.776e+05$$

$$g = 0.000367$$

$$h = 2722$$

$$k = 0.005931$$

$$p = 1.64$$

60 °C:

$$f(x) = a \cdot \exp(-8e6 \cdot x) + c \cdot \exp(-3.6e5 \cdot x) + e \cdot \exp(-f \cdot x) + g \cdot \exp(-h \cdot x)$$

$$a = 0.01334$$

$$c = -0.006047$$

$$e = -0.004079$$

$$f = 6.652e+05$$

$$g = 0.0006545$$

$$h = 1413$$

$$k = 0.006136$$

$$p = 3.899$$

3.7. References

- [1] K. Miyazaki, "A hyperthermophilic laccase from *Thermus thermophilus* HB27," *Extremophiles*, vol. 9, no. 6, pp. 415–425, Dec. 2005, doi: 10.1007/s00792-005-0458-z.
- [2] L. Zheng, U. Baumann, and J. L. Reymond, "An efficient one-step site-directed and site-saturation mutagenesis protocol," *Nucleic Acids Res.*, vol. 32, no. 14, 2004, doi: 10.1093/nar/gnh110.
- [3] H. Serrano-Posada, S. Centeno-Leija, S. P. Rojas-Trejo, C. Rodríguez-Almazán, V. Stojanoff, and E. Rudiño-Piñera, "X-ray-induced catalytic active-site reduction of a multicopper oxidase: Structural insights into the proton-relay mechanism and O₂-reduction states," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 71, no. Pt 12, pp. 2396–2411, Nov. 2015, doi: 10.1107/S1399004715018714.
- [4] F. N. Castellano, J. D. Dattelbaum, and J. R. Lakowicz, "Long-lifetime Ru(II) complexes as labeling reagents for sulfhydryl groups," *Anal. Biochem.*, vol. 255, no. 2, pp. 165–170, Jan. 1998, doi: 10.1006/abio.1997.2468.
- [5] B. P. Sullivan, D. J. Salmon, and T. J. Meyer, "Mixed Phosphine 2,2'-Bipyridine Complexes of Ruthenium," *Inorg. Chem.*, vol. 17, no. 12, pp. 3334–3341, 1978, doi: 10.1021/ic50190a006.
- [6] M. E. Ener, H. B. Gray, and J. R. Winkler, "Hole Hopping through Tryptophan in Cytochrome P450," *Biochemistry*, vol. 56, no. 28, pp. 3531–3538, Jul. 2017, doi: 10.1021/acs.biochem.7b00432.

- [7] G. Bulaj, T. Kortemme, and D. P. Goldenberg, "Ionization-reactivity relationships for cysteine thiols in polypeptides," *Biochemistry*, vol. 37, no. 25, pp. 8965–8972, Jun. 1998, doi: 10.1021/bi973101r.
- [8] F. Roulling, A. Godin, A. Cipolla, T. Collins, K. Miyazaki, and G. Feller, "Activity–stability relationships revisited in blue oxidases catalyzing electron transfer at extreme temperatures," *Extremophiles*, vol. 20, no. 5, pp. 621–629, Sep. 2016, doi: 10.1007/s00792-016-0851-9.
- [9] S. M. Jones and E. I. Solomon, "Electron transfer and reaction mechanism of laccases," *Cellular and Molecular Life Sciences*, vol. 72, no. 5. Birkhauser Verlag AG, pp. 869–883, Mar. 01, 2015, doi: 10.1007/s00018-014-1826-6.
- [10] G. Metzker, I. De Aguiar, S. C. Martins, M. S. Schultz, L. C. G. Vasconcellos, and D. W. Franco, "Electrochemical and chemical aspects of ruthenium(II) and (III) amines in basic solution: The role of the ruthenium(IV) species," *Inorganica Chim. Acta*, vol. 416, pp. 142–146, May 2014, doi: 10.1016/j.ica.2014.03.013.
- [11] L. Michaelis and E. S. Hill, "The viologen indicators," *J. Gen. Physiol.*, vol. 16, no. 6, pp. 859–873, Jul. 1933, doi: 10.1085/jgp.16.6.859.
- [12] P. C. Jocelyn, "The Standard Redox Potential of Cysteine-Cystine from the Thiol-Disulphide Exchange Reaction with Glutathione and Lipoic Acid," *Eur. J. Biochem.*, vol. 2, no. 3, pp. 327–331, Oct. 1967, doi: 10.1111/j.1432-1033.1967.tb00142.x.
- [13] R. C. Murray and P. A. Rock, "The determination of the ferrocyanide-ferricyanide standard electrode potential at 25°C in cells without liquid junction using cation-sensitive glass electrodes," *Electrochim. Acta*, vol. 13, no. 4, pp. 969–975, Apr. 1968, doi: 10.1016/0013-4686(68)85028-5.

- [14] S. Katayama, A. Myoga, and Y. Akahori, "Swelling behavior of amphoteric gel and the volume phase transition," *J. Phys. Chem.*, vol. 96, no. 11, pp. 4698–4701, 1992, doi: 10.1021/j100190a101.
- [15] M. Z. Hoffman, F. Bolletta, L. Moggi, and G. L. Hug, "Rate Constants for the Quenching of Excited States of Metal Complexes in Fluid Solution," *J. Phys. Chem. Ref. Data*, vol. 18, no. 1, pp. 219–543, Oct. 1989, doi: 10.1063/1.555840.
- [16] Y. T. Liao, L. Zang, and D. L. Akins, "Catalytic reduction of methyl viologen by sulfide ion within MCM-41," *Catal. Commun.*, vol. 6, no. 2, pp. 141–145, Feb. 2005, doi: 10.1016/j.catcom.2004.11.015.
- [17] C. Aubert, M. H. Vos, P. Mathis, A. P. M. Eker, and K. Brettel, "Intraprotein radical transfer during photoactivation of DNA photolyase," *Nature*, vol. 405, no. 6786, pp. 586–570, Jun. 2000, doi: 10.1038/35014644.
- [18] C. Johannes and A. Majcherczyk, "Laccase activity tests and laccase inhibitors," *J. Biotechnol.*, vol. 78, no. 2, pp. 193–199, Mar. 2000, doi: 10.1016/S0168-1656(00)00208-X.

Chapter 4

**Trp/Tyr pair protects *Thermus thermophilus* HB27 laccase
from oxidative damage**

4.1. Abstract

Dioxygen-utilizing enzymes run the risk of damage by reactive oxygen species (ROS) produced in uncoupled catalytic turnover. Tryptophan (Trp) and tyrosine (Tyr) residues might act as internal antioxidants to protect oxidases and oxygenases from ROS damage. *Thermus thermophilus* HB27 laccase (*Tth-lac*) is a multicopper oxidase (MCO) containing a blue (type 1) copper and a trinuclear copper (TNC) cluster that couple substrate oxidation with dioxygen reduction to water. There is a Trp/Tyr pair near (3.5 Å) the TNC; and there also is a chain of closely spaced Trp and Tyr residues from the TNC to the surface. LCMS followed by peptide analysis of deactivated samples identified the TNC-proximal Trp/Tyr pair as an oxidation site, and curvature in the activity assay profile as well as a reduction in total turnover number for the double mutant (W133F/Y134F) indicate that W133 and Y134 move oxidizing holes away from the TNC to sites that are less critical for enzyme function.

4.2. Introduction

Thermus thermophilus HB27 laccase (*Tth-lac*), an enzyme from a bacterium that grows optimally at 65°C, is active even at elevated temperatures (1). *Tth-lac* belongs to a family of three-domain multicopper oxidases (MCOs) that have four copper sites involved in electron transfer and enzyme catalysis (2,3). According to the consensus mechanism (4-6), substrate oxidation occurs at a type 1 copper (Cu_{T1}), followed by long-range electron transfer (13 Å) to a trinuclear copper (TNC) cluster, where dioxygen is reduced to water (Figure 4.1).

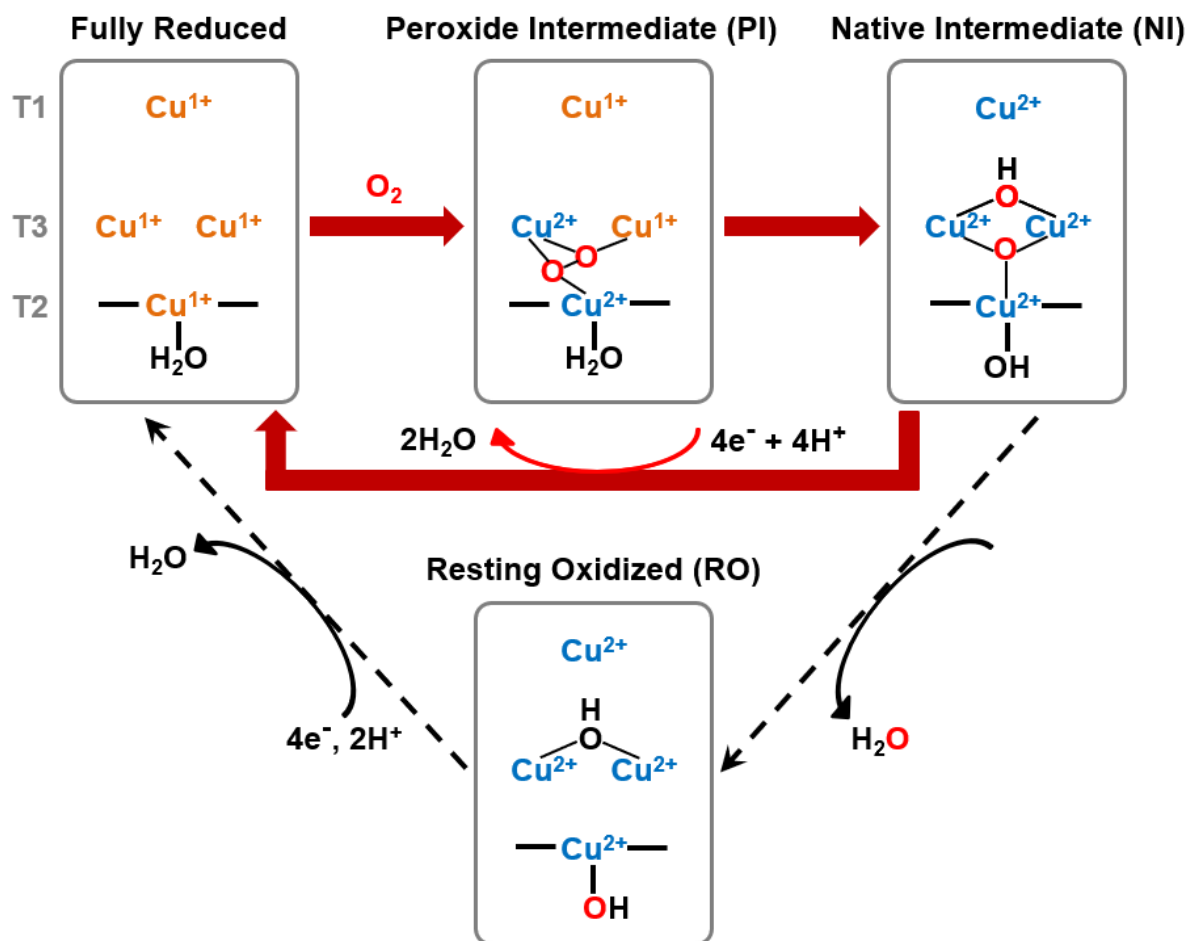


Figure 4.1. Proposed steps in the catalytic cycle (red arrows). NI decay to RO is off-path; upon 4-electron reduction, RO re-enters the cycle (black dashed arrows) (5).

When an easily oxidized substrate is present in high concentration, the “ping-pong bi bi” mechanism (4,7) outlined in Figure 4.1 leads to efficient 4-electron oxidase chemistry. With low concentrations of refractory substrates under normal dioxygen tension, however, the reaction of partially reduced enzyme with O_2 could lead to deleterious side reactions. Our interest centers on what happens when the enzyme fails to oxidize substrate during turnover, as it is likely that reactive oxygen species (ROS) such as O_2^- or H_2O_2 would be produced. In related work on cytochrome P450, we suggested that Trp/Tyr chains move oxidizing holes from the heme to distant sites near cellular reductants, thereby preventing ROS formation and

subsequent enzyme inactivation (8-11). An analogous protective antioxidant role has been implicated for Tyr residues in the two-domain laccase from *Streptomyces coelicolor* (12) and the six-domain blood protein ceruloplasmin (13).

We draw attention to a Trp/Tyr pair near (3.5 Å (3)) the *Tth*-lac TNC (W133-Y134) as well as a chain of closely spaced Trp/Tyr residues (W96, Y120, W118) that could couple the copper cluster to potential reductants at the enzyme surface (Figure 4.2). These Trp/Tyr chains are found in most MCOs; indeed, among 25 three-domain laccase structures available from the Protein Data Bank, all except two have Trp or Trp/Tyr pairs adjacent to the TNC (8). In addition to protecting MCOs from oxidative degradation, these chains could play a role in enzyme catalysis (14).

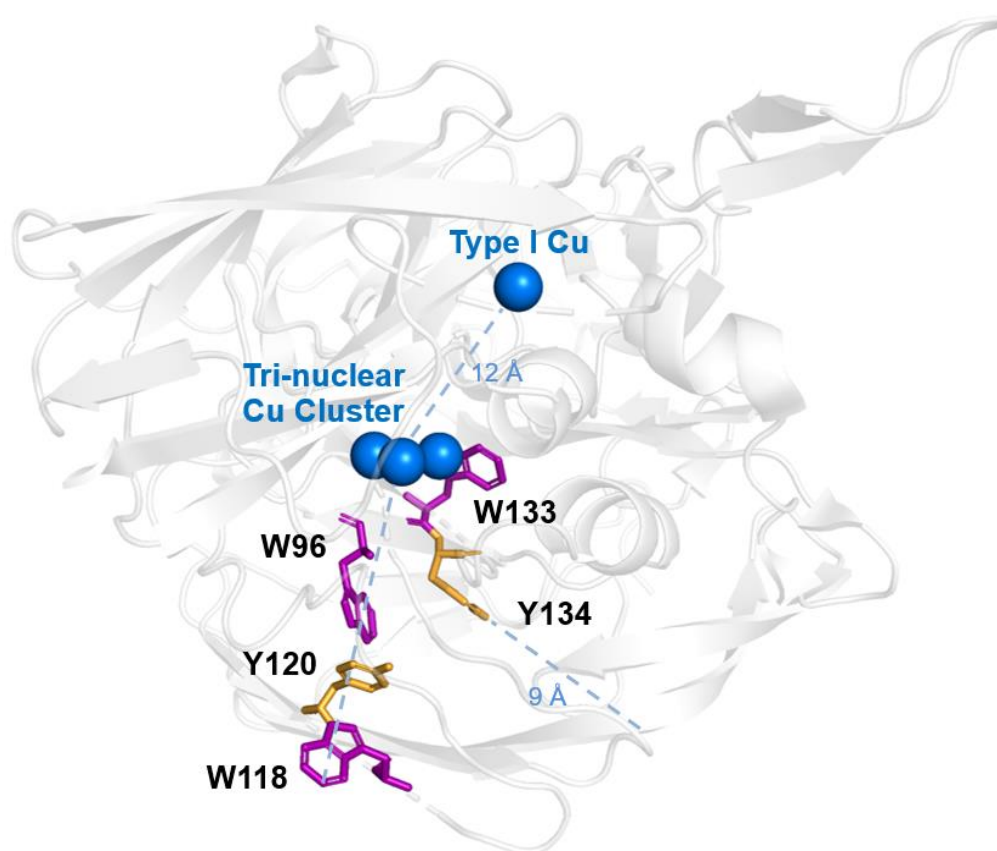


Figure 4.2. A Trp/Tyr chain (W96/Y120/W118) and a Trp/Tyr pair (W133-Y134) near the trinuclear copper cluster (PDB: 2XU9) (3).

4.3. Results and Discussion

Enzyme activity. We examined the impact on enzyme activity of Trp and Tyr mutations near the TNC in *Tth-lac*. In the aerobic oxidation of ABTS^{2-} catalyzed by WT and four single-site enzyme mutants (W133F, Y134F, W118F, W96F), the absorbance of $\text{ABTS}^{\bullet-}$ increases linearly during the first 5 min of reaction (Figure 4.3A). Of interest is that at early stages, the reaction catalyzed by the W133F/Y134F double mutant exhibits slight curvature (Figure 4.3B), indicative of a modified reaction pathway or gradual enzyme deactivation. As the activities of WT and all mutants exhibit saturation with increasing ABTS^{2-} concentration, we fit the observed kinetics to a Michaelis-Menten model (Figure 4.4). At high ABTS^{2-} concentrations, the initial rates for WT and the four single-site mutants plateau in the range $5\text{-}7 \times 10^{-8} \text{ M s}^{-1}$ ($[\textit{Tth-lac}] = 10^{-7} \text{ M}$), whereas for the W133F/Y134F double mutant, the limiting rate ($2 \times 10^{-8} \text{ M s}^{-1}$) is less than half that of the WT enzyme (Table 4.1). Virtually identical circular dichroism spectra of WT and W133F/Y134F confirm that the native enzyme structure is intact in the double mutant (Figure S4.10). When either of the residues in the pair is still present, the activity is close to that of WT, demonstrating that changes at either W133 or Y134 site alone do not impair enzyme function.

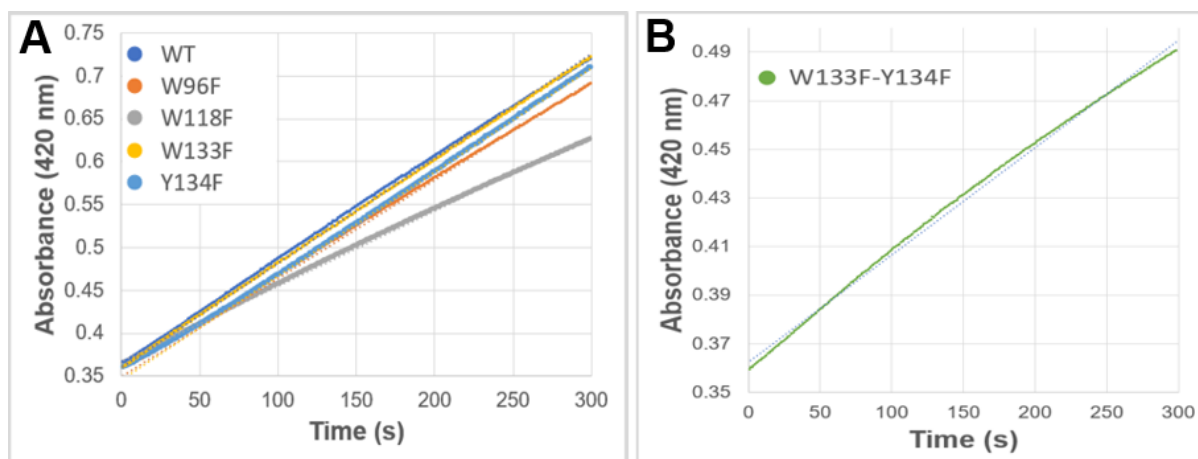


Figure 4.3. Activity assay profiles at early times (initial 5 min) for WT and mutant enzymes (500 nM [E], 3 mM ABTS^{2-} in MES buffer, pH 5.3, $[\text{O}_2] = 240 \mu\text{M}$).

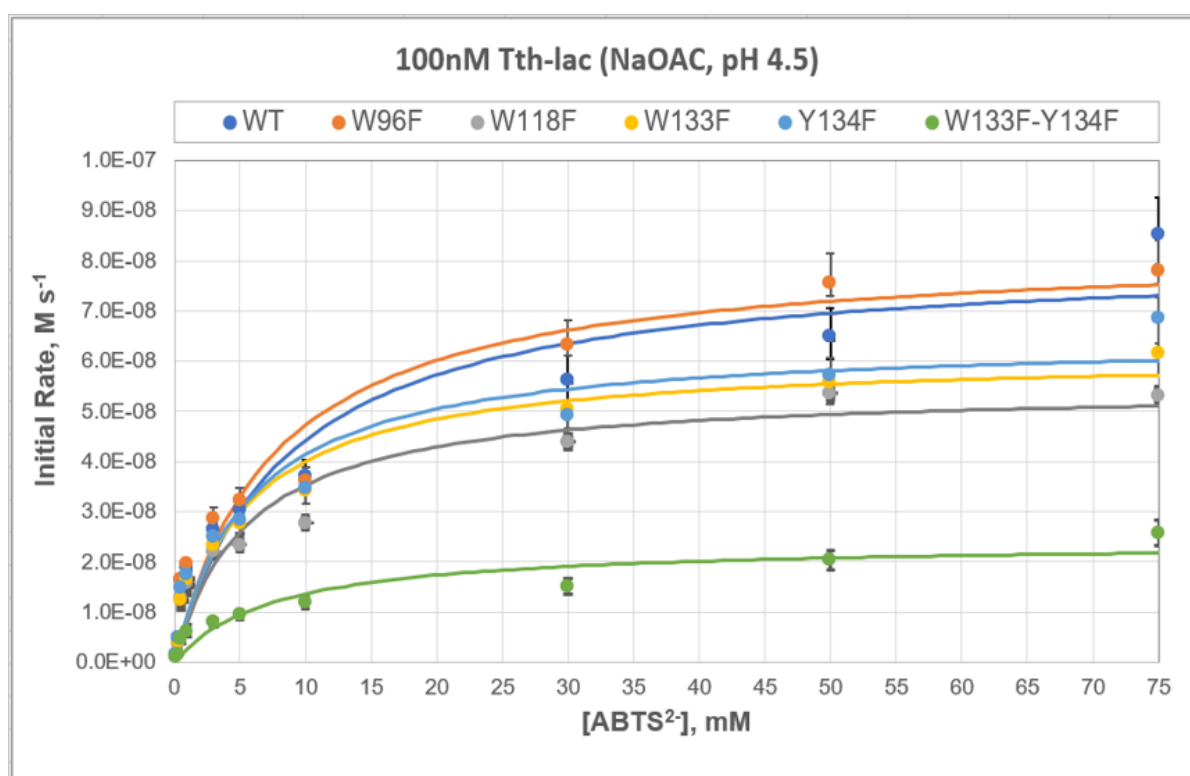


Figure 4.4. Enzyme activity assays with ABTS^{2-} (mM, $\pm 0.2 \%$) at pH 4.5 for WT and Trp/Tyr mutants. Solid lines are least squares fits to a Michaelis-Menten model. (Measurements made in MES buffer at pH 5.3 are shown in Table S4.7 and Figure S4.9.)

Table 4.1. Catalytic efficiencies: WT and the Trp/Tyr mutants in NaOAc, pH 4.5.

	WT	W96F	W118F	W133F	Y134F	W133F/Y134F
k_{cat} (s^{-1})	8.0×10^{-1} $\pm 8 \times 10^{-2}$	8.3×10^{-1} $\pm 7 \times 10^{-2}$	5.5×10^{-1} $\pm 1 \times 10^{-2}$	6.1×10^{-1} $\pm 5 \times 10^{-2}$	6.5×10^{-1} $\pm 6 \times 10^{-2}$	2.4×10^{-1} $\pm 2 \times 10^{-2}$
K_{m} (M)	7.9×10^{-3} $\pm 2 \times 10^{-4}$	7.5×10^{-3} $\pm 2 \times 10^{-4}$	5.7×10^{-3} $\pm 3 \times 10^{-4}$	5.0×10^{-3} $\pm 2 \times 10^{-4}$	5.7×10^{-3} $\pm 2 \times 10^{-4}$	7.2×10^{-3} $\pm 5 \times 10^{-4}$
$k_{\text{cat}}/K_{\text{m}}$ $\text{M}^{-1}\text{s}^{-1}$	9.8×10^1 $\pm 3 \times 10^0$	1.1×10^2 $\pm 1 \times 10^1$	9.6×10^1 $\pm 4 \times 10^0$	1.2×10^2 $\pm 1 \times 10^0$	1.1×10^1 $\pm 1 \times 10^0$	3.3×10^1 $\pm 6 \times 10^0$

Tth-lac catalysis involves two substrates, ABTS^{2-} and O_2 (Figure 4.1). Employing the W133F/Y134F double mutant, we examined ABTS^{2-} oxidation kinetics at reduced O_2 concentration. Importantly, in the saturation regime of substrate concentration ($[\text{ABTS}^{2-}] = 75$ mM; $[\textit{Tth}\text{-lac}] = 500$ nM), the reaction rate is unaffected by a fourfold decrease in O_2 concentration (Table 4.2), demonstrating that the reduction in k_{cat} is not due to less efficient reaction with co-substrate O_2 . It is unlikely that ABTS^{2-} oxidation by the Cu_{T1} center would be impacted by the W133F/Y134F mutation, but we cannot rule out potential effects on Cu_{T1} -to-TNC electron transfer kinetics. Notably, neither W133F nor Y134F loses WT activity, confirming that only one of the two redox-active residues is required for enzyme catalysis.

Table 4.2. Activity assay: double mutant under O_2 limiting conditions.

[ABTS] = 75 mM	[O_2]	Initial Rate, (M/s)
W133F/Y134F (500 nM)	240 μM (100% aerated buffer)	1.22×10^{-7} $\pm 2 \times 10^{-9}$
	60 μM (25% aerated buffer)	1.22×10^{-7} $\pm 2 \times 10^{-9}$

Enzyme Deactivation. If the Tyr/Trp pair near the TNC is providing protection from enzyme damage by ROS, we anticipated that replacing these residues with Phe would lead to attenuated enzyme survival. We ran the *Tth*-lac catalyzed ABTS²⁻ oxidation with WT, W133F, and W133F/Y134F enzymes until they were no longer active. Total turnovers for WT and W133F were virtually identical (WT, 8.5×10^3 ; W133F, 8.0×10^3 turnovers), but the double mutant survived for just 4.7×10^3 turnovers.

LCMS data for inactivated WT and W133F enzymes revealed mass increases of 32 D for WT and 16 D for W133F, consistent with enzyme oxidation (Figure S4.2). To identify the oxidation sites, we examined the peptides from tryptic digestion of inactivated enzymes. MS/MS spectra revealed oxidation on a peptide containing residues 127 through 141 (ELAGTFWYHPHLHGR) (Figure S4.3). The mass of the fragment from the WT enzyme is 32 D greater than predicted on the basis of the enzyme sequence; and a mass increase of 16 D was found for the same fragment from the inactivated W133F mutant (Figure S4.4 and Figure S4.5). The results confirm that oxidation occurs at W133 in the WT enzyme; and Y134 is the second site of oxidation, which in turn indicates that these residues may aid enzyme survival by directing highly oxidizing holes away from the TNC when the enzyme fails to oxidize substrate. The oxidation of Y134 during W133F enzyme turnover clearly demonstrates that the oxidizing hole can hop from the TNC to Y134 even in the absence of W133; and, as Y134 is less than 10 Å from the enzyme surface, rapid electron transfer from solution reductants would prolong enzyme activity.

4.4. Concluding Remarks

Multicopper oxidases have the remarkable capacity to catalyze the aerobic oxidation of diverse high-potential substrates. However, as a tradeoff, the utilization of dioxygen can make the enzyme highly susceptible to oxidative damage. Since we found that total turnover numbers for the W133F/Y134F mutant are dramatically lower than those for WT, we conclude that the Trp/Tyr pair prolongs enzyme survival by scavenging highly oxidizing holes generated in the TNC. When the enzyme fails to oxidize substrates during catalysis, these strongly oxidizing holes are rapidly reduced by one of the residues in the pair.

Tryptophan and tyrosine, along with methionine and selenocysteine, are believed to be the last four amino acids introduced into the genetic code (15), possibly in response to the atmospheric transformation produced by oxygenic photosynthesis (16). Whatever the cause, there is mounting evidence that Trp, Tyr, and Met residues perform important protective (antioxidant) roles in a wide variety of metalloenzymes (8-13, 17-23).

4.5. Materials and Methods

Enzyme Preparation. *Thermus thermophilus* HB27 laccase was expressed in *E. coli* and purified following a previously reported procedure with slight modifications (24). Each mutant was obtained using standard site-directed mutagenesis protocols with modifications in the PCR parameters depending on the K_m of the primers designed for each mutant (25).

Activity Assays. Prior to activity assay measurements, a 30 μ M enzyme stock solution with 1 mM CuSO_4 in 20 mM Tris at pH 8 was prepared. The substrate for activity assays was the dianion of 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS^{2-}), a common

oxidase substrate (1,26). For a 3 mL assay sample containing 100 nM enzyme with varying ABTS²⁻ concentrations, 10 μ L of 30 μ M enzyme stock solution were added to 1990 μ L buffer (sodium acetate (NaOAc), pH 4.5 or MES, pH 5.3) immediately before measurements. Reaction was initiated by adding 1 mL ABTS²⁻ stock solution of the appropriate concentration. Absorption increases accompanying formation of ABTS^{•-} were monitored every 3 s for 2 - 3 min (and also in separate experiments at 5 min) at 24 °C using a UV-vis spectrometer equipped with a thermostatted sample compartment. Initial rates (M/s) were determined from the slopes of absorbance vs. time plots (first 10 – 30 s) for ABTS^{•-} (420 nm, $\epsilon = 3.6 \times 10^4 \text{ M}^{-1}\text{cm}^{-1}$; 550 nm, $\epsilon = 5.9 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$) (27).

Total Turnover Measurements. Enzyme (500 nM in 10 mL 20 mM Tris buffer at pH 8 with 1 mM CuSO₄) was placed in a dialysis cassette with a 10 kD molecular weight cutoff membrane. A sponge was attached to the top of the dialysis cassette to let it float in a beaker containing 750 mL of 1 mM ABTS²⁻ and 1 mM CuSO₄ in 25 mM NaOAc buffer, pH 5.3. Each beaker with the cassette containing wildtype (WT) or mutant enzyme was stirred at room temperature (20 °C). ABTS²⁻/CuSO₄ solutions were exchanged multiple rounds with fresh ABTS²⁻ until there was no further absorption increase at 420 nm, indicating that enzyme catalysis had terminated. Reactions proceeded for one day on the first round and two days for each of the next two rounds of solution exchanges. The amount of ABTS^{•-} produced in each round was estimated from the 420 nm absorbance of the dialysate. The stability of ABTS^{•-} radical as well as the degree of ABTS²⁻ autooxidation was tested with a partially oxidized ABTS sample solution having the 420 nm absorbance of ABTS^{•-} radical at around 1. The degree of ABTS²⁻ autooxidation was negligible and the ABTS^{•-} radical was stable over time (Figure S4.7), giving the same absorbance measurement after 2 days (Table S4.1). Since one laccase catalytic cycle

requires four electrons, total turnovers were defined as one fourth of the ratio of total moles ABTS^{•-} produced to the moles of enzyme in the cassette. Deactivated enzymes were isolated by syringe from the dialysis cassette, washed with MES buffer (pH 5.3) in multiple spin filtration cycles, and then used for LC-MS and peptide analysis.

LC-MS and Peptide Analysis. Trypsin digestion of samples from total turnover experiments was performed according to standard protocols (28). HPLC was run for each sample to remove salts from the digested peptides, and purified peptide samples were loaded on a QE HF mass spectrometer for analysis. To account for oxidation artifacts that frequently arise in LC-MS analyses of proteolysis fragments (29, 30), fresh enzyme control samples were analyzed to distinguish adventitious oxidation from that occurring during enzyme turnover. LC-MS analyses of inactivated enzymes and peptide segments were performed in the Caltech Mass Spectrometry Facility and the Beckman Institute Proteome Exploration Laboratory.

4.6. Acknowledgments

We thank Mona Shahgholi for LC-MS data, Brett Lomenick for assistance with the QE HF MS, and BL, Jeff Jones and Tsui-Fen Chou for their help with peptide analysis.

4.7. Funding Sources

This work was supported by the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health under award number R01DK019038. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

4.8. References

1. Miyazaki, K. A hyperthermophilic laccase from *Thermus thermophilus* HB27. *Extremophiles* **2005**, 9(6):415-425. [Doi: 10.1007/s00792-005-0458-z](https://doi.org/10.1007/s00792-005-0458-z)
2. Nakamura, K.; Go, N. Function and molecular evolution of multicopper blue proteins. *Cell. Mol. Life Sci.* **2005**, 62, 2050–2066. [Doi: 10.1007/s00018-004-5076-x](https://doi.org/10.1007/s00018-004-5076-x)
3. Serrano-Posada, H.; Centeno-Leija, S.; Rojas-Trejo, S.P.; Rodriguez-Almazan, C.; Stojanoff, V.; Rudino-Pinera, E. X-ray-induced catalytic active-site reduction of a multicopper oxidase: structural insights into the proton-relay mechanism and O₂-reduction states. *Acta Crystallogr D Biol Crystallogr.* **2015**, 71, 2396-2411. [Doi: 10.1107/s1399004715018714](https://doi.org/10.1107/s1399004715018714)
4. Petersen, L.C.; Degn, H. Steady-state kinetics of laccase from *Rhus vernicifera*. *Biochim. Biophys Acta.* **1978**, 526, 85-92. [Doi: 10.1016/0005-2744\(78\)90292-9](https://doi.org/10.1016/0005-2744(78)90292-9)
5. Solomon, E.I.; Sundaram, U.M.; Machonkin, T.E. Multicopper Oxidases and Oxygenases. *Chem. Rev.* **1996**, 96, 7, 2563-2605. [Doi: 10.1021/cr950046o](https://doi.org/10.1021/cr950046o)
6. Galli, I.; Musci, G.; Bonaccorsi di Patti, M.C. Sequential reconstitution of copper sites in the multicopper oxidase CueO. *J Biol Inorg Chem.* **2004**, 9(1):90-5. [Doi: 10.1007/s00775-003-0501-4](https://doi.org/10.1007/s00775-003-0501-4)
7. Hijazi, N.H.; Laidler, K.J. Transient-Phase and Steady-State Kinetics for Enzyme Systems Involving Two Substrates. *Can. J. Biochem.* **1972**, 51, 832-840. [Doi: 10.1139/o73-103](https://doi.org/10.1139/o73-103)
8. Gray, H. B.; Winkler, J. R. Hole hopping through tyrosine/ tryptophan chains protects proteins from oxidative damage. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, 112, 10920–10925. [Doi: 10.1073/pnas.1512704112](https://doi.org/10.1073/pnas.1512704112)

9. Ener, M. E.; Gray, H. B.; Winkler, J. R. Hole hopping through tryptophan in cytochrome P450. *Biochemistry* **2017**, *56*, 3531–3538. Doi: 10.1021/acs.biochem.7b00432
10. Gray, H.B.; Winkler, J.R. Living with Oxygen. *Acc. Chem. Res.* **2018**, *51*, 8, 1850-1857. Doi: 10.1021/acs.accounts.8b00245
11. Sorensen, M.L.H; Sanders, B.C.; Hicks, L.P.; Rasmussen, M.H.; Vishart, A.L.; Kongsted, J.; Winkler, J.R.; Gray, H.B.; Hansen, T. Hole Hopping through Cytochrome P450. *J. Phys. Chem. B.* **2020**, *124*, 3065-3073. Doi: 10.1021/acs.jpcc.9b09414
12. Gupta, A.; Nederlof, I.; Sottini, S.; Tepper, A.W.J.W.; Groenen, E.J.J.; Thomassen, E.A.J.; Canters, G.W. Involvement of Tyr108 in the Enzyme Mechanism of the Small Laccase from *Streptomyces coelicolor*. *J. Am. Chem. Soc.* **2012**, *134*, 44, 18213–18216. Doi: 10.1021/ja3088604
13. Tian, S.L.; Jones, S.M.; Solomon, E.I. Role of Tyrosine Radical in Human Ceruloplasmin Catalysis. *ACS Central Science* **2020**, *6*, 1835-1843. Doi: 10.1021/acscentsci.0c00953
14. Winkler, J.R.; Gray, H.B. Could tyrosine and tryptophan serve multiple roles in biological redox processes? *Phil. Trans. R. Soc. A.* **2015**, *373*: 20140178. Doi: 10.1098/rsta.2014.0178
15. Moosmann, B.; Schindeldecker, M.; Hajieva, P. Cysteine, glutathione and a new genetic code: biochemical adaptations of the primordial cells that spread into open water and survived biospheric oxygenation. *Biol. Chem.* **2020**, *401*, 213-231. Doi: 10.1515/hsz-2019-0232
16. Moosmann, B. Redox Biochemistry of the Genetic Code. *Trends Biochem. Sci.* **2021**, *46*, 83-86. Doi: 10.1016/j.tibs.2020.10.008

17. Yee, E.; Dzikovski, B.; Crane, B.R. Tuning Radical Relay Residues by Proton Management Rescues Protein Electron Hopping. *J. Am. Chem. Soc.* **2019**, *141*, 17571-17587. [Doi: 10.1021/jacs.9b05715](https://doi.org/10.1021/jacs.9b05715)
18. Weber, D.S.; Warren, J.J. The interaction between methionine and two aromatic amino acids is an abundant and multifunctional motif in proteins. *Arch. Biochem. Biophys.* **2019**, *672*, 108053. [Doi: 10.1016/j.abb.2019.07.018](https://doi.org/10.1016/j.abb.2019.07.018)
19. Teo, R.D.; Wang, R.B.; Smithwick, E.R.; Migliore, A.; Therien, M.J.; Beratan, D.N. Mapping hole hopping escape routes in proteins. *Proc. Natl. Acad. Sci. USA.* **2019**, *116*, 15811-15816. [Doi: 10.1073/pnas.1906394116](https://doi.org/10.1073/pnas.1906394116)
20. Granold, M.; Hajieva, P.; Ioana, M.; Florin-Dan, I.; Moosmann, B. Modern diversification of the amino acid repertoire driven by oxygen. *Proc. Natl. Acad. Sci. USA.* **2018**, *115*, 41-46. [Doi: 10.1073/pnas.1717100115](https://doi.org/10.1073/pnas.1717100115)
21. Meena, K.; English, A.M. LC-MS/MS Proteoform Profiling Exposes Cytochrome c Peroxidase Self-Oxidation in Mitochondria and Functionally Important Hole Hopping from Its Heme. *J. Am. Chem. Soc.* **2018**, *140*, 12033-12039. [Doi: 10.1021/jacs.8b05966](https://doi.org/10.1021/jacs.8b05966)
22. Njuma, O.J.; Davis, I.; Ndontsa, E.N.; Krewall, J.R.; Liu, A.M.; Goodwin, D.C. Mutual synergy between catalase and peroxidase activities of the bifunctional enzyme KatG is facilitated by electron hole-hopping within the enzyme. *J. Biol. Chem.* **2017**, *292*, 18408-18421. [Doi: 10.1074/jbc.M117.791202](https://doi.org/10.1074/jbc.M117.791202)
23. Meena, K.; English, A.M. LC-MS/MS suggests that hole hopping in cytochrome c peroxidase protects its heme from oxidative modification by excess H₂O₂. *Chem. Sci.* **2017**, *8*, 1152-1162. [Doi: 10.1039/C6SC03125K](https://doi.org/10.1039/C6SC03125K)
24. Shin, J.; Gray, H.B.; Winkler, J.R. Stability/activity tradeoffs in *Thermus thermophilus* HB27 laccase. *J Biol Inorg Chem.* **2020**, *25*(2):233-238. [Doi: 10.1007/s00775-020-01754-7](https://doi.org/10.1007/s00775-020-01754-7)

25. Zheng, L.; Baumann, U.; Reymond, J.L. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucl. Acids Res.* **2004**, 32:e115. [Doi: 10.1093/nar/gnh110](https://doi.org/10.1093/nar/gnh110)
26. Liu, X.; Gillespie, M.; Ozel, A.D.; Dikici, E.; Daunert, S.; Bachas, L.G. Electrochemical properties and temperature dependence of a recombinant laccase from *Thermus thermophilus*. *Anal Bioanal Chem.* **2011**, 399(1): 361-366. [Doi: 10.1007/s00216-010-4345-9](https://doi.org/10.1007/s00216-010-4345-9)
27. Collins, P.J.; Dobson, A.D.W.; Field, J.A. Reduction of the 2,2,'-Azinobis(3-Ethylbenzthiazoline-6-Sulfonate) Cation Radical by Physiological Organic Acids in the Absence and Presence of Manganese. *Appl Environ Microbiol.* **1998**, 64(6):2026-2031. [Doi: 10.1128/AEM.64.6.2026-2031.1998](https://doi.org/10.1128/AEM.64.6.2026-2031.1998)
28. León, I.R.; Schwämmle, V.; Jensen, O.N.; Sprenger, R.R. Quantitative assessment of in-solution digestion efficiency identifies optimal protocols for unbiased protein analysis. *Mol Cell Proteomics.* **2013**. 12(10):2992-3005. [Doi: 10.1074/mcp.M112.025585](https://doi.org/10.1074/mcp.M112.025585)
29. Bettinger, J.Q.; Welle, K.A.; Hryhorenko, J.R.; Ghaemmaghami, S. Quantitative Analysis of in Vivo Methionine Oxidation of the Human Proteome. *J Proteome Res.* **2020**, 19(2):624-633. [Doi: 10.1021/acs.jproteome.9b00505](https://doi.org/10.1021/acs.jproteome.9b00505)
30. Mouchahoir, T.; Schiel, J.E. Development of an LC-MS/MS peptide mapping protocol for the NISTmAb. *Anal Bioanal Chem.* **2018**, 410, 2111–2126. [Doi: 10.1007/s00216-018-0848-6](https://doi.org/10.1007/s00216-018-0848-6)

Supporting Information

**Trp-Tyr pair protects *Thermus thermophilus* HB27 laccase
from oxidative damage**

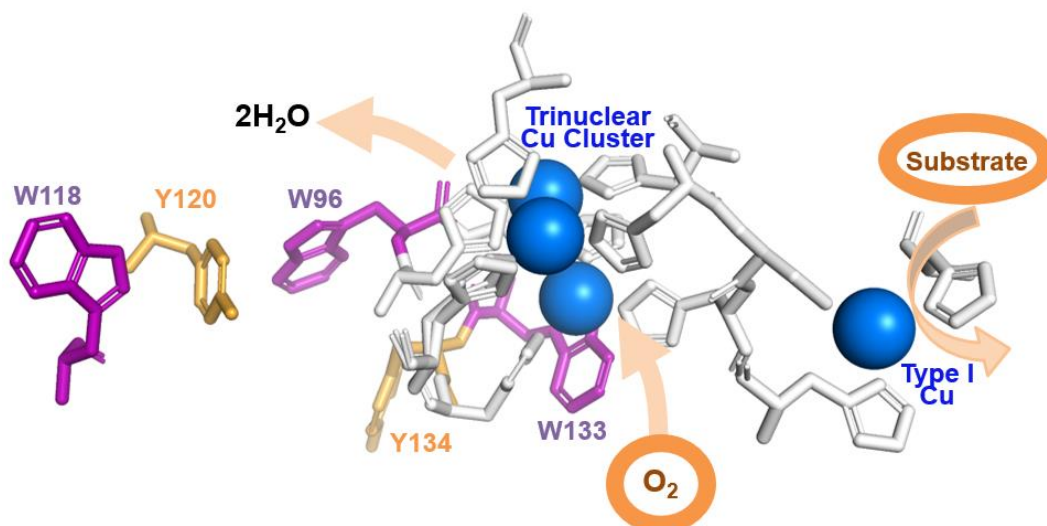


Figure S4.1. Graphic showing the active site coppers (blue) with coordinating residues (grey) and Trp/ Tyr residues near the trinuclear copper cluster.

Mass-spec before and after the Enzyme Deactivation

	WT	W133F
Expected Mass (D)	49585	49558
After Total Turnover		
Observed Mass (D)	49617	49574
	Mass Shift After Enzyme Turnover	
Discrepancy (D)	+ 32	+ 16

Figure S4.2. LC-MS results before and after enzyme total turnovers.

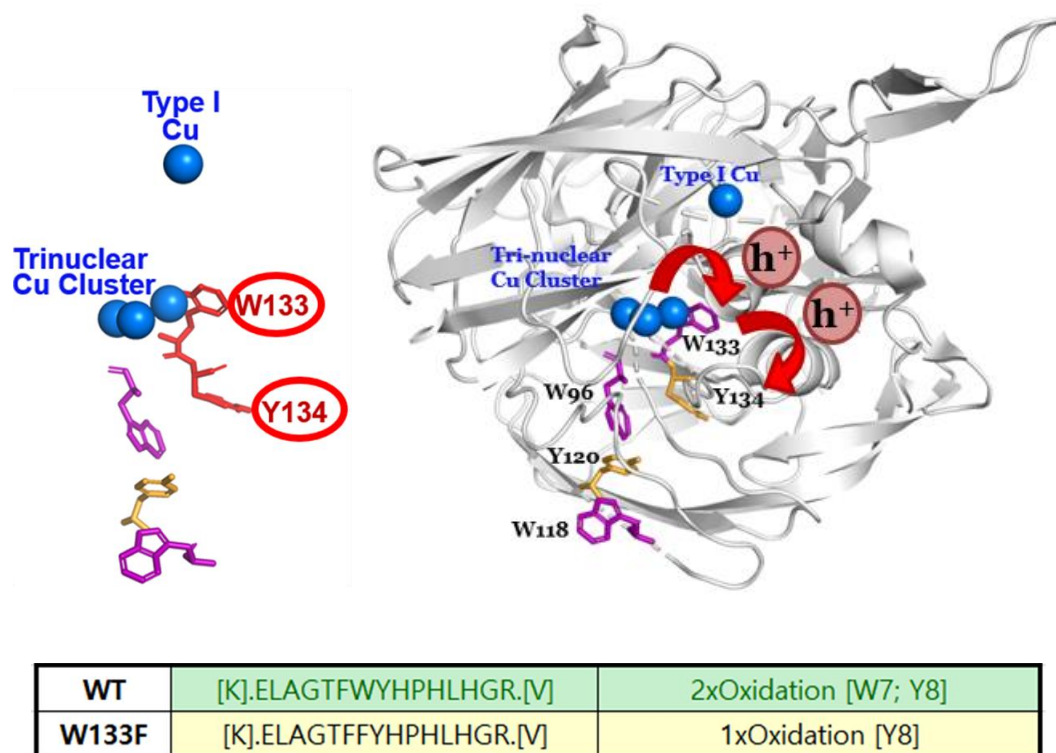


Figure S4.3. Graphic showing a Trp/Tyr (W133/Y134) pair near the Tth-lac TNC. Peptide analysis indicates that the W133/Y134 pair is oxidized during WT catalytic turnover, and Y134 is oxidized during W133F turnover.

Sequence: ELAGTFWYHPHLHGR, W7-Oxidation (15.99492 Da), Y8-Oxidation (15.99492 Da)

Charge: +3

Proteins (1): Tth-lac_WT

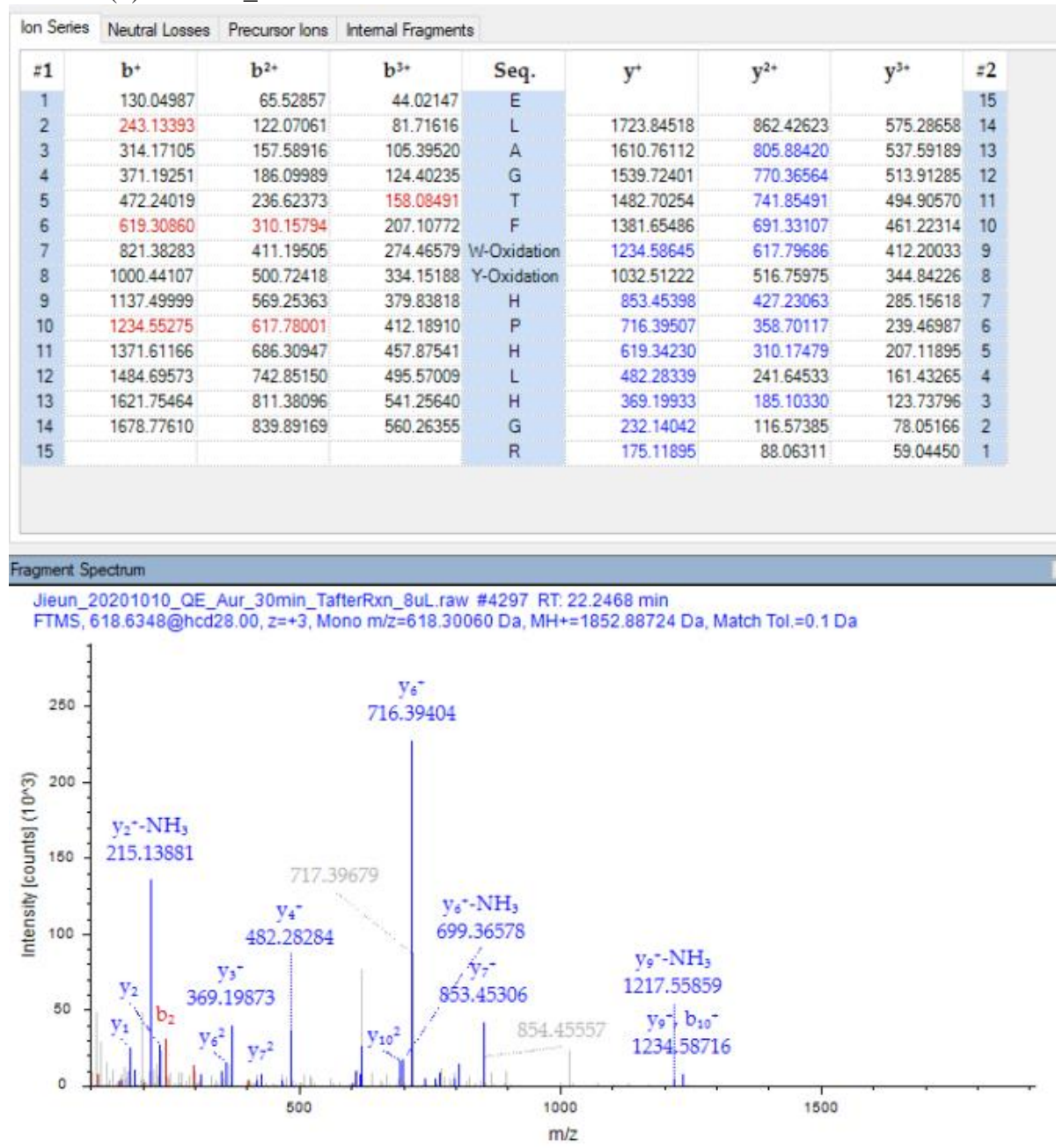


Figure S4.4. MS/MS Data for the Oxidized Peptide Fragments in *Tth*-lac WT.

Sequence: ELAGTFFYHPHLHGR, Y8-Oxidation (15.99492 Da)
Charge: +3

Proteins (2): Tth-lac_W133F

Ion Series	Neutral Losses	Precursor Ions	Internal Fragments					
#1	b ⁺	b ²⁺	b ³⁺	Seq.	y ⁺	y ²⁺	y ³⁺	#2
1	130.04987	65.52857	44.02147	E				15
2	243.13393	122.07061	81.71616	L	1668.83937	834.92332	556.95131	14
3	314.17105	157.58916	105.39520	A	1555.75531	778.38129	519.25662	13
4	371.19251	186.09989	124.40235	G	1484.71819	742.86273	495.57758	12
5	472.24019	236.62373	158.08491	T	1427.69673	714.35200	476.57043	11
6	619.30860	310.15794	207.10772	F	1326.64905	663.82816	442.88787	10
7	766.37702	383.69215	256.13052	F	1179.58064	590.29396	393.86506	9
8	945.43526	473.22127	315.81660	Y-Oxidation	1032.51222	516.75975	344.84226	8
9	1082.49417	541.75072	361.50291	H	853.45398	427.23063	285.15618	7
10	1179.54694	590.27711	393.85383	P	716.39507	358.70117	239.46987	6
11	1316.60585	658.80656	439.54013	H	619.34230	310.17479	207.11895	5
12	1429.68991	715.34859	477.23482	L	482.28339	241.64533	161.43265	4
13	1566.74882	783.87805	522.92113	H	369.19933	185.10330	123.73796	3
14	1623.77029	812.38878	541.92828	G	232.14042	116.57385	78.05166	2
15				R	175.11895	88.06311	59.04450	1

Fragment Spectrum

Jieun_20201010_QE_Aur_30min_TafterRxn_BuL.raw #5791 RT: 27.0919 min
FTMS, 599.9655@hcd28.00, z=+3, Mono m/z=599.96552 Da, MH+=1797.88199 Da, Match Tol.=0.1 Da

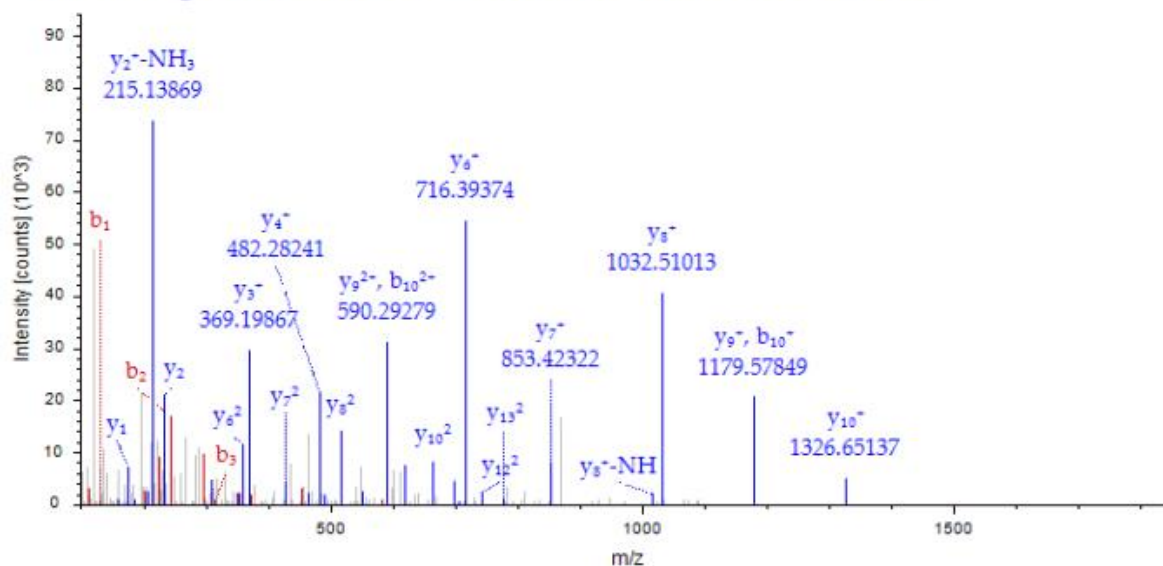


Figure S4.5. MS/MS Data for the Oxidized Peptide Fragments in *Tth*-lac W133F Mutant.

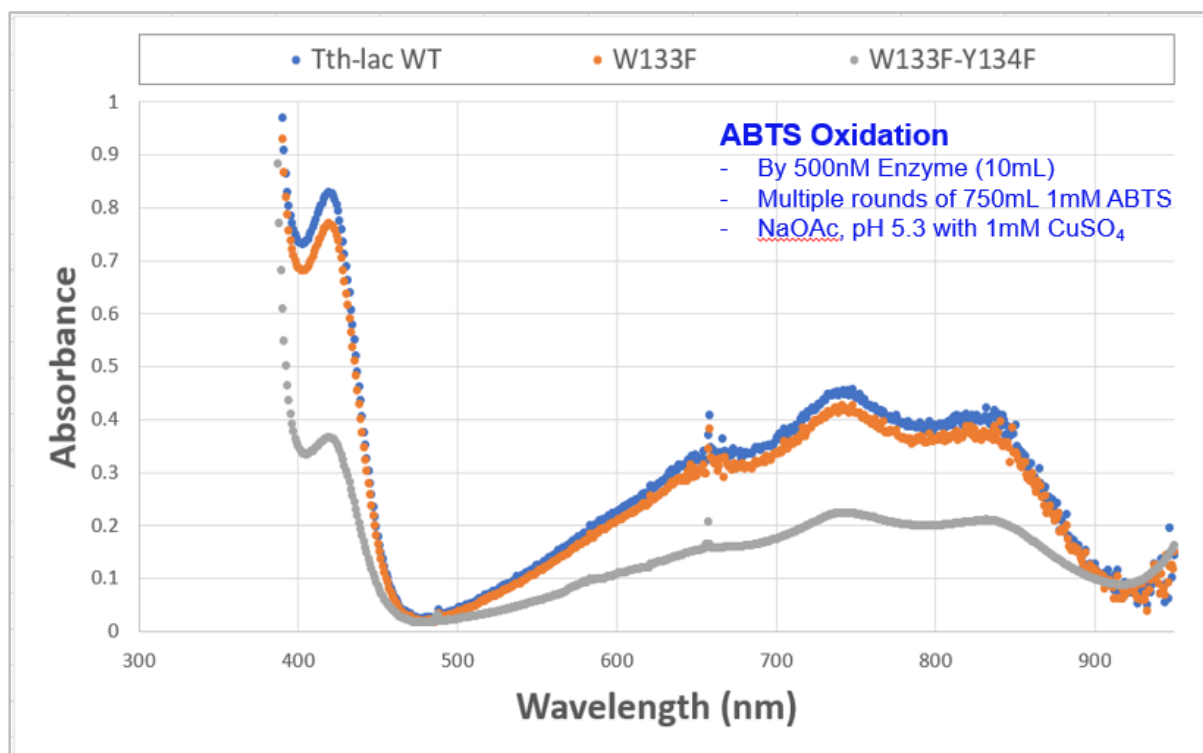


Figure S4.6. UV-vis spectra showing the relative amount of ABTS^{•-} formed after enzyme total turnovers (solutions from all three rounds were diluted to give the 420 nm absorbances shown).

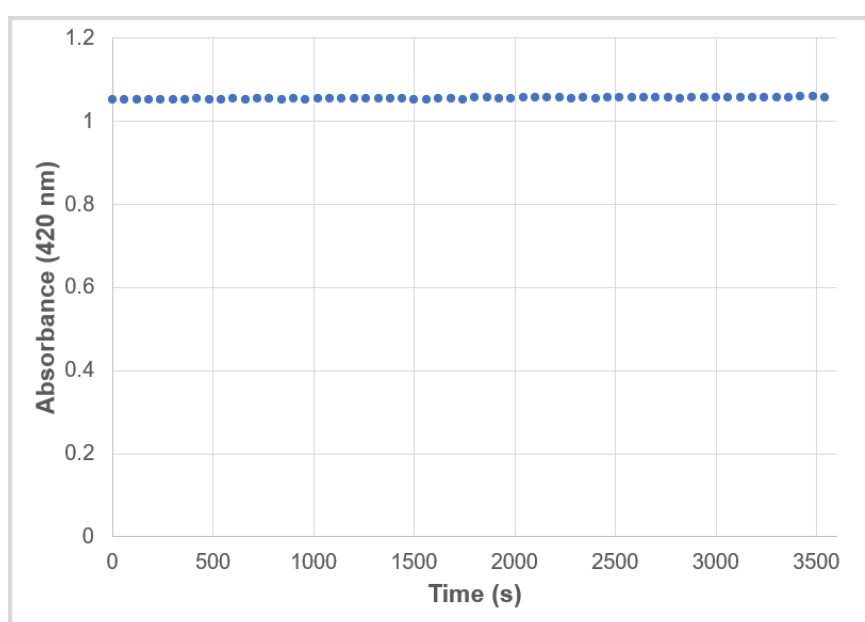


Figure S4.7. Absorbance of ABTS^{•-} at 420 nm monitored for 1 h at 20 °C.

Table S4.1. Absorbance of ABTS^{•-} at 420 nm in partially oxidized [1 mM]_{initial} ABTS solution containing 1 mM CuSO₄ in 25 mM NaOAc buffer, pH 5.3 (same as the total turnover experiment).

	Initial	After 1 day	After 2 days
Absorbance (420 nm)	1.03	1.04	1.05

Activity assays were done with three different batches of proteins expressed at different times (and five different sets purified separately). A reasonable degree of reproducibility was attained by expressing and purifying all mutants as one set at the same time and by storing the enzyme in Tris, pH 8 with 1mM CuSO₄.

[Set 1]

Measurements were made in 25 mM sodium acetate buffer, pH 4.5.

Table S4.2. Michaelis-Menton kinetics in NaOAc buffer, pH 4.5.

	[E], nM	[ABTS], mM									
	100	75	50	30	10	5	3	1	0.5	0.3	0.1
Initial Rate, M s⁻¹	WT	8.5x10 ⁻⁸	6.5x10 ⁻⁸	5.6x10 ⁻⁸	3.7x10 ⁻⁸	3.0x10 ⁻⁸	2.6x10 ⁻⁸	1.9x10 ⁻⁸	1.6x10 ⁻⁸	4.4x10 ⁻⁹	1.6x10 ⁻⁹
	W96F	7.8x10 ⁻⁸	7.6x10 ⁻⁸	6.3x10 ⁻⁸	3.6x10 ⁻⁸	3.2x10 ⁻⁸	2.9x10 ⁻⁸	2.0x10 ⁻⁸	1.6x10 ⁻⁸	4.9x10 ⁻⁹	1.4x10 ⁻⁹
	W118F	5.3x10 ⁻⁸	5.3x10 ⁻⁸	4.4x10 ⁻⁸	2.8x10 ⁻⁸	2.3x10 ⁻⁸	2.2x10 ⁻⁸	1.6x10 ⁻⁸	1.3x10 ⁻⁸	3.6x10 ⁻⁹	1.2x10 ⁻⁹
	W133F	6.2x10 ⁻⁸	5.6x10 ⁻⁸	5.0x10 ⁻⁸	3.4x10 ⁻⁸	2.8x10 ⁻⁸	2.3x10 ⁻⁸	1.6x10 ⁻⁸	1.3x10 ⁻⁸	4.3x10 ⁻⁹	1.5x10 ⁻⁹
	Y134F	6.8x10 ⁻⁸	5.7x10 ⁻⁸	4.9x10 ⁻⁸	3.5x10 ⁻⁸	2.8x10 ⁻⁸	2.5x10 ⁻⁸	1.7x10 ⁻⁸	1.5x10 ⁻⁸	4.8x10 ⁻⁹	1.5x10 ⁻⁹
	W133F/Y134F	2.6x10 ⁻⁸	2.0x10 ⁻⁸	1.5x10 ⁻⁸	1.2x10 ⁻⁸	9.5x10 ⁻⁹	8.0x10 ⁻⁹	6.1x10 ⁻⁹	4.6x10 ⁻⁹	1.6x10 ⁻⁹	8.9x10 ⁻¹⁰

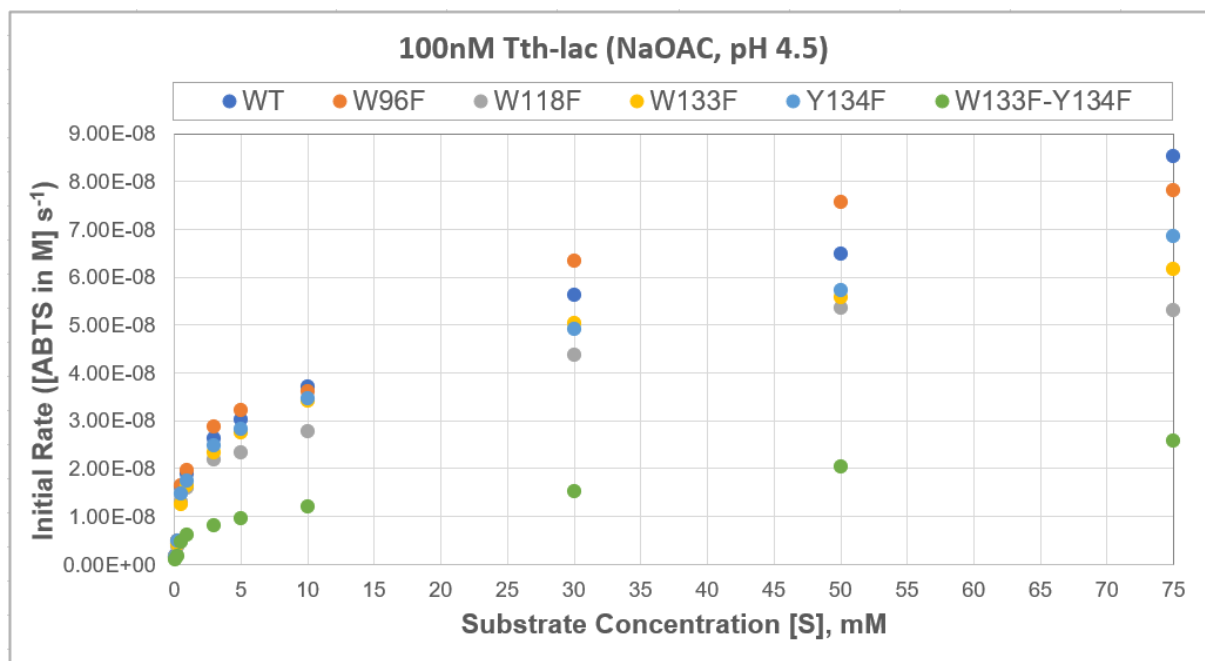


Figure S4.8. Enzyme activity assays (substrate ABTS at pH 4.5) for Trp/Tyr chain mutants.

Reproducibility with Different Batches of Proteins

Each batch was expressed at different times as a set with Cu_{T1}-depleted (T1D) and T1D-W118F mutants as exceptions (T1D and T1D-W118F mutants were expressed separately twice.)

[Set 2, 3, 4]

The measurements were made in 25 mM sodium acetate buffer, pH 4.5.

Table S4.3. Activity assay with different batches of proteins to check reproducibility.

		Batch 1	Batch 2	Batch 3	Batch 1	Batch 2	Batch 3
	[E], nM	[ABTS], mM					
	100	3	3	3	1	1	1
Initial Rate, M s ⁻¹	WT	3.0 x10 ⁻⁸	2.8 x10 ⁻⁸	3.1 x10 ⁻⁸	1.9 x10 ⁻⁸	1.8 x10 ⁻⁸	1.9 x10 ⁻⁸
	W96F	3.1 x10 ⁻⁸	2.9 x10 ⁻⁸	2.8 x10 ⁻⁸	2.0 x10 ⁻⁸	2.0 x10 ⁻⁸	1.9 x10 ⁻⁸
	W118F	2.2 x10 ⁻⁸	2.2 x10 ⁻⁸	2.2 x10 ⁻⁸	1.4 x10 ⁻⁸	1.3 x10 ⁻⁸	1.5 x10 ⁻⁸
	W133F	2.6 x10 ⁻⁸	2.5 x10 ⁻⁸	2.6 x10 ⁻⁸	1.7 x10 ⁻⁸	1.6 x10 ⁻⁸	1.7 x10 ⁻⁸
	Y134F	2.6 x10 ⁻⁸	2.3 x10 ⁻⁸	2.6 x10 ⁻⁸	1.6 x10 ⁻⁸	1.6 x10 ⁻⁸	1.7 x10 ⁻⁸
	W133F/ Y134F	8.4 x10 ⁻⁹	7.6 x10 ⁻⁹	9.2 x10 ⁻⁹	5.8 x10 ⁻⁹	5.0 x10 ⁻⁹	7.5 x10 ⁻⁹

Protein Concentration Dependence

The protein concentration was doubled for all mutants to test the protein concentration dependence of enzyme activity towards ABTS. The reaction kinetics appear to be the first order with respect to enzyme concentration.

Table S4.4. Activity assay results with a doubled enzyme concentration.

		Batch 1	
	[E], nM	[ABTS], mM	
	200	3	1
Initial Rate, M s ⁻¹	WT	6.1 x10 ⁻⁸	4.1 x10 ⁻⁸
	W96F	6.1 x10 ⁻⁸	4.2 x10 ⁻⁸
	W118F	4.4 x10 ⁻⁸	3.3 x10 ⁻⁸
	W133F	4.7 x10 ⁻⁸	3.3 x10 ⁻⁸
	Y134F	5.3 x10 ⁻⁸	4.1 x10 ⁻⁸
	W133F/Y134F	1.7 x10 ⁻⁸	1.2 x10 ⁻⁸

Table S4.5. Uncertainties of catalytic efficiencies: WT and the Trp/Tyr mutants in NaOAc, pH 4.5 (the lower bound and the upper bound estimated for each parameter, and the larger of the two was assigned as an uncertainty).

	WT	W96F	W118F	W133F	Y134F	W133F/Y134F
k_{cat} (s⁻¹)	8.0×10^{-1} $\pm 0.8 \times 10^{-1}$	8.3×10^{-1} $\pm 0.7 \times 10^{-1}$	5.5×10^{-1} $\pm 0.1 \times 10^{-1}$	6.1×10^{-1} $\pm 0.5 \times 10^{-1}$	6.5×10^{-1} $\pm 0.6 \times 10^{-1}$	2.4×10^{-1} $\pm 0.2 \times 10^{-1}$
K_m (M)	7.9×10^{-3} $\pm 0.2 \times 10^{-3}$	7.5×10^{-3} $\pm 0.2 \times 10^{-3}$	5.7×10^{-3} $\pm 0.3 \times 10^{-3}$	5.0×10^{-3} $\pm 0.2 \times 10^{-3}$	5.7×10^{-3} $\pm 0.2 \times 10^{-3}$	7.2×10^{-3} $\pm 0.5 \times 10^{-3}$
k_{cat}/K_m M⁻¹s⁻¹	9.8×10^1 $\pm 0.3 \times 10^1$	1.1×10^2 $\pm 0.1 \times 10^2$	9.6×10^1 $\pm 0.4 \times 10^1$	1.2×10^2 $\pm 0.1 \times 10^1$	1.1×10^1 $\pm 0.1 \times 10^1$	3.3×10^1 $\pm 0.6 \times 10^1$
Upper Bound						
	WT	W96F	W118F	W133F	Y134F	W133F/Y134F
k_{cat} (s⁻¹)	8.8×10^{-1}	9.0×10^{-1}	5.5×10^{-1}	6.3×10^{-1}	6.9×10^{-1}	2.6×10^{-1}
K_m (M)	8.1×10^{-3}	7.7×10^{-3}	5.4×10^{-3}	4.9×10^{-3}	5.8×10^{-3}	6.7×10^{-3}
k_{cat}/K_m M⁻¹s⁻¹	1.1×10^2	1.2×10^2	1.0×10^2	1.3×10^2	1.2×10^1	3.9×10^1
Lower Bound						
	WT	W96F	W118F	W133F	Y134F	W133F/Y134F
k_{cat} (s⁻¹)	7.5×10^{-1}	8.0×10^{-1}	5.6×10^{-1}	5.6×10^{-1}	5.9×10^{-1}	2.2×10^{-1}
K_m (M)	7.8×10^{-3}	7.5×10^{-3}	6.0×10^{-3}	4.8×10^{-3}	5.5×10^{-3}	7.6×10^{-3}
k_{cat}/K_m M⁻¹s⁻¹	9.6×10^1	1.1×10^2	9.3×10^1	1.2×10^2	1.1×10^1	2.9×10^1

The upper bound and lower bound values were determined from upper limit and lower limit of each data point (shown as error bars in Figure 4) by least squares fits to the Michaelis-Menten model.

[Set 5]

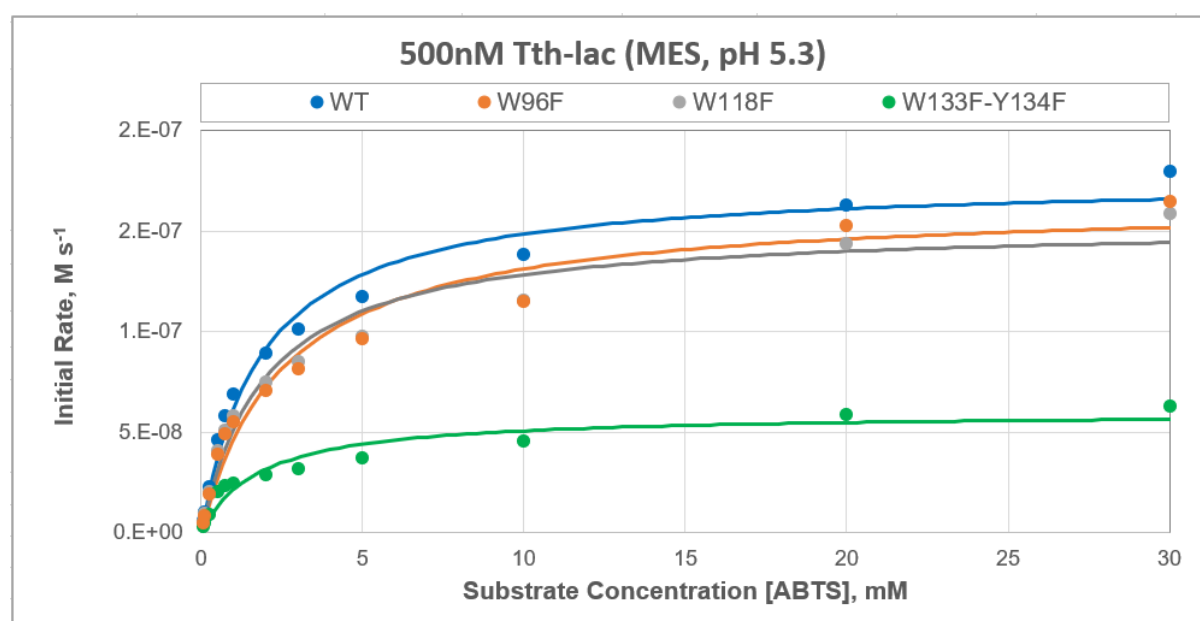
Buffer Dependence

Activity assays were also performed in MES non-coordinating buffer (only initial rates were monitored). Although the optimal pH for an ABTS assay is pH 4.5, pH 5.3 was used with MES buffer (the low end of the MES buffering range), and the enzyme concentration in the assay sample was increased to 500 nM.

Table S4.6. Michaelis-Menton Kinetics in MES buffer, pH 5.3.

	[E], nM	[ABTS], mM						
	500	75	50	30	20	10	5	3
Initial Rate, M s⁻¹	WT	5.5×10^{-7}	3.6×10^{-7}	1.8×10^{-7}	1.6×10^{-7}	1.4×10^{-7}	1.2×10^{-7}	1.0×10^{-7}
	W118F	3.9×10^{-7}	3.2×10^{-7}	1.6×10^{-7}	1.4×10^{-7}	1.2×10^{-7}	9.8×10^{-8}	8.5×10^{-8}
	W133F/Y134F	1.6×10^{-7}	1.3×10^{-7}	6.3×10^{-8}	5.9×10^{-8}	4.5×10^{-8}	3.7×10^{-8}	3.2×10^{-8}
	W96F	5.1×10^{-7}	4.1×10^{-7}	1.6×10^{-7}	1.5×10^{-7}	1.2×10^{-7}	9.6×10^{-8}	8.2×10^{-8}

	[E], nM	[ABTS], mM						
	500	2	1	0.75	0.5	0.25	0.1	0.05
Initial Rate, M s⁻¹	WT	8.9×10^{-8}	6.9×10^{-8}	5.8×10^{-8}	4.6×10^{-8}	2.3×10^{-8}	1.0×10^{-8}	5.9×10^{-9}
	W118F	7.5×10^{-8}	5.9×10^{-8}	5.1×10^{-8}	4.1×10^{-8}	2.1×10^{-8}	9.6×10^{-9}	5.5×10^{-9}
	W133F/Y134F	2.9×10^{-8}	2.4×10^{-8}	2.4×10^{-8}	2.0×10^{-8}	9.3×10^{-9}	4.9×10^{-9}	3.2×10^{-9}
	W96F	7.1×10^{-8}	5.5×10^{-8}	4.9×10^{-8}	3.9×10^{-8}	2.0×10^{-8}	8.7×10^{-9}	5.1×10^{-9}

**Figure S4.9.** Enzyme activity assays (substrate ABTS at pH 5.3) for WT and Trp/Tyr chain mutants. Curved lines are least squares fits to a Michaelis-Menten model.

As can be seen in Figure S9, the same trend in activity profile is observed as in Figure S4, which indicates that the observed decrease in the catalytic efficiency of the Trp/Tyr double mutant is not due to the assay conditions.

Table S4.7. Catalytic efficiencies: WT and the Trp/Tyr mutants in MES, pH 5.3.

	WT	W96F	W118F	W133F/Y134F
k_{cat} (s^{-1})	3.5×10^{-1}	3.3×10^{-1}	3.1×10^{-1}	1.2×10^{-1}
K_m (M)	1.9×10^{-3}	2.6×10^{-3}	2.0×10^{-3}	1.8×10^{-3}
Catalytic Efficiency (k_{cat}/K_m), $M^{-1}s^{-1}$	184	126	155	67

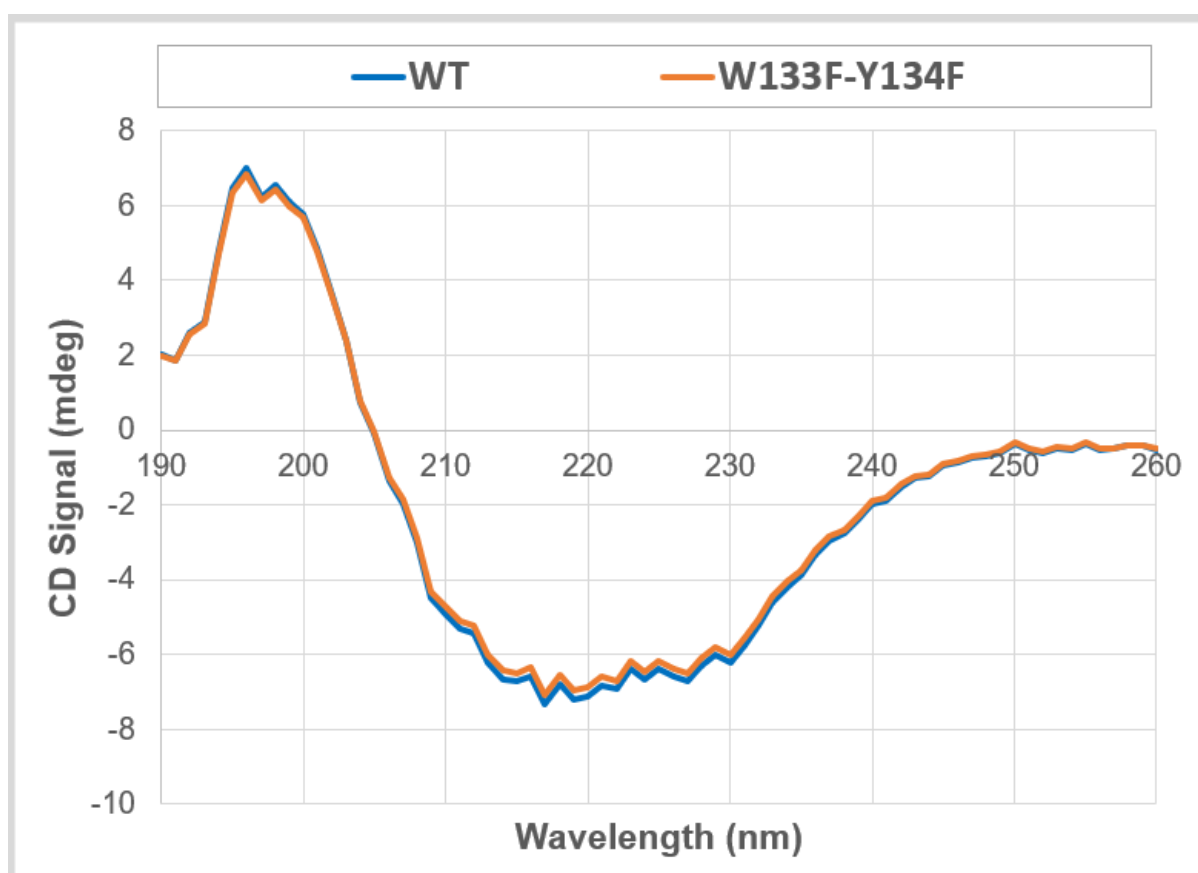


Figure S4.10. Circular dichroism spectra of 5 μ M *Tth*-lac WT and W133F-Y134F in sodium phosphate buffer, pH 6.5.

Chapter 5

Investigating the Factors Affecting Laccase Catalysis

- Active-site Potentials & Structural Motifs -

5.1. Potential Discrepancy

According to the consensus mechanism for MCOs, the type 1 copper (Cu_{T1}) is believed to be the substrate oxidation site [1], [2] (see Chapter 2 and Chapter 4 for more details). The Cu_{T1} potential values of fungal laccases from *Pycnoporus cinnabarinus*, *Neurospora crassa* and *Trametes versicolor* are 750 mV [3], 780 mV [4] and 785 mV [5] respectively. The Cu_{T1} potentials of bacterial laccases tend to be lower than that of fungal laccases, which mostly range between 0.4 – 0.5 V [6]–[9] with 375 mV for a small laccase from *Streptomyces sviveus* [6], 480 mV for a thermophilic laccase from *Thermus thermophilus* HB27 (*Tth*-lac) [7] and 455 mV for CotA laccase from *Bacillus Subtilis* (CotA-lac) [8].

Laccases are capable of aerobic oxidation of lignin as their primary function [10], but are capable of oxidizing diverse substrates with a wide potential range due to their natures of substrate non-specificity (Table 5.1). It is notable that the potential of Cu_{T1} in bacterial laccases is more than 500 mV lower than that expected for one-electron oxidation of polyphenolic substrates ($E^\circ(\text{PhO}/\text{PhOH}) = 1.1\text{V}$) [11]. The Cu_{T1} potentials of bacterial laccases even fall short of the potential of ABTS (2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic) acid, 680 mV vs. NHE [12], [13]) which is the most commonly used substrate for MCOs (Table 5.1).

Table 5.1. Various laccase substrates with a wide range of potentials.

Substrates	$[\text{Ru}(\text{NH}_3)_5(\text{py})]\text{Cl}_2$	$[\text{Ru}(\text{NH}_3)_4(\text{bpy})]\text{Cl}_2$	ABTS	2,6-DMP	Guaiacol	Promazine
E° (mV), pH 5	300	500	680	690	800	817
λ_{max} (nm)	407	525	420	470	520	514
Ref	[14]	[14]	[13], [15]	[16]	[17]	[18]

Therefore, in order to resolve this potential discrepancy, it is of interest to explore the roles of redox-active amino acid residues and possible formation of high potential intermediates during enzyme catalysis.

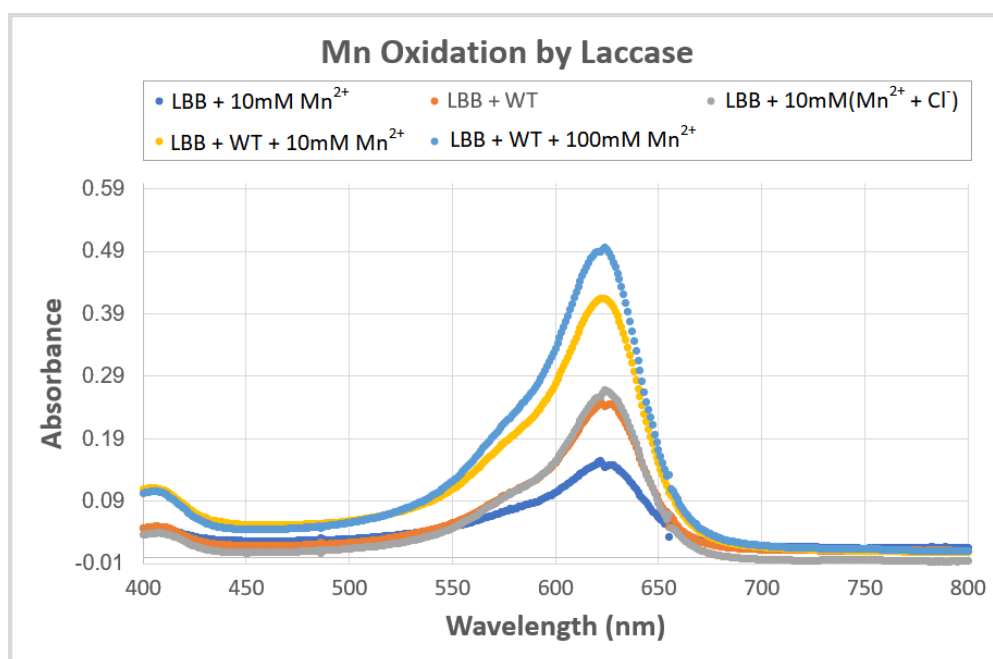


Figure 5.1. Oxidation of leucoberbelin blue (LBB) and manganese by *Tth*-lac, pH 4.5.

Activity assays can be done with the substrates listed in Table 5.1 and many other, and Figure 5.1 illustrates the substrate non-specific nature of *Tth*-lac. Leucoberbelin blue (LBB) is a commonly used dye for detecting high-valent manganese species with the absorption increase at 620 nm upon its oxidation by high-valent manganese formed in aqueous solution [19]. In Figure 5.1, there is a little bit of absorbance at 620 nm even for the sample with only LBB and 10mM Mn²⁺, since the commercial LBB product (from Sigma) comes as slightly oxidized already. When wild-type *Tth*-lac is added to LBB, the absorbance at 620 nm increases, which

is indicative of LBB oxidation. When Mn^{2+} is added to the assay media, the higher degree of absorbance increase is observed for the same duration of time, as *Tth*-lac can oxidize both Mn and LBB, and the oxidized Mn species can further oxidize more LBB. When $MnCl_2$ is used instead of $MnSO_4$, the absorbance does not increase as much (Figure 5.1), since chlorides act as laccase inhibitors [20]. If the reaction is left to proceed for a few days, brown precipitates are formed as manganese oxides get produced with further oxidation.

The observed substrate variability cannot fully be explained by the consensus mechanism which suggests the Cu_{T1} as a substrate oxidation site in MCOs. In *Tth*-lac in particular, the Cu_{T1} is deeply embedded in the protein scaffold (see Chapter 2 for more details), which hinders the binding of bulky substrates. Therefore, low solvent exposure as well as the insufficient Cu_{T1} potential ($E^\circ(Cu^{2+}/Cu^+) = 480$ V vs. NHE), pH 5.5 [7] acts unfavorably for *Tth*-lac to oxidize high potential substrates (Table 5.1).

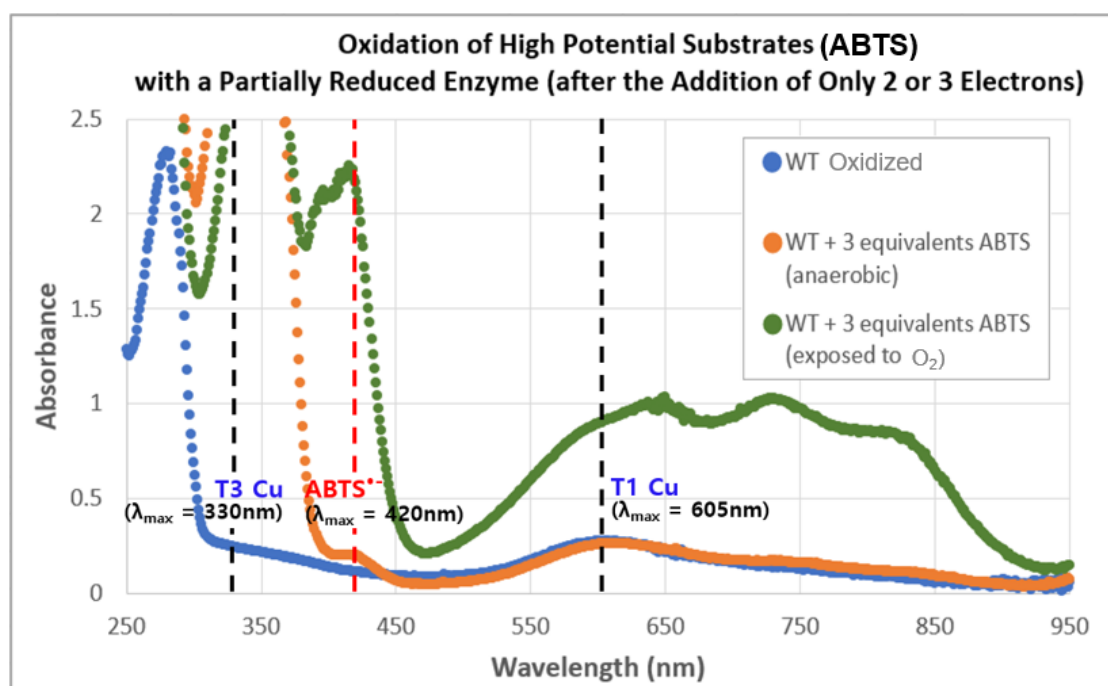


Figure 5.2. Anaerobic titration of three reducing equivalents of ABTS to the wild-type *Tth*-lac.

To test if the TNC itself has high enough potential to oxidize ABTS, redox titrations were done with the known stoichiometric ratios of enzyme to substrate. When two or three reducing equivalents of ABTS were added to a fully oxidized *Tth*-lac WT sample which had been deoxygenated, no ABTS oxidation was observed (Figure 5.2). When the sample was exposed to O₂, the enzyme finally started turning over. From this observation, it is clear that the enzyme is capable of oxidizing high potential substrates like ABTS only in the presence of oxygen. It could also be a likely scenario that the potentials of active-site coppers and/or substrates are altered upon binding of substrates and dioxygen to the enzyme, but the consensus mechanism suggests that the oxygen does not bind at the TNC until all the four coppers are fully reduced. Therefore, these uncertainties associated with the potential gap certainly raise further questions to the consensus mechanism that has been proposed so far.

Since all four active-site coppers in *Tth*-lac do not seem to possess high enough potentials to perform lignin oxidation, the roles of redox-active amino acids such as W and Y are considered. The underlying motivation behind was the study on the evolutionary convergence of lignin degrading enzymes, which suggested that the surface tryptophan in both lignin peroxidase and versatile peroxidase acts as a catalytic site facilitating long range electron transfer reactions [21], [22]. Although these are heme peroxidases and it may not be the case for multicopper oxidases, the involvement of surface Trp (W118) in *Tth*-lac in substrate oxidation was also speculated, as it is located at the end of the Trp/Tyr chain stretching from the TNC (Figure 4.2).

To further investigate the role of surface Trp (W118), it was mutated to Phe which is redox inactive, and the mutant (W118F) activity was tested using high potential substrates (ABTS and Promazine Hydrochloride). We expected the Phe mutant to have a much lower activity, if

the surface Trp was involved in the binding and oxidation of high potential substrates. However, there was no significant difference between the activity of W118F and of WT toward ABTS which has around 200 mV higher potential than the *Tth*-lac Cu_{T1}. (see Chapter 4 for more details). Therefore, we needed a substrate which has even higher potential than ABTS to test this hypothesis. Promazine hydrochloride was a suitable candidate, as it is water soluble and has a potential of about 820 mV versus NHE [18] which is about 340 mV higher than the potential of the Cu_{T1}. Oxidation of Promazine can be monitored by the absorbance increase at 520 nm ($\epsilon = 8900 \text{ M}^{-1}\text{cm}^{-1}$) [18]. When the activity of W118F toward promazine was compared with that of WT, there was again only a minute difference, though slightly lower (Table 5.2, Table 5.3 and Figure 5.3).

Table 5.2. Substrate concentration dependence of W118F activity toward Promazine hydrochloride.

	[E], nM	[Promazine], mM								
	500	300	200	100	50	25	10	5	2	1
Initial Rate, Abs(514nm)/s	WT	scatter	lower	0.0011	0.00117	0.0015	0.00074	0.000394	0.000189	0.000104
	W118F	scatter	lower	0.000998	0.00118	0.00128	0.000601	0.000301	0.000138	0.0000759

Table 5.3. Protein concentration dependence of W118F activity toward Promazine hydrochloride.

	[Promazine], mM	[E], nM	
		250	500
Initial Rate, Abs(514nm)/s	WT	0.000343	0.00074
	W118F	0.000303	0.000601

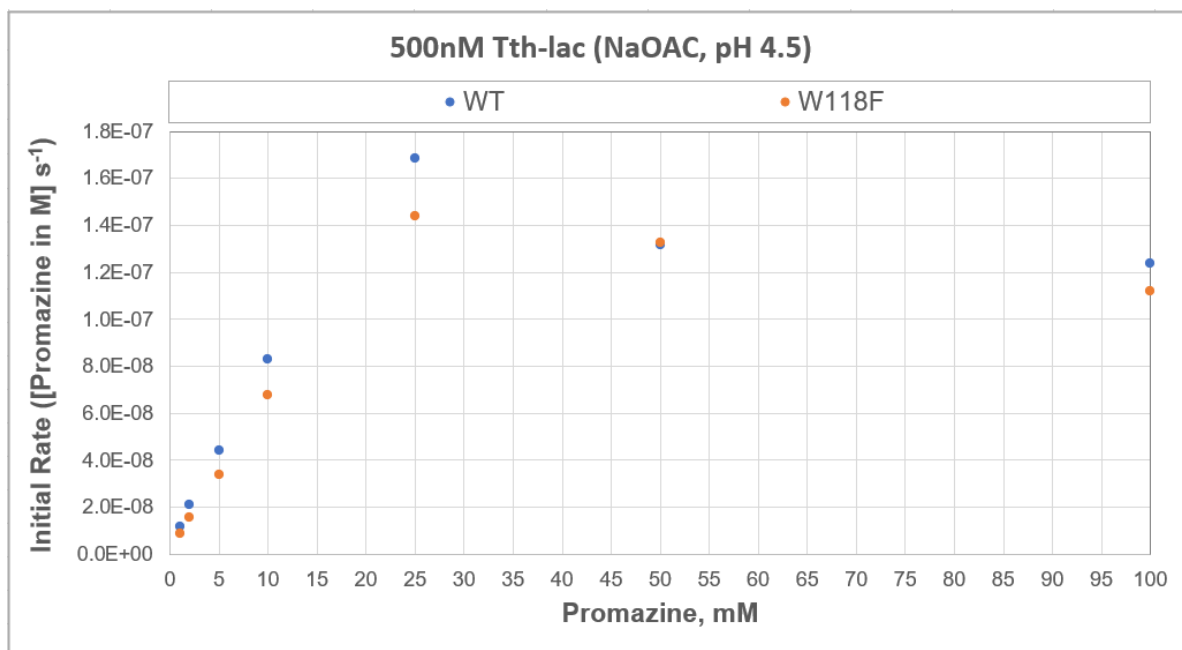


Figure 5.3. Activity assay of *Tth*-lac WT and the W118F mutant with promazine hydrochloride as a substrate, pH 4.5.

To figure out whether this slightly lower activity (Figure 5.3) is of any significance, another possible experiment to try was making a Cu_{T1} depleted enzyme (T1D) and testing its activity toward high potential substrates. Examining the behaviors of the T1D mutant was expected to enhance our understanding about whether the Cu_{T1} is absolutely necessary even for the oxidation of high potential substrates.

5.2. T1D & T1D-W118F

The Cu_{T1} depleted (T1D) laccase was produced by mutating Cys 445 to serine to prevent copper incorporation to the Cu_{T1} site as previously reported [23]. The C445S mutant indeed lacks the Cu_{T1} absorption band at 605 nm. When the activity of T1D laccase toward ABTS was tested, even the T1D laccase was found to be active toward ABTS (1% activity at k_{cat} of *Tth*-lac WT). However, since it is likely that this minute activity is due to the presence of a small amount of

Cu_{T1}-bound enzymes at equilibrium, another confirmation was needed to test out the possibility of another pathway and the potential involvement of redox active amino acids. To cut off both presumable pathways, a surface Trp mutation (W118F) was made to the T1D laccase, and the resulting mutant activity was examined to see if the 1% activity of T1D laccase relative to WT activity goes away (Figure 5.4). The result was that both T1D laccase and T1D-W118F exhibited the identical activity corresponding to 1% of WT activity toward ABTS in the saturating regime of ABTS concentration (k_{cat} of WT). From these observations, it seems like the surface mutation does not impair the enzyme activity to a significant degree.

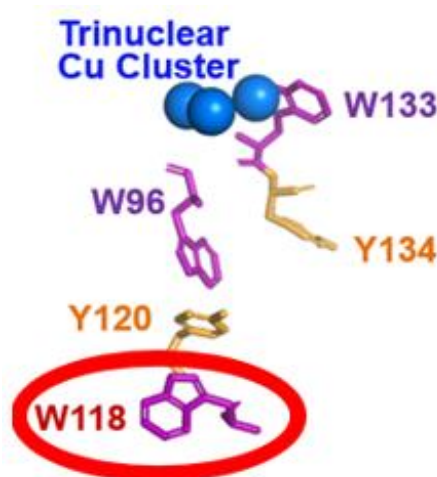


Figure 5.4. T1D enzyme (C445S) and its surface Trp mutant (C445S-W118F).

Going back to the consensus mechanism of MCOs, another likely candidate that can provide enough driving force for the oxidation of high potential substrates could be the peroxide intermediate (PI) which has been proposed as a catalytically relevant precursor to the native intermediate (NI) [1]. The PI is designated as a peroxide-level species with two coppers of the TNC oxidized and peroxide-bridged while one copper still remains to be reduced (see chapter 4 for more details). The PI can be generated by reacting a fully reduced T1D laccase with dioxygen [1], and two charge transfer bands arise at 340 nm and 480 nm as distinct spectral

features [24]. If the PI has enough potential to reduce high potential substrates, electrons could go directly into the TNC and the W/Y chain could still run the catalytic cycle (as a fewer-than-four-electron process) even without the involvement of Cu_{T1} . However, no spectral feature of the PI was detected from *Thh*-lac even after multiple trials with both T1D laccase and T1D-W118F. The reasons could be the instability of the PI in *Thh*-lac or the existence of alternative reaction pathways which deviate from the consensus for MCO catalysis. Other ways of generating a peroxide intermediate-like species reported are adding one reducing equivalent of H_2O_2 to the fully oxidized enzyme or reacting a partially reduced (with two reducing equivalents) laccase with O_2 [25]. The transient species generated from *Rhus* laccase with these methods resembled the PI generated with the T1D laccase, and seemed to persist up to around 24 h [25]. Therefore, a future direction is to try to generate this PI-like species, characterize it with the resonance Raman spectroscopy and test its reactivity with ABTS or promazine hydrochloride. Although it is currently still not conclusive whether the PI plays a role in oxidizing high potential substrates or not, the reaction with dioxygen is certainly a key for laccase function, as the enzyme starts turning over only in the presence of O_2 and none of the active site coppers alone has enough potential to perform the job on its own.

It should also be noted that the substrate potential itself may also be significantly altered upon binding of substrates on the enzyme surface or to the catalytic sites. In order to identify the catalytic sites for different substrates, structural analysis of the enzyme with X-ray crystallography or molecular docking simulation might be helpful. However, the enzyme crystallization for both WT and the mutants (W118F and W133F) has not been successful even with hundreds of screening conditions. A small crystal was obtained with 0.2 M ammonium sulfate, 0.1 M bis-tris, pH 6.5 and 25 w/v polyethylene glycol 3350, but it was too small to resolve the structure even with the micro-ED. Molecular docking simulations can be of a

benefit to calculate energy difference upon substrate binding and estimate the potential change of substrates during catalysis.

5.3. Met-rich Loop Mutant

A burst of activity for *Tth*-lac WT in the presence of excess copper was reported by a former study, and the highest degree of activity enhancement was reported with 1 mM CuSO₄ supplementation [26]. As this phenomenon is observed even for the fully metalated enzymes, the presence of extra transient metal binding sites has been speculated, and the most likely candidate is the Met-rich loop in *Tth*-lac. The Met-rich loop is observed as a common structural motif in several multi-copper oxidases such as CueO and McoA [27]–[29], and there exists a Met-rich loop in *Tth*-lac as well close to the Cu_{T1}, which consists of eleven methionines and one histidine and one aspartate (Figure 5.5). Since the idea of an extra transient metal binding site coordinated by Met, His and Asp contributing to enzyme activity has been suggested in other copper enzymes such as CueO, it was tested if a similar scenario applies to *Tth*-lac as well.

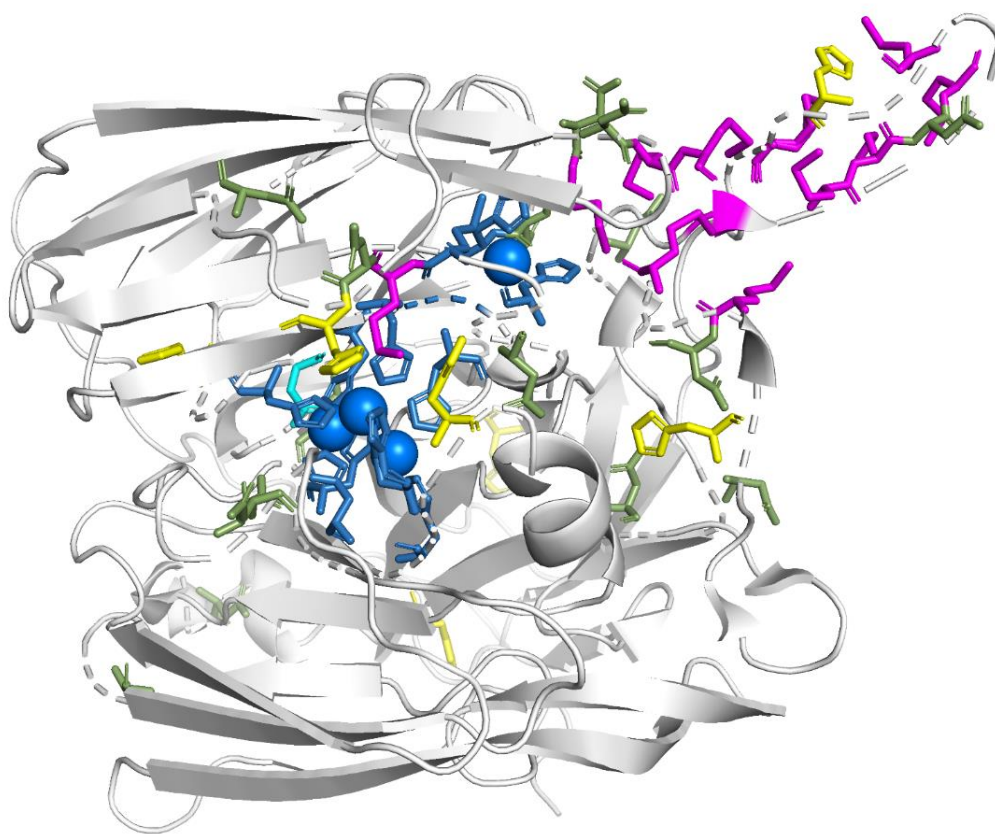


Figure 5.5. Distribution of Met (magenta), His (yellow), Asp (green), and active-site copper-coordinating residues (blue) in *Tth*-lac (PDB: 2YAE [2]).

To investigate the roles of metal binding residues in the Met-rich loop present in *Tth*-lac, the 6-His-tag was removed, since histidines in the tag may affect the enzyme activity in the presence of excess copper. All Met, Asp and His residues in the loop were replaced by Ala which is a metal non-coordinating residue to compare the activity with WT in the presence and absence of excess copper in the media.

Original Sequence (*Tth*-lac WT)

GPSFPEPKVRSQGGLLSLKLSATPTPLA IAGQRATLLTYGGSFPGPTLR
 VRPRDTVRLTLENRLPEPTNLHWHGLP I SPKVDDPFLE I PPGESWTYEFT
 VPKELAGTFWYHPHLHGRVAPQLFAGLLGALVVESSDA I PELREAEEHL
 LVLKDLALQGGRPAPHTPMDWNGKEGLVLVNGALRPTLVAQKATLRLR
 LLNASNARYRLALQDHPLYL I AADGGFLEEPLEVSELLAPGERAEVLV
 RLRKEGRFLLQALPYDRGAMMDMGGMAHAM PQGSPRPETLLYL I APKN
 PKPLPLPKALSPFPTLPAPVVTRRLVLTEDMMAARFF I NGQVFDHRRVDL
 KGQAQTVEVWEVENQGDADHPFHLHVHPFQVLSVGGRPFPYRAWKDVVNL
 KAGEVARLLVPLREKGRTVFFCH I VEHEDRGMIGVLEVG

Met-rich-loop – Ala mutant (*Tth*-lac MA Mutant)

GPSFPEPKVRSQGGLLSLKLSATPTPLA IAGQRATLLTYGGSFPGPTLR
 VRPRDTVRLTLENRLPEPTNLHWHGLP I SPKVDDPFLE I PPGESWTYEFT
 VPKELAGTFWYHPHLHGRVAPQLFAGLLGALVVESSDA I PELREAEEHL
 LVLKDLALQGGRPAPHTPADWANGKEGLVLVNGALRPTLVAQKATLRLR
 LLNASNARYRLALQDHPLYL I AADGGFLEEPLEVSELLAPGERAEVLV
 RLRKEGRFLLQALPYDRGAGAAAAGGAAAPQGSPRPETLLYL I APKN
 PKPLPLPKALSPFPTLPAPVVTRRLVLTEDAAAARFF I NGQVFDHRRVDL
 KGQAQTVEVWEVENQGDADHPFHLHVHPFQVLSVGGRPFPYRAWKDVVNL
 KAGEVARLLVPLREKGRTVFFCH I VEHEDRGMIGVLEVG

The enzymes were metalated in cells by adding 0.5 mM CuSO₄ to the growth media following the published metalation procedure for CotA laccase with slight modification [30]. Metalated enzymes were purified in MES buffer, pH 6 using a HP SP cationic exchange column with a 500 mM NaCl gradient over the duration of 2 hrs. Enzymes were stored in Tris, pH 8 with 1 mM CuSO₄ until use.

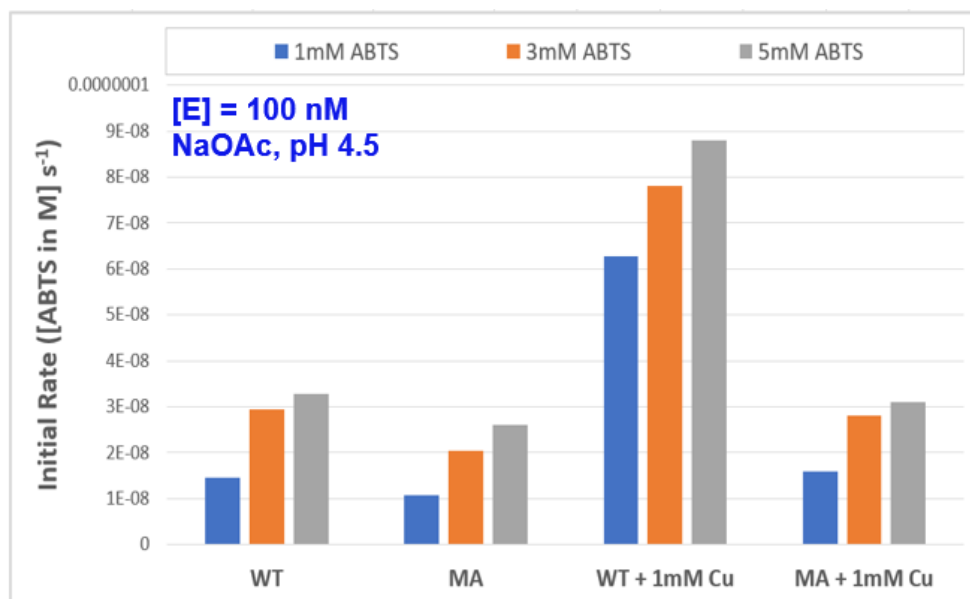


Figure 5.6. Activity comparison of *Tth*-lac WT and *Tth*-lac MA (all Met, Asp and His in the Met-rich loop mutated to Ala) in the presence and absence of excess copper in the activity assay media (NaOAc, pH 4.5).

Both *Tth*-lac WT and the *Tth*-lac MA mutant exhibited almost identical activities in the absence of excess CuSO_4 in the assay media (Figure 5.6). However, with 1 mM CuSO_4 supplementation, a burst of activity was observed only for the WT sample as previously reported [26]. To elucidate the reason for this observation, another mutant (*Tth*-lac DH) which had only Asp and His in the loop substituted to Ala was also made and tested for its activity. Since Met can only bind Cu^{1+} , if extra Cu^{2+} ions transiently bound in the presence of excess CuSO_4 are responsible for the activity enhancement, Asp and His were the likely candidates that could provide a cavity for Cu^{2+} binding.

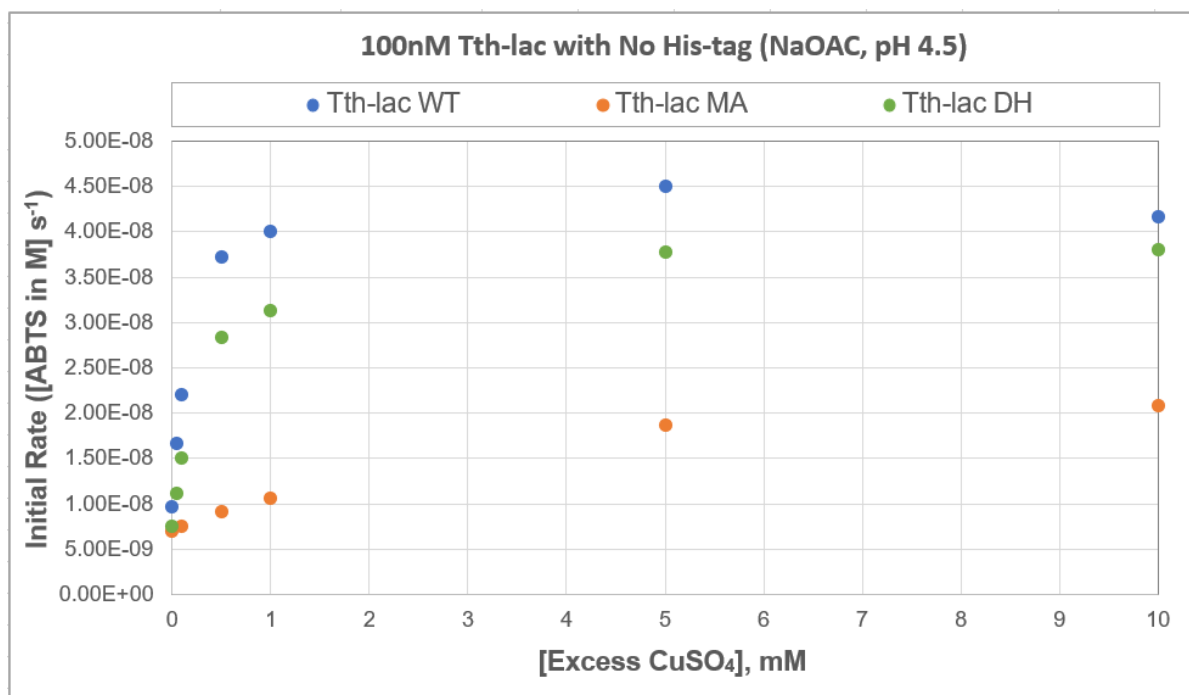


Figure 5.7. Effects of excess copper concentration on the degree of activity enhancement for *Tth-lac* WT, *Tth-lac* MA (All Met, His and Asp in the loop mutated to Ala) and *Tth-lac* DH (Only His and Asp mutated to Ala).

Surprisingly, the activity enhancement was observed (at a slightly smaller degree) even for the *Tth-lac* DH mutant (Figure 5.7). Although the burst of activity is not observed for the *Tth-lac* MA mutant up to around 1 mM Cu²⁺ supplementation in the assay media, the increase in activity became more distinct at higher excess copper concentration (5 mM, 10 mM). Whereas for *Tth-lac* WT and *Tth-lac* DH, the activity reached the max with <500 μM excess copper concentration, as previously observed by a former study on *Tth-lac* [26]. Moreover, at low copper concentration (50 μM, 100 μM), the degree of activity enhancement for *Tth-lac* DH was smaller than that for *Tth-lac* WT. This may imply that Asp and His are somewhat involved in transient copper binding, but the activity enhancement with excess copper can still occur only with methionines in the loop. There are two more aspartates on the surface close to the loop and the Cu_{T1}, so it can be speculated that those residues may also be involved, but since they

are too close to the Cu_{T1}, mutations on those residues might severely disrupt the structure or affect the potential of the Cu_{T1}. Current efforts are on trying to get crystal structures of the Met-rich loop mutants to examine how the loop conformation changes for each mutant and to see if copper can bind to any cavities on the loop.

5.4. CotA Laccase from *Bacillus Subtilis* (CotA-lac)

Structural Homology

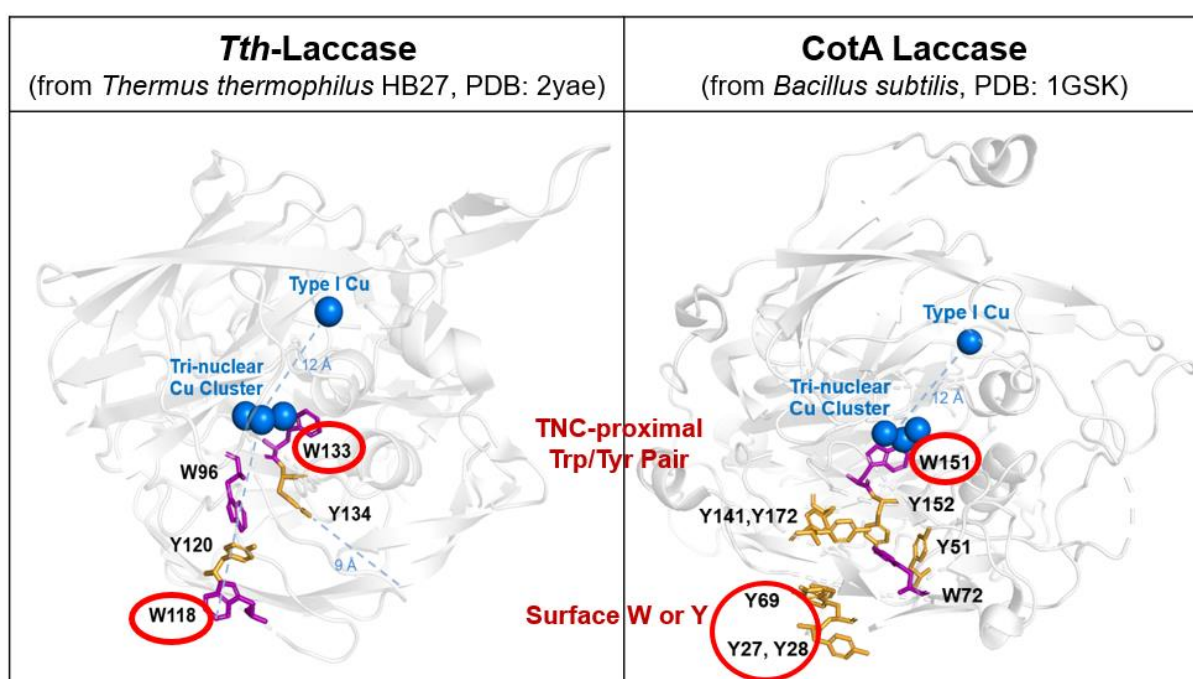


Figure 5.8. Structures of *Tth*-lac and CotA-lac (PDB: 2YAE [2], 1GSK [31])

CotA laccase (CotA-lac) is a spore coat protein from *Bacillus subtilis*, known to exhibit bilirubin oxidase activity (oxidation of bilirubin to biliverdin) [32], [33]. Since *Tth*-lac and CotA-lac are both three domain laccases structurally similar to each other, gene synthesis of CotA-lac was done to draw some parallel conclusions. In CotA laccase as well, there are chains of Trp and Tyr residues stretching from the TNC as well as a TNC-proximal Trp/Tyr pair which is closely positioned from other Tyr and Trp residues (Figure 5.8).

CotA-lac gene was cloned between NdeI and XhoI sites of the pET-22b(+) vector. By choosing the 5'NdeI cloning site, the pelB leader sequence was detached from the gene for expression in the cytoplasmic region of the cell. A C-terminal stop codon was added to prevent the 6 x his-tag fusion. After the gene synthesis, the protein expression was done following the published procedure with slight modification [30]. The TB media with 0.4% glycerol supplementation gave higher yield of the protein than the LB media, and CuSO₄ was added to make the final concentration of 0.5 – 0.7 mM in the media depending on the cell density ($0.6 < OD < 1$). The enzyme purification was done with a HP SP cationic exchange column following the published procedure with slight modifications [34], [35]. Although the published procedures use the buffer pH of 7.6, buffers at lower pH (Tris pH 7, MES pH 6 or MOPS pH 6.5) were used to facilitate better binding of the CotA-lac to the cationic exchange column.

The TNC-proximal Trp mutation (W151F) was made to see if it lowers the total turnover number. As expected, the W151F point mutation did not cause any change in activity and total turnover number due to the presence of a Tyr/Trp chain stretching from W151 out to the enzyme surface. A W151F/Y152F double mutant is of interest just like in *Tth*-lac (see Chapter 4 for more details), but in CotA-lac, there also exist Y51 and W72 along the chain following the W151/Y152 pair, and these residues could potentially participate in hole hopping. Making a quadruple mutant with all of these residues mutated to Phe may cause severe structural disruption, which will require structural analysis as well.

Just like in *Tth*-lac, testing the roles of surface-exposed redox-active amino acid at the end of the Trp/Tyr chain was of interest. In CotA-lac, there exist three tyrosines exposed on the surface: Y27, Y28 and Y69. These residues are all within 5Å from each other forming a surface Tyr

cluster. Each of these residues were mutated to a redox-inactive Phe to see if any of the mutations knocks out the enzyme activity. No significant difference was observed between the activities of WT and Y69F. In addition, although both Phe and Ala substitutions were tried, the enzyme did not express with all of the Y27F, Y28F, Y27F-Y28F, Y27A and Y28A mutant plasmids.

5.5. References

- [1] S. M. Jones and E. I. Solomon, “Electron transfer and reaction mechanism of laccases,” *Cellular and Molecular Life Sciences*, vol. 72, no. 5. Birkhauser Verlag AG, pp. 869–883, Mar. 01, 2015, doi: 10.1007/s00018-014-1826-6.
- [2] H. Serrano-Posada, S. Centeno-Leija, S. P. Rojas-Trejo, C. Rodríguez-Almazán, V. Stojanoff, and E. Rudiño-Piñera, “X-ray-induced catalytic active-site reduction of a multicopper oxidase: Structural insights into the proton-relay mechanism and O₂-reduction states,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 71, no. Pt 12, pp. 2396–2411, Nov. 2015, doi: 10.1107/S1399004715018714.
- [3] K. Brijwani, A. Rigdon, and P. V. Vadlani, “Fungal laccases: Production, function, and applications in food processing,” *Enzyme Research*, vol. 2010. 2010, doi: 10.4061/2010/149748.
- [4] K. Piontek, M. Antorini, and T. Choinowski, “Crystal structure of a laccase from the fungus *Trametes versicolor* at 1.90-Å resolution containing a full complement of coppers,” *J. Biol. Chem.*, vol. 277, no. 40, pp. 37663–37669, Oct. 2002, doi: 10.1074/jbc.M204571200.
- [5] S. Kurniawati and J. A. Nicell, “Characterization of *Trametes versicolor* laccase for the transformation of aqueous phenol,” *Bioresour. Technol.*, vol. 99, no. 16, pp. 7825–7834, Nov. 2008, doi: 10.1016/j.biortech.2008.01.084.
- [6] M. Gunne, A. Höppner, P. L. Hagedoorn, and V. B. Urlacher, “Structural and redox properties of the small laccase Ssl1 from *Streptomyces sviveus*,” *FEBS J.*, vol. 281, no. 18, pp. 4307–4318, Sep. 2014, doi: 10.1111/febs.12755.
- [7] J. Shin, H. B. Gray, and J. R. Winkler, “Stability/activity tradeoffs in *Thermus*

- thermophilus HB27 laccase,” *J. Biol. Inorg. Chem.*, vol. 25, no. 2, 2020, doi: 10.1007/s00775-020-01754-7.
- [8] P. Durão, I. Bento, A. T. Fernandes, E. P. Melo, P. F. Lindley, and L. O. Martins, “Perturbations of the T1 copper site in the CotA laccase from *Bacillus subtilis*: Structural, biochemical, enzymatic and stability studies,” *J. Biol. Inorg. Chem.*, vol. 11, no. 4, pp. 514–526, Jun. 2006, doi: 10.1007/s00775-006-0102-0.
- [9] J. Gallaway, I. Wheeldon, R. Rincon, P. Atanassov, S. Banta, and S. C. Barton, “Oxygen-reducing enzyme cathodes produced from SLAC, a small laccase from *Streptomyces coelicolor*,” *Biosens. Bioelectron.*, vol. 23, no. 8, pp. 1229–1235, Mar. 2008, doi: 10.1016/j.bios.2007.11.004.
- [10] Y. Zhu, X. Ouyang, Y. Zhao, L. Jiang, H. Guo, and X. Qiu, “Oxidative depolymerization of lignin improved by enzymolysis pretreatment with laccase,” *J. Energy Chem.*, vol. 27, no. 3, pp. 801–805, May 2018, doi: 10.1016/j.jechem.2017.04.018.
- [11] J. Lind, X. Shen, T. E. Eriksen, and G. Merényi, “The One-Electron Reduction Potential of 4-Substituted Phenoxy Radicals in Water,” *J. Am. Chem. Soc.*, vol. 112, no. 2, pp. 479–482, 1990, doi: 10.1021/ja00158a002.
- [12] X. Liu, M. Gillespie, A. D. Ozel, E. Dikici, S. Daunert, and L. G. Bachas, “Electrochemical properties and temperature dependence of a recombinant laccase from *Thermus thermophilus*,” *Anal. Bioanal. Chem.*, vol. 399, no. 1, pp. 361–366, Jan. 2011, doi: 10.1007/s00216-010-4345-9.
- [13] R. Bourbonnais, D. Leech, and M. G. Paice, “Electrochemical analysis of the interactions of laccase mediators with lignin model compounds,” *Biochim. Biophys.*

- Acta - Gen. Subj.*, vol. 1379, no. 3, pp. 381–390, Mar. 1998, doi: 10.1016/S0304-4165(97)00117-7.
- [14] T. Matsubara and P. C. Ford, “Some Applications of Cyclic Voltammetry to the Reactions and Properties of Ruthenium Ammine Complexes. Reduction Potentials and Rate Studies,” *Inorg. Chem.*, vol. 15, no. 5, pp. 1107–1110, May 1976, doi: 10.1021/ic50159a025.
- [15] S. L. Scott, W. J. Chen, A. Bakac, and J. H. Espenson, “Spectroscopic parameters, electrode potentials, acid ionization constants, and electron exchange rates of the 2,2'-azinobis(3-ethylbenzothiazoline-6-sulfonate) radicals and ions,” *J. Phys. Chem.*, vol. 97, no. 25, pp. 6710–6714, 1993, doi: 10.1021/j100127a022.
- [16] A. S. Pavitt, E. J. Bylaska, and P. G. Tratnyek, “Oxidation potentials of phenols and anilines: correlation analysis of electrochemical and theoretical values,” *Environ. Sci. Process. Impacts*, vol. 19, no. 3, pp. 339–349, Mar. 2017, doi: 10.1039/c6em00694a.
- [17] T. Kenzom, P. Srivastava, and S. Mishra, “Structural insights into 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS)-mediated degradation of reactive blue 21 by engineered *Cyathus bulleri* laccase and characterization of degradation products,” *Appl. Environ. Microbiol.*, vol. 80, no. 24, pp. 7484–7495, Dec. 2014, doi: 10.1128/AEM.02665-14.
- [18] J. Kulys, K. Krikstopaitis, and A. Ziemys, “Kinetics and thermodynamics of peroxidase- and laccase-catalyzed oxidation of N-substituted phenothiazines and phenoxazines,” *J. Biol. Inorg. Chem.*, vol. 5, no. 3, pp. 333–340, 2000, doi: 10.1007/PL00010662.
- [19] H. J. Altmann, “Bestimmung von in Wasser gelöstem Sauerstoff mit

- Leukoberbelinblau I - Eine schnelle Winkler-Methode," *Fresenius' Zeitschrift für Anal. Chemie*, vol. 262, no. 2, pp. 97–99, Jan. 1972, doi: 10.1007/BF00425919.
- [20] P. P. Champagne, M. E. Nesheim, and J. A. Ramsay, "A mechanism for NaCl inhibition of Reactive Blue 19 decolorization and ABTS oxidation by laccase," *Appl. Microbiol. Biotechnol.*, vol. 97, no. 14, pp. 6263–6269, Jul. 2013, doi: 10.1007/s00253-012-4525-y.
- [21] I. Ayuso-Fernández, F. J. Ruiz-Dueñas, and A. T. Martínez, "Evolutionary convergence in lignin-degrading enzymes," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 115, no. 25, pp. 6428–6433, Jun. 2018, doi: 10.1073/pnas.1802555115.
- [22] L. T. Mai Pham, S. J. Kim, and Y. H. Kim, "Improvement of catalytic performance of lignin peroxidase for the enhanced degradation of lignocellulose biomass based on the imbedded electron-relay in long-range electron transfer route," *Biotechnol. Biofuels*, vol. 9, no. 1, p. 247, Nov. 2016, doi: 10.1186/s13068-016-0664-1.
- [23] A. W. J. W. Tepper, T. J. Aartsma, and G. W. Canters, "Channeling of electrons within SLAC, the small laccase from *Streptomyces coelicolor*," *Faraday Discussions*, vol. 148, no. 0. The Royal Society of Chemistry, pp. 161–171, Dec. 14, 2011, doi: 10.1039/c002585b.
- [24] A. E. Palmer, Sang Kyu Lee, and E. I. Solomon, "Decay of the peroxide intermediate in laccase: Reductive cleavage of the O-O bond," *J. Am. Chem. Soc.*, vol. 123, no. 27, pp. 6591–6599, 2001, doi: 10.1021/ja010365z.
- [25] M. Goldberg, O. Farver, and I. Pecht, "Interaction of Rhus laccase with dioxygen and its reduction intermediates.," *J. Biol. Chem.*, vol. 255, no. 15, pp. 7353–7361, Aug. 1980, doi: 10.1016/s0021-9258(20)79711-5.

- [26] K. Miyazaki, “A hyperthermophilic laccase from *Thermus thermophilus* HB27,” *Extremophiles*, vol. 9, no. 6, pp. 415–425, Dec. 2005, doi: 10.1007/s00792-005-0458-z.
- [27] K. Y. Djoko, # Lee, X. Chong, A. G. Wedd, and Z. Xiao, “Reaction mechanisms of the multicopper oxidase CueO from *Escherichia coli* support its functional role as a cuprous oxidase,” doi: 10.1021/ja9091903.
- [28] S. K. Singh *et al.*, “Crystal structures of multicopper oxidase CueO bound to copper(I) and silver(I): functional role of a methionine-rich sequence,” *J Biol Chem.*, vol. 286, no. 43, pp. 37849-57, Oct. 2011, doi: 10.1074/jbc.M111.293589.
- [29] P. T. Borges *et al.*, “Methionine-rich loop of multicopper oxidase McoA follows open-to-close transitions with a role in enzyme catalysis,” *ACS Catal.*, vol. 10, no. 13, pp. 7162–7176, Jul. 2020, doi: 10.1021/acscatal.0c01623.
- [30] P. Durão *et al.*, “Copper incorporation into recombinant CotA laccase from *Bacillus subtilis*: Characterization of fully copper loaded enzymes,” *J. Biol. Inorg. Chem.*, vol. 13, no. 2, pp. 183–193, Feb. 2008, doi: 10.1007/s00775-007-0312-0.
- [31] F. J. Enguita, L. O. Martins, A. O. Henriques, and M. A. Carrondo, “Crystal structure of a bacterial endospore coat component: A laccase with enhanced thermostability properties,” *J. Biol. Chem.*, vol. 278, no. 21, pp. 19416–19425, May 2003, doi: 10.1074/jbc.M301251200.
- [32] N. Mano, “Features and applications of bilirubin oxidases,” *Applied Microbiology and Biotechnology*, vol. 96, no. 2. Springer, pp. 301–307, Oct. 03, 2012, doi: 10.1007/s00253-012-4312-9.
- [33] S. I. Sakasegawa, H. Ishikawa, S. Imamura, H. Sakuraba, S. Goda, and T. Ohshima,

- “Bilirubin oxidase activity of *Bacillus subtilis* CotA,” *Appl. Environ. Microbiol.*, vol. 72, no. 1, pp. 972–975, Jan. 2006, doi: 10.1128/AEM.72.1.972-975.2006.
- [34] I. Bento, L. O. Martins, G. G. Lopes, M. A. Carrondo, and P. F. Lindley, “Dioxygen reduction by multi-copper oxidases; a structural perspective,” *Dalt. Trans.*, no. 21, pp. 3507–3513, Nov. 2005, doi: 10.1039/b504806k.
- [35] L. O. Martins *et al.*, “Molecular and biochemical characterization of a highly stable bacterial laccase that occurs as a structural component of the *Bacillus subtilis* endospore coat,” *J. Biol. Chem.*, vol. 277, no. 21, pp. 18849–18859, May 2002, doi: 10.1074/jbc.M200827200.

Chapter 6

Investigating the Mechanistic Details of Lytic Polysaccharide Monooxygenases (LPMOs)

6.1. Introduction

Although details about the LPMO catalysis still remain debatable, $\text{Cu}^{3+}(\text{O})$ has been involved in both O_2 mechanism and H_2O_2 mechanism as a potential reactive intermediate [1], [2] (refer to Chapter 1 for more details). Two cellulose-active LPMOs (ScLPMO10B and ScLPMO10C) and one chitin-active LPMO (BILPMO10A) were studied in order to elucidate if $\text{Cu}^{3+}(\text{O})$ can be generated and tested for its reactivity toward polysaccharides.

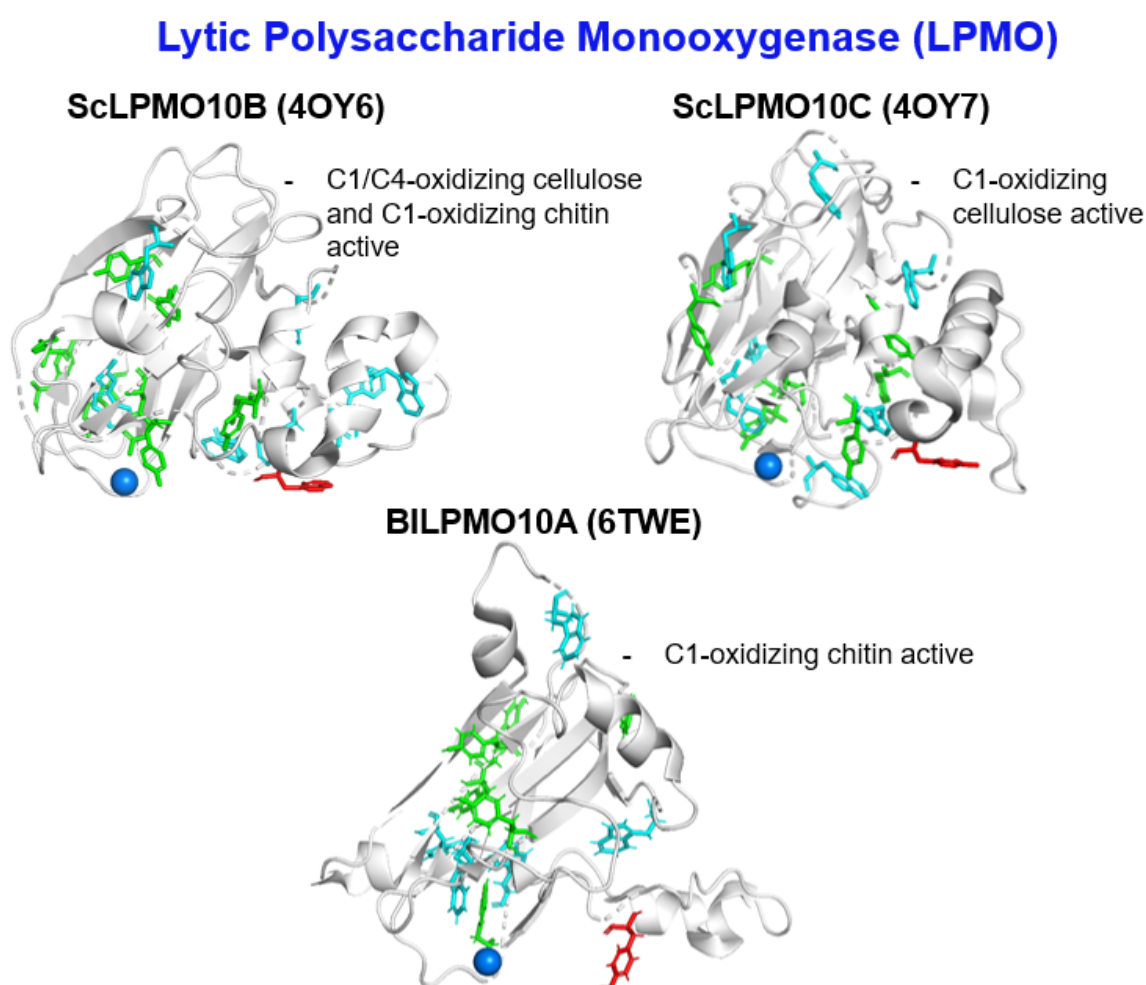


Figure 6.1. Structures of bacterial LPMOs (PDB: 4OY6 [3], 4OY7 [3], 6TWE [4]) active towards cellulose and chitin: an aromatic amino acid on the cellulose/chitin binding domain (red), Trp (cyan) and Try (green) residues.

As can be seen in Figure 6.1., LPMOs are all structurally very similar in a way that there is a surface-exposed active site type 2 copper ($\text{Cu}_{\text{T}2}$) and a surface-exposed aromatic amino acid which form a putative cellulose binding domain together. Just like MCOs, these enzymes also have chains of closely spaced tryptophans and tyrosines stretching from the active site copper on the surface (Figure 6.1).

In efforts to elucidate how LPMO operates, spectroscopic approaches were taken to photochemically trigger the redox chemistry and to transiently monitor the generation of Cu^{3+} species. Moreover, after the photochemical generation of Cu^{3+} species, their reactivity toward cellulose and chitin was tested in photodegradation experiments to determine if Cu^{3+} species participate in degrading polysaccharides. Enzymes have been successfully expressed in *E. coli* and were purified with the use of Avicel or the ionic exchange columns after the osmotic shock procedure to obtain periplasmic extracts.

6.2. Enzyme Preparation

An example of a LPMO vector construct

- **Gene Synthesis**

- **Gene name:** scLPMO10B
- **#bp:** 575
- **Cloning:** Custom cloning Vector name: pET-26b(+) [Kanamycin-resistant]
- **Cloning site 5':** NcoI **Cloning site 3':** XhoI

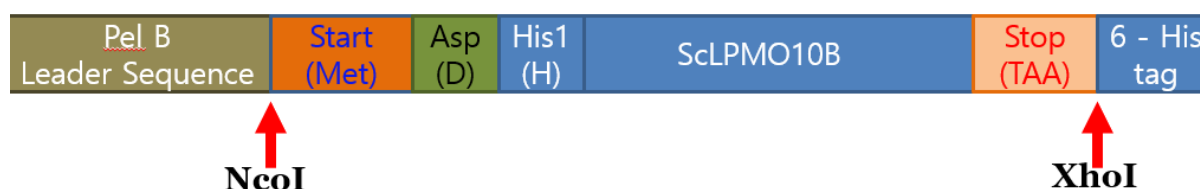


Figure 6.2. A schematic diagram of ScLPMO10B gene construct.

Sequence details

gccagccggcgATGgc/ATGgAT/catgtagtgtggttgatccggccagccgtaattatggttctggtggaacgctggggc
 gatgatttcagaatccggcaatggccgatgaagatccgatgtgctggcaggcctggcaggatgatccgaatgcaatgtggaattgga
 atggtctgtatcgaatggcagtgacggtgactttgaagccgtgtccggatggccagctgtgcagcgggtgctgcaccgaaagtgg
 ccgttataatagcctggatgccgtgggcccgtggcagaccaccgatgtgaccgatgattttaccgtgaaactgcatgatcaggccagtc
 atggcggcgattattttctggtgtatgtgaccaaacagggctttgatccggcaaccaggccctgacctggggtgaaactgcagcaggtt
 gcacgtaccggtagctatggtccgagccagaattatgaaattccggttagcaccagtggctgaccggccgtcatgtggtttataaccatt
 ggcaggcaagccatattgatcagacctatttctgttagtgatggtgattttggcTAActcgag

The texts in orange color in the beginning of the sequence above indicate the **pelB leader sequence** which exists in the **pET-26b vector**. This signaling sequence should remain intact, since it will direct the over-expressed recombinant protein to the periplasmic membrane and be cleaved off to produce the enzyme of interest. The pelB leader sequence itself has the **start codon ATG** indicated in orange bold.

In order to clone a gene into a specific vector, we should do it between certain restriction sites. In this case, the scLPMO10B gene sequence (black letters) was cloned between the **Nco I** (ccATGg) and **Xho I** (ctcgag) sites underlined and highlighted in yellow.

Once the pelB leader sequence gets cleaved off, the protein starts with what is left in the Nco I sequence ATGg. (The first two orange letters ‘cc’ of the Nco I site also belong to the pelB leader sequence.) Since three bases make up an amino acid, the 4th base ‘g’ in ATGg will cause a frame shift. (e.g. Instead of producing a protein that starts with the sequence ‘ATG/cat/ggt/agt/gtg/gtt,’ this leftover ‘g’ will produce ‘ATG/gca/tgg/tag/tgt/ggt’ which will give completely different sets of amino acids.) Therefore, the two protect-bases ‘AT’ in green

are added, so that it can make up an Asp amino acid (gAT). This Asp residue as well as Met from the start codon in the Nco I site (indicated in the sequence above as ‘/ATGgAT/’) was removed via site-directed mutagenesis to ensure that the N-terminal α -amino group of His (the next residue in the sequence shown as ‘cat’) is freely available for copper binding.

The protein expression can be achieved even without the start codon ATG in the Nco I site (in red bold letters) due to the presence of another start codon (ATG) in the pelB leader sequence (in orange bold letters). As mentioned earlier, the pelB leader sequence eventually gets cleaved off, so that the remaining protein can have the sequence in black letters (before the stop codon (TAA)).

Trypsin digestion

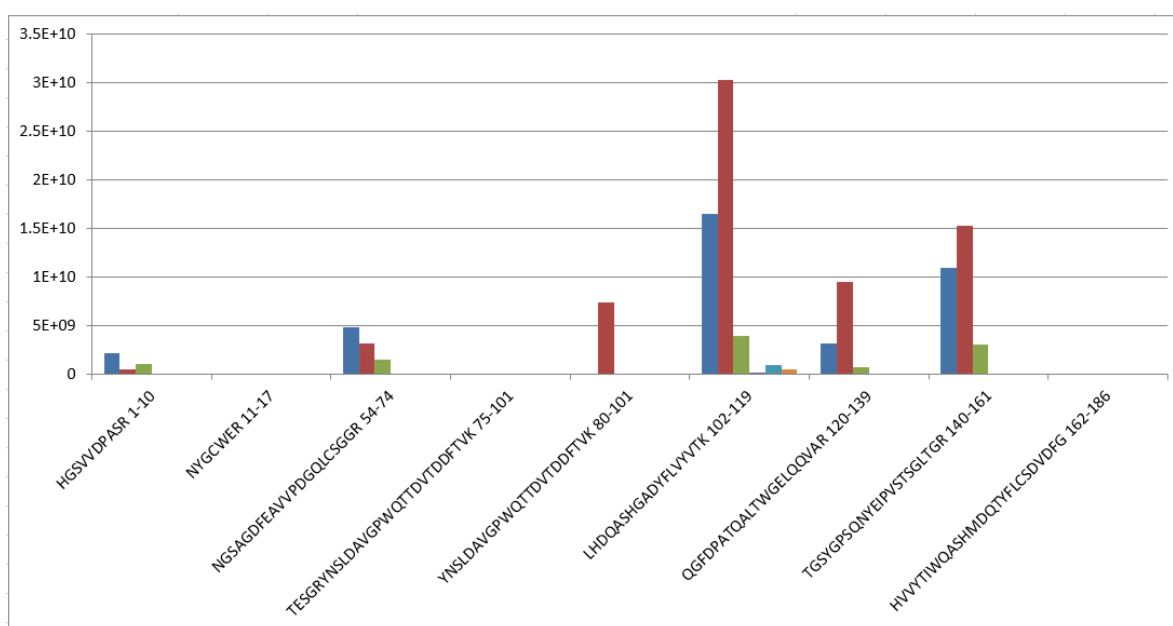


Figure 6.3. Sequence analysis of ScLPMO10B by trypsin digestion.

Since the expression yield of these bacterial LPMOs tend to be very low (5 mg/L), it was quite difficult to isolate the pure enzyme in the beginning. Before optimizing purification conditions,

it was made sure that an intact LPMO is getting expressed by confirming the sequence coverage with Trypsin digestion (Figure 6.3).

Avicel purification

The His-tag attachment should generally be avoided for copper proteins due to histidines on the tag competing for coppers being incorporated to the active sites. However, attaining high purity just with the FPLC purification using ionic exchange and size exclusion columns can be a challenging task, when the properties of other *E. coli* proteins in the periplasmic extract exhibit similar chemical properties with LPMOs. Another interesting way to purify LPMOs is using Avicel [5]. Avicel is a microcrystalline cellulosic substrate to which LPMOs can bind. With the addition of Avicel to the periplasmic extract obtained by applying osmotic shock, LPMOs selectively bind to Avicel. By centrifugation, LPMO-bound Avicel can be separated from the rest of the periplasmic extract. Upon the addition of glucose at close to 1 M concentration, selective elution of LPMOs can be achieved (Figure 6.4).

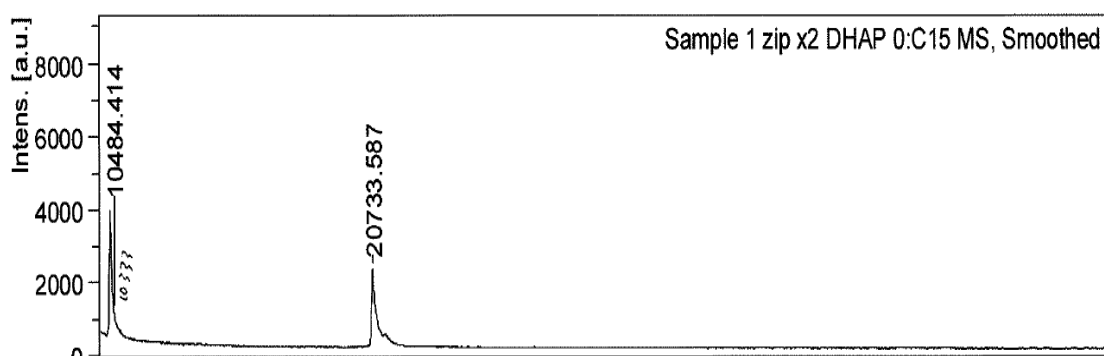


Figure 6.4. MALDI-TOF analysis of LPMO purified by the Avicel purification method.

However, the typical yield of ScLPMO10B was unfortunately less than 5 mg/L even after optimizing the conditions: with both TB and LB, with and without glucose or glycerol supplementation, different induction temperature from 25°C to 37°C and duration from 3 hr to

12 hr, and different IPTG concentration (from 0.1 mM to 1 mM). Therefore, after discussions with Dr. Vincent Eijsink and Dr. Zarah Forsberg at Norwegian University of Life Sciences (NMBU), the plasmids of ScLPMO10C and BILPMO10A which are known to express better than ScLPMO10B were obtained from the Eijsink group. Both genes are in the pRSETB vector which is ampicillin resistant, and the both plasmids contain a signaling sequence from SmLPMO10A (also known as CBP21) which functions similar to the pelB leader sequence.

Signal sequence from SmLPMO10A (CBP21)

ATGAACAAAACCTTCCCGTACCCTGCTCTCTCTGGGCCTGCTGAGCGCGGCCATGTTCCGGCGTTTCGCAACAGGC

BILPMO10A Sequence

HGFIEKPGSRAALCSEAFGFLNLCGSVMYEPQSLEAKKGFPHSGPADGQIASAGGLFGGILDQQSENRFKHIM
TGGEHTFTWYTTAPHNTSQWHYYITKKGWDPDKPLKRADFELIGAVPHDGSPPASRNLSHHIYIPEDRLLGYHVILA
VVDVADTENAFYQVIDVDLVNK

ScLPMO10C Sequence

HGVAMMPGSRTYLCQLDAKTGTGALDPTNPACQAALDQSGATALYNWFAVLDSNAGGRGAGYVPDGTLCGAGDRS
PYDFSAYNAARSDWPRTHLTSGATIPVEYSNWAHPGDFRVYLTKEGWSPTSELGWDDLELIQTVTNPPQQGSPG
TDGGHYYWDLALPSGRSGDALIFMQWVRSQENFFSCSDVVFDTGGNGEVTGIRGSGSTPDPDPTPTPTDPTTPTP
THTGSCMAVYSVENSWSGGFQGSVEVMNHGTEPLNGWAVQWQPGGGTTLGGVWNGSLTSGSDGTVTVRNVVDHNRV
VPPDGSVTFGFTATSTGNDFPVDSIGCVAP

AA10-linker-CBM2

The enzymes were expressed and purified following the published procedures with slight modifications [3], [6]. After giving osmotic shock to the harvested cells with Tris buffer, pH 8 containing 0.5 M sucrose and 100 μ M EDTA, periplasmic extracts were obtained and purified further with ion exchange chromatography and size exclusion chromatography. The enzymes have been found to be less stable in acidic solutions with low pH. Therefore, instead of using citrate buffer at pH 3.5 as suggested in the published protocol, Tris buffer at pH 7 or 8 were used for purification with anionic exchange chromatography.

6.3. Activity Assays

Cellulose degradation reactions / product analysis

For standard cellulose degradation reactions, general published procedures were followed [6]. To briefly outline, Avicel and phosphoric acid swollen cellulose (PASC) were incubated with 1 μ M LPMO in 20mM sodium phosphate at pH 6 in the presence of 2 mM ascorbic acid. The sample in an Ependorf tube was gently shaken at 150 rpm for 24 hours in the incubator set to 30°C to keep Avicel and PASC suspended in the LPMO solution. The degradation products of Avicel and PASC were analyzed by MALDI-TOF mass spectrometry. To investigate the rate of degradation, samples can be collected every hour or at certain time intervals. To analyze double-oxidized products at C1 and C4 both, *T. reesei* cellobiohydrolase and *thermophilum* cellobiose dehydrogenase (*MtCDH*) can be added to split the product into two singly oxidized monosaccharides [3], [6].

MALDI-TOF

The mass of the degradation products can be detected using the matrix-assisted laser desorption/ionization – time of flight (MALDI-TOF) mass spectrometry. The mass values of the products can be compared with the expected mass values to monitor the degree of cellulose degradation [6]. The same procedure was adopted for the β -chitin substrate as well to monitor the BILPMO10A activity.

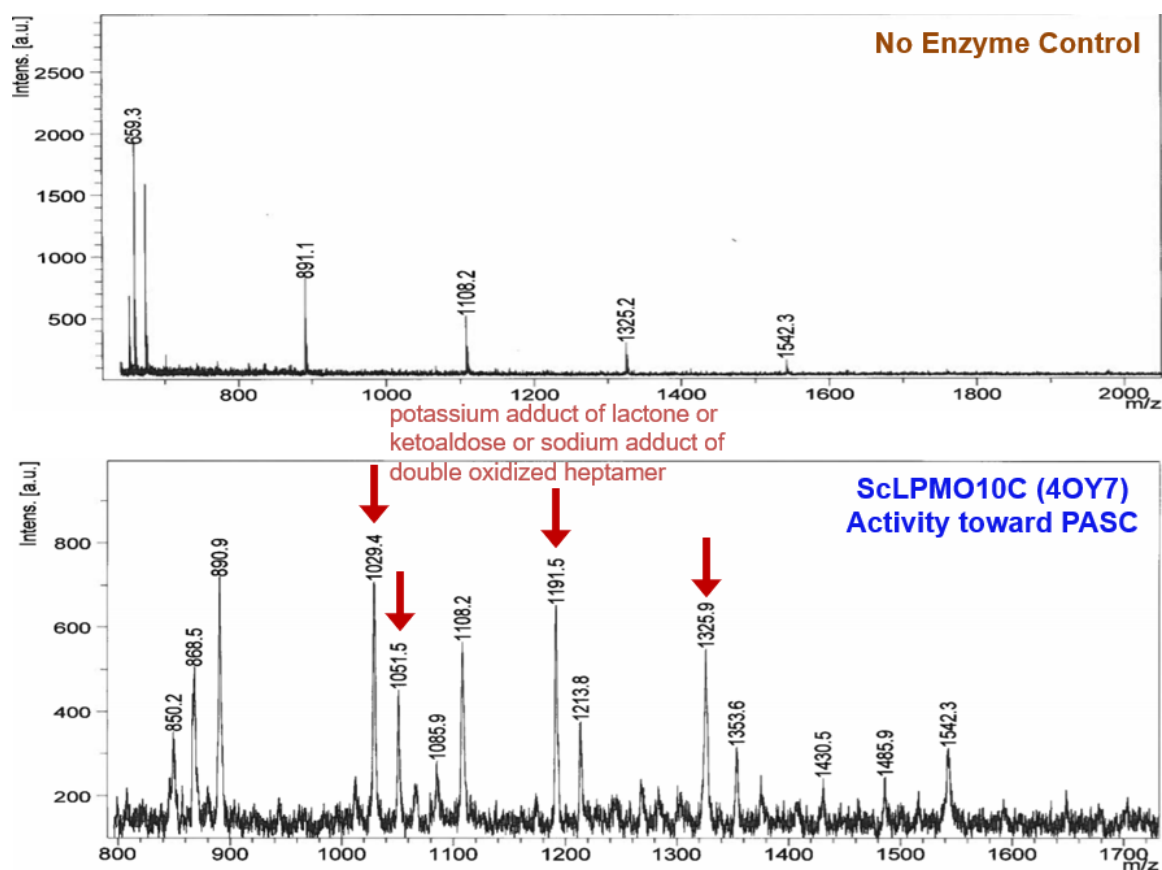
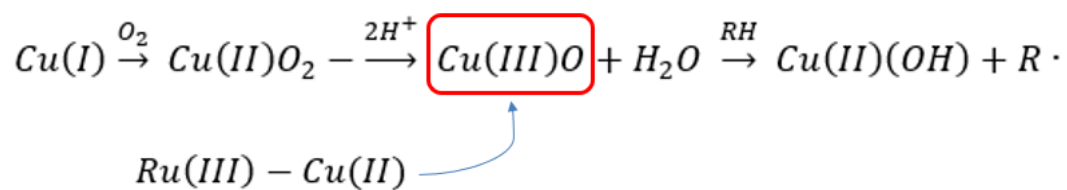


Figure 6.5. MALDI-TOF data showing degradation products of PASC (phosphoric acid swollen cellulose) in a standard activity assay with ScLPMO10C.

After the cellulose or chitin degradation experiments, product analysis was done with MALDI-TOF, and each degradation product with a known mass was identified (Figure 6.5.) in comparison to published values [6].

6.4. Photodegradation Experiment

Cu(III)(O) Intermediate Species



Although whether these enzymes are oxygenases or peroxygenases is still a hot debate, both of these postulated mechanisms include the generation of Cu(III) intermediates resulting from the oxidation of Cu(I) by O₂ or H₂O₂, followed by O-O bond cleavage. Therefore, in efforts to elucidate how LPMO operates, a spectroscopic approach was taken to photochemically trigger the redox chemistry and to transiently monitor the generation of Cu(III) species.

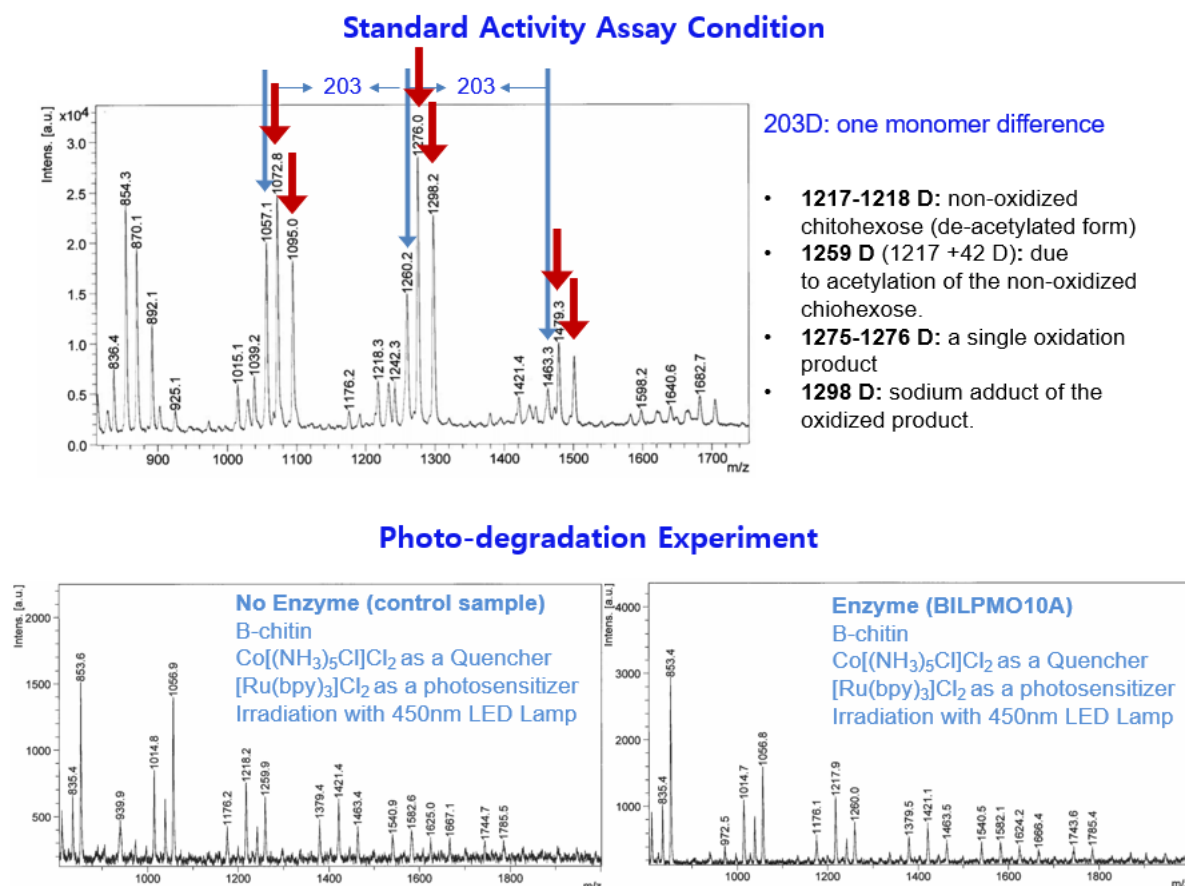


Figure 6.6. MALDI-TOF data showing degradation products of β -chitin in a standard activity assay with BILPMO10A (top) and no product formation in a photo-degradation experiment (bottom).

After confirming by product analysis with MALDI-TOF that my enzymes can indeed degrade recalcitrant cellulose or chitin substrates, photochemical studies were conducted to test if photochemically generated Cu^{3+} species are reactive toward cellulosic substrates. Samples were prepared as a mixture containing cellulose/chitin as a substrate, protein, Ru-photosensitizer ($\text{Ru}(\text{bpy})_3\text{Cl}_2$), and the irreversible cobalt quencher $[\text{Co}(\text{NH}_3)_5\text{Cl}]\text{Cl}_2$, and were deoxygenated. Irradiation of Ru-photosensitizer in the sample can generate Ru^{2+} excited state, and the cobalt quencher can produce more oxidizing Ru^{3+} species which could potentially

oxidize Cu^{2+} to Cu^{3+} . To ensure that sufficient driving force is applied by the photogenerated Ru^{3+} species, experiments were conducted with $\text{Ru}(\text{bpy})_3^{3+}$ ($E^\circ = 1.26 \text{ V vs. NHE}$), $\text{Ru}(\text{bpm})_3^{3+}$ ($E^\circ = 1.69 \text{ V vs. SSCE}$) and $\text{Ru}(\text{bpz})_3^{3+}$ ($E^\circ = 1.98 \text{ V vs. SSCE}$) [7], [8].

If Cu^{3+} species being generated are reactive toward cellulose/chitin, we should still be able to detect the soluble degradation products even from these photo-excited samples which lack both O_2 and reductants (ascorbic acid). Unfortunately, no product was detected from the photo-degradation experiment even with concentrated samples (no product peak for the photo-excited samples as opposed to the standard activity assay sample (Figure 6.6)), which may indicate that Cu^{3+} intermediates are not the active intermediates responsible for cellulose degradation.

6.5. Cu^{3+} Transient Absorption Spectrum

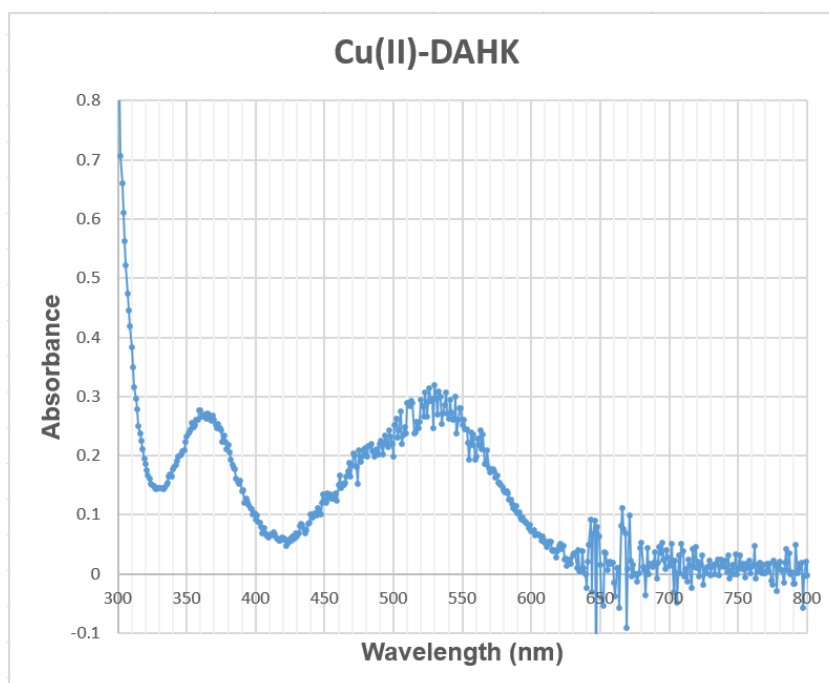


Figure 6.7. UV-Vis absorption spectrum of $\text{Cu}(\text{II})$ -bound DAHK peptide.

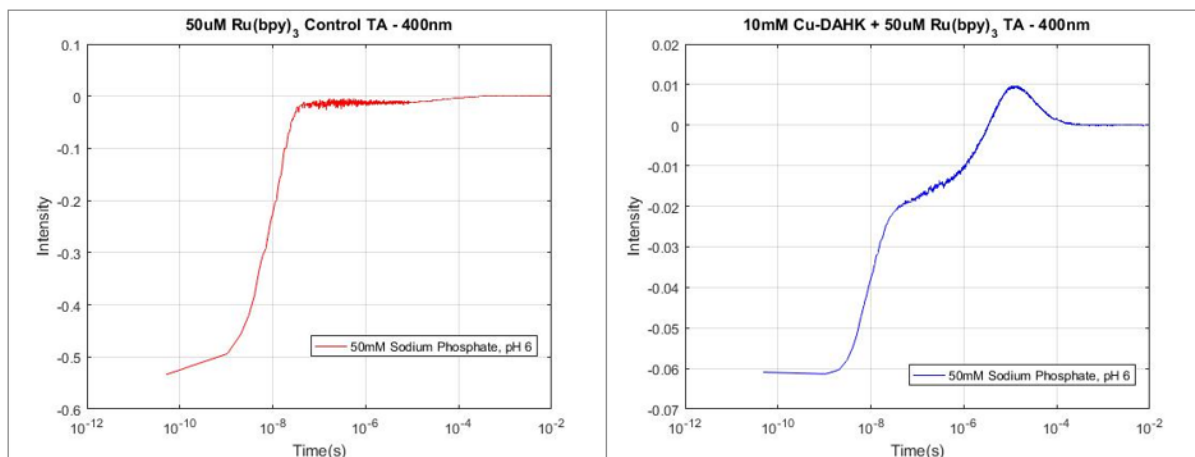


Figure 6.8. Transient absorption kinetics spectra of the photochemically generated DAHK-Cu(III) at 400 nm (right).

Although no activity was observed by the photo-excited samples, it was necessary to confirm that Cu^{3+} species were indeed being generated with the addition of Ru-compounds as photosensitizers. If not, it is also possible that the holes may have just ended up oxidizing a nearby Trp or Tyr. Therefore, I wanted to check the hole transfer pathway with transient absorption spectroscopy. The model study done with a copper bound DAHK peptide segment shows that the Cu^{3+} species have an absorption feature at around 350-400 nm (Figure 6.8). Since the coordination environment for copper is different in LPMOs, different absorption features at shifted wavelengths may arise, which could be more difficult to detect. But, even then, the decay kinetics of Ru^{3+} species can be compared between the protein sample and the control sample to see how the redox activity of Ru^{3+} changes in the presence of LPMOs upon Cu^{3+} generation. When the transient absorption spectroscopic experiments were tried with both ScLPMO10C and BILPMO10A samples, a longer lasting Ru^{3+} bleach signal was observed compared to the control sample, which could be an indication that it is due to the generation of Cu^{3+} . However, no positive absorption feature was observed anywhere between the 350-555 nm wavelength region where Cu^{3+} and Trp radical features are expected (Figures 6.9-6.14).

Making the XANES measurements could potentially provide us an evidence for Cu^{3+} generation, but it can also be difficult to resolve the spectra when the concentration of Cu^{3+} being formed is not high enough.

6.6. Appendices

Bimolecular quenching experiments with LPMOs and Ru-complexes as photosensitizers were monitored with the transient absorption spectroscopy at different wavelengths to examine the hole transfer kinetics. However, the kinetic profiles obtained with reversible quenchers ($[\text{Ru}(\text{NH}_3)_6]\text{Cl}_3$ and methyl viologen) were obscured by back reaction kinetics (Figures 6.9 - 6.14). No positive absorption features were observed anywhere between the 350 - 550 nm wavelength regions. When an irreversible quencher ($[\text{Co}(\text{NH}_3)_5\text{Cl}]\text{Cl}_2$) was tried, a longer lasting bleach for Ru^{3+} species was observed with the LPMO sample (Figure 6.16) compared to the control sample (Figure 6.15), which could be indicative of Cu^{3+} generation. But again, no positive absorption features distinctive for either Cu^{3+} or Trp radical species were detected.

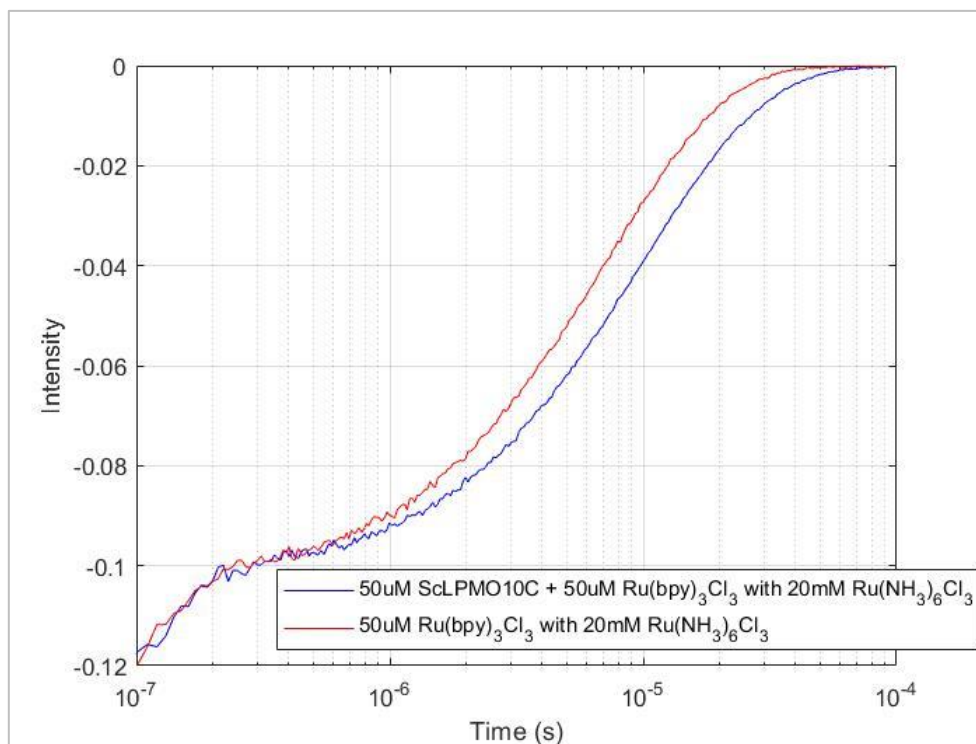


Figure 6.9. Single wavelength transient absorption spectroscopy with ScLPMO10C and Ru(NH₃)₆Cl₃ (440 nm probed).

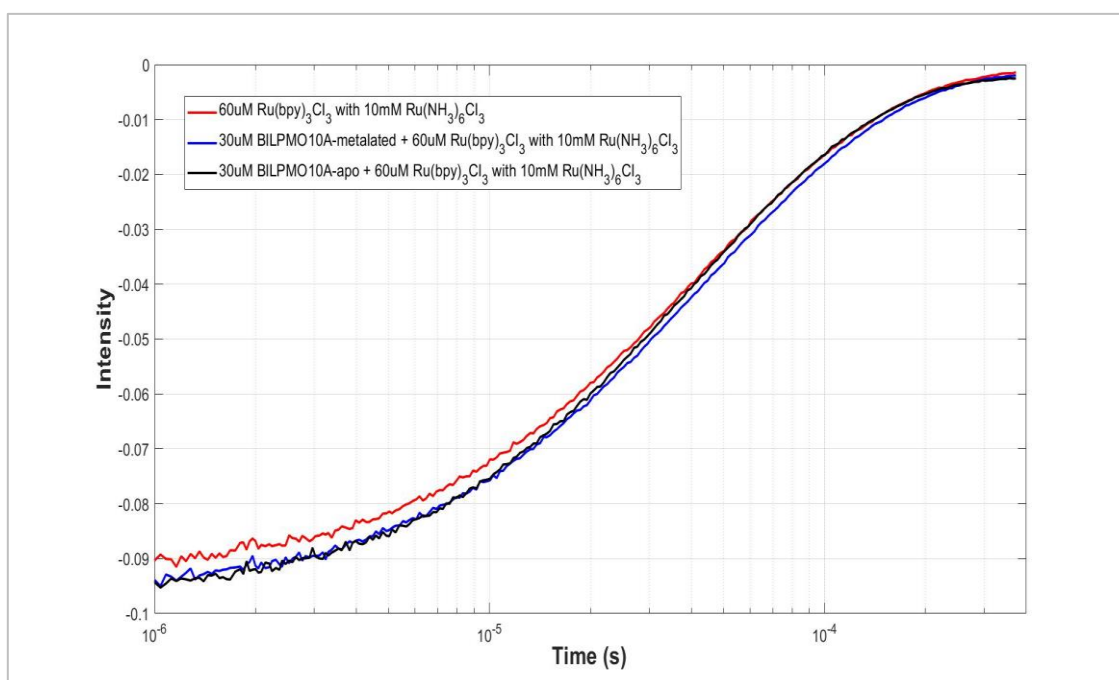


Figure 6.10. Single wavelength transient absorption spectroscopy with BILPMO10A and Ru(NH₃)₆Cl₃ (440 nm probed).

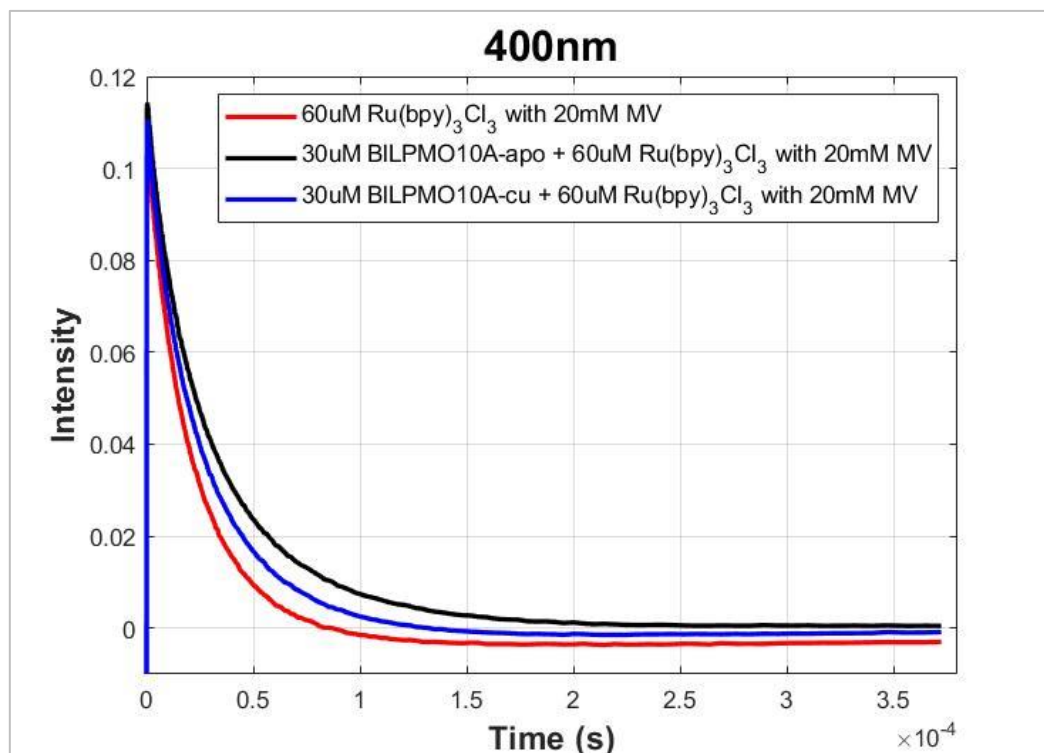


Figure 6.11. Single wavelength transient absorption spectroscopy with BILPMO10A and methyl viologen (400 nm probed).

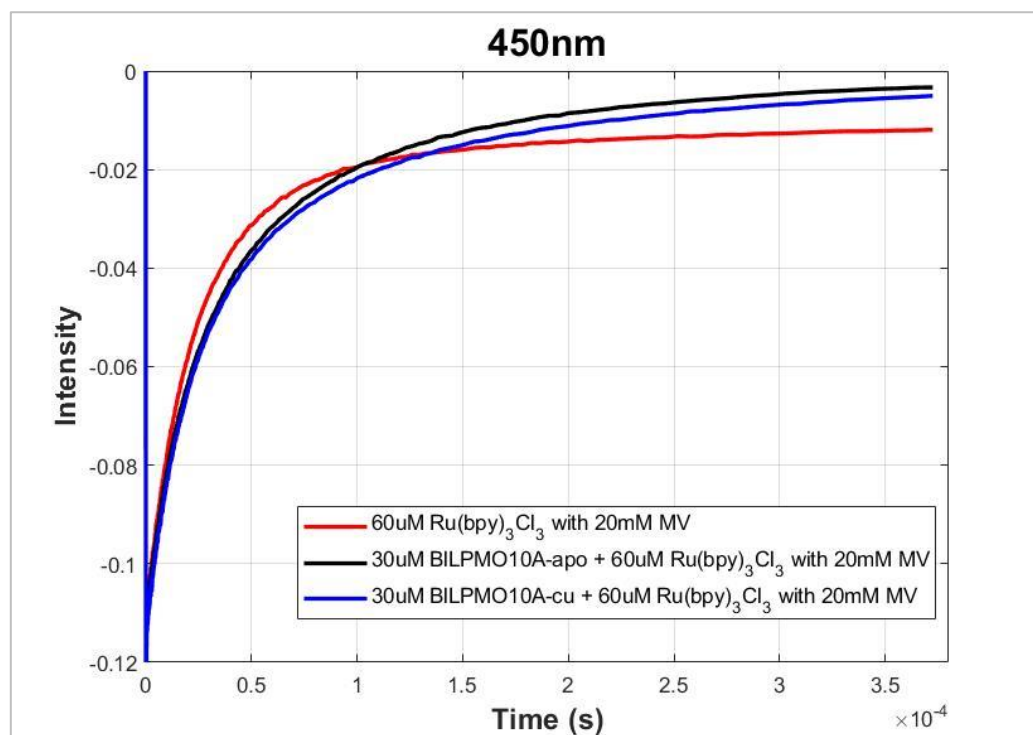


Figure 6.12. Single wavelength transient absorption spectroscopy with BILPMO10A and methyl viologen (450 nm probed).

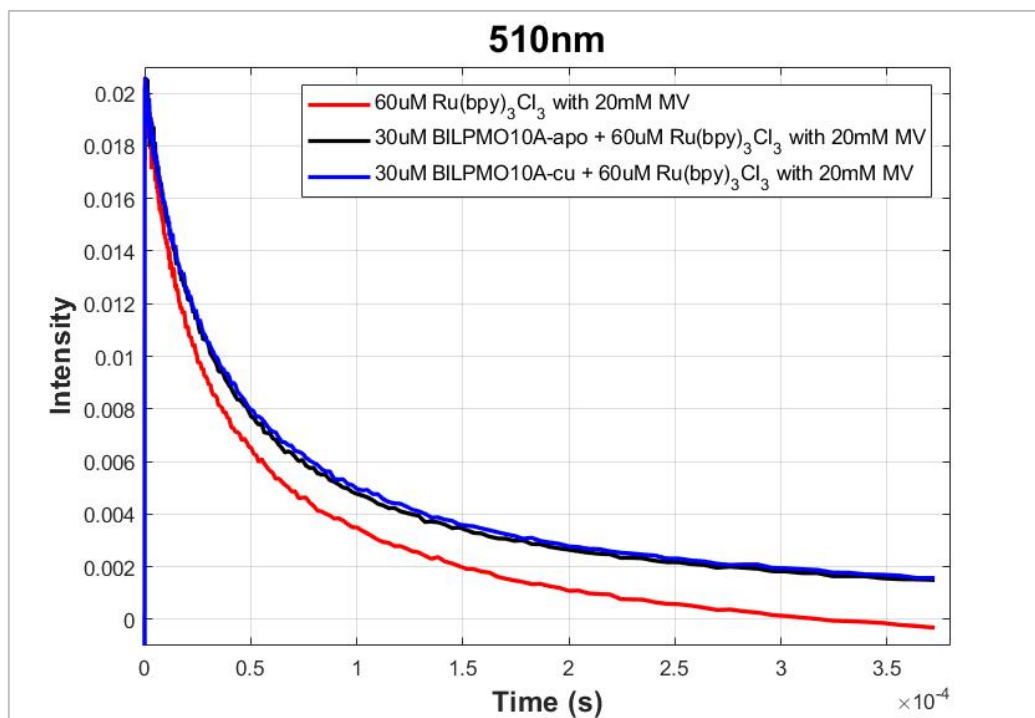


Figure 6.13. Single wavelength transient absorption spectroscopy with BILPMO10A and methyl viologen (510 nm probed).

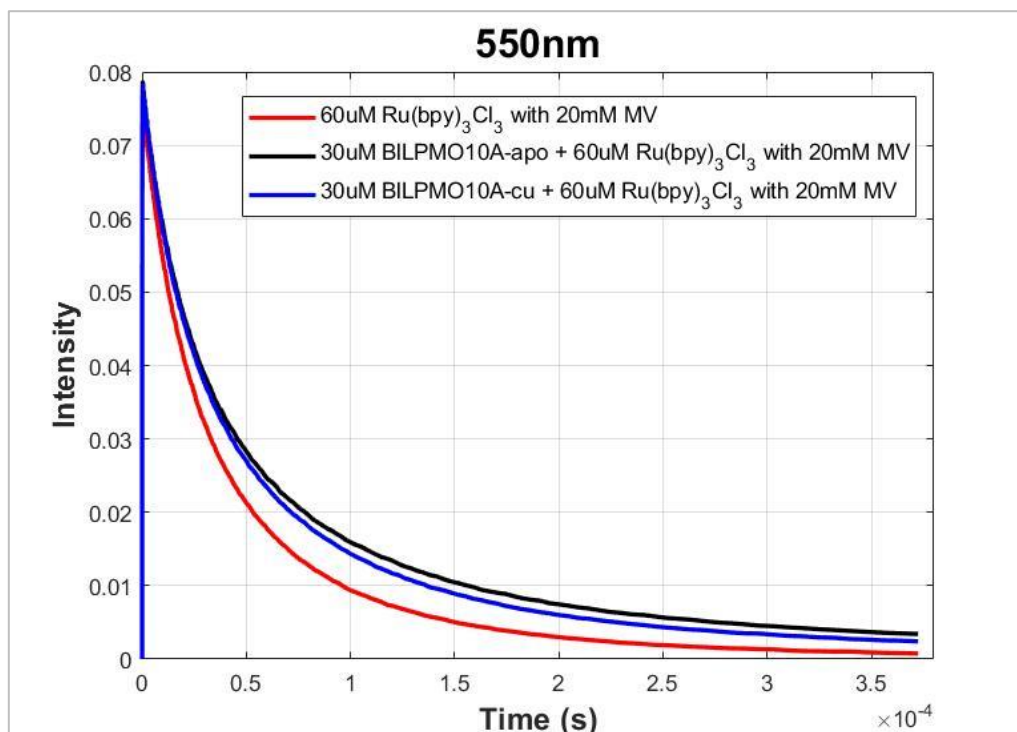


Figure 6.14. Single wavelength transient absorption spectroscopy with BILPMO10A and methyl viologen (550 nm probed).

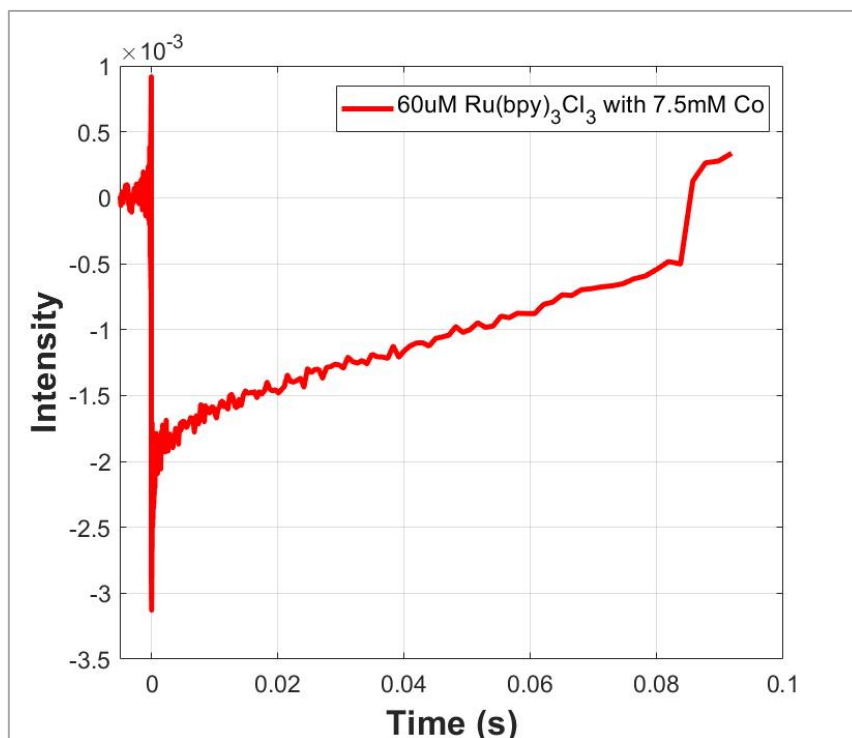


Figure 6.15. Single wavelength transient absorption spectroscopy with Ru(bpy)₃Cl₃ and [Co(NH₃)₅Cl]Cl₂ (450 nm probed).

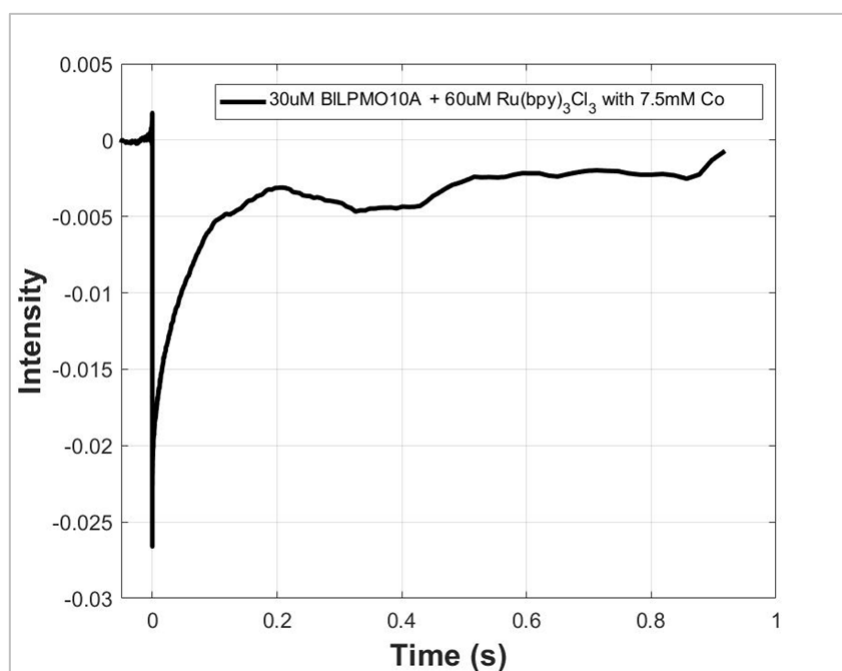


Figure 6.16. Single wavelength transient absorption spectroscopy with BILPMO10A and [Co(NH₃)₅Cl]Cl₂ (450 nm probed).

6.7. Acknowledgments

I thank Zarah Forsberg and Vincent Eijsink from Norwegian university of life sciences (NMBU) for providing me the plasmids for ScLPMO10C and BILPMO10A and for all the helpful discussions which helped me to successfully express and purify active enzymes with better yields.

6.8. References:

- [1] K. E. H. Frandsen *et al.*, “The molecular basis of polysaccharide cleavage by lytic polysaccharide monooxygenases,” *Nat. Chem. Biol.*, vol. 12, no. 4, pp. 298–303, Apr. 2016, doi: 10.1038/nchembio.2029.
- [2] B. Bissaro *et al.*, “Oxidative cleavage of polysaccharides by monocopper enzymes depends on H₂O₂,” *Nat. Chem. Biol.*, vol. 13, no. 10, pp. 1123–1128, Oct. 2017, doi: 10.1038/nchembio.2470.
- [3] Z. Forsberg *et al.*, “Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 23, pp. 8446–8451, Jun. 2014, doi: 10.1073/pnas.1402771111.
- [4] G. Courtade *et al.*, “Mechanistic basis of substrate-O₂ coupling within a chitin-active lytic polysaccharide monooxygenase: An integrated NMR/EPR study,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 117, no. 32, pp. 19178–19189, Aug. 2020, doi: 10.1073/pnas.2004277117.
- [5] S. V. Valenzuela, G. Ferreres, G. Margalef, and F. I. J. Pastor, “Fast purification

- method of functional LPMOs from *Streptomyces ambofaciens* by affinity adsorption,” *Carbohydr. Res.*, vol. 448, pp. 205–211, Aug. 2017, doi: 10.1016/j.carres.2017.02.004.
- [6] Z. Forsberg *et al.*, “Comparative study of two chitin-active and two cellulose-active AA10-type lytic polysaccharide monooxygenases,” *Biochemistry*, vol. 53, no. 10, pp. 1647–1656, Mar. 2014, doi: 10.1021/bi5000433.
- [7] N. Sutin and C. Creutz, “Properties and Reactivities of the Luminescent Excited States of Polypyridine Complexes of Ruthenium(II) and Osmium(II),” 1978, pp. 1–27.
- [8] D. Paul Rillema, G. Allen, T. J. Meyer, and D. Conrad, “Redox Properties of Ruthenium(II) Tris Chelate Complexes Containing the Ligands 2,2'-Bipyrazine, 2,2'-Bipyridine, and 2,2'-Bipyrimidine,” *Inorg. Chem.*, vol. 22, no. 11, pp. 1617–1622, 1983, doi: 10.1021/ic00153a012.

Chapter 7

Occurrence of Amino Acid Pairs and Clusters in Biological Systems

7.1. Motivation

It is notable that there exists a tyrosine cluster which consists of three tyrosine residues (Y27, Y28 and Y69) within 5Å from each other on the surface of CotA laccase. This cluster is located at the end of the chain of redox active amino acids (Trp and Tyr) stretching from the trinuclear copper cluster (TNC) where dioxygen reduction occurs to produce water. In addition to the potential roles of these chains providing protection to the enzyme (see Chapter 4 for more details), it is plausible that this surface tyrosine cluster in CotA laccase may contribute to facilitating the enzyme function of bilirubin oxidation. In efforts to elucidate biological roles of this tyrosine cluster, bioinformatics approaches were taken to examine the occurrence of amino acid pairs and clusters in proteins.

7.2. Surface Vicinal Tyrosines

Among 92099 proteins whose X-ray crystallographic structures are available in the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB, www.rcsb.org), 44163 proteins have one or more vicinal tyrosines which are less than 5 Å distance apart on the surface of the protein. Among those 44163 proteins which have at least one surface vicinal tyrosines, 27476 proteins belong to enzyme subclasses categorized into 7 different classes: 1. Oxidoreductases, 2. Transferases, 3. Hydrolases, 4. Lyases, 5. Isomerases, 6. Ligases, 7. Translocases.

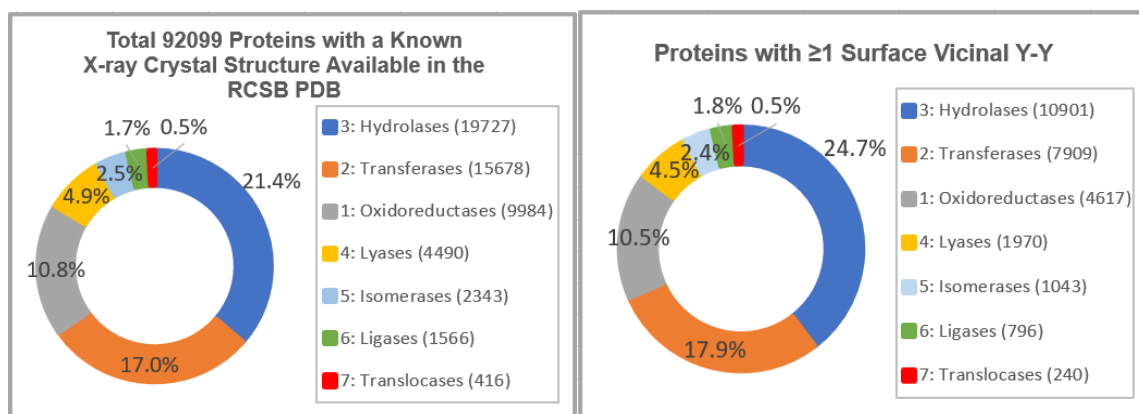


Figure 7.1. Distribution of proteins (left) and proteins with more than or equal to one surface vicinal Y-Y (right) in different protein classes.

Table 7.1. Surface vicinal Y-Y.

	Classification of 92099 proteins into 7 Different Categories	Percentage Distribution	Classification of 44163 Surface Vicinal Y-Y Containing Proteins into 7 Different Categories	Percentage Distribution	Percentage of Surface Vicinal Y-Y Containing Proteins in Each Category Relative to the Total Protein Number in Each Category
	3: Hydrolases (19727)	21.4%	3: Hydrolases (10901)	24.7%	55.3%
	2: Transferases (15678)	17.0%	2: Transferases (7909)	17.9%	50.4%
	1: Oxidoreductases (9984)	10.8%	1: Oxidoreductases (4617)	10.5%	46.2%
	4: Lyases (4490)	4.9%	4: Lyases (1970)	4.5%	43.9%
	5: Isomerases (2343)	2.5%	5: Isomerases (1043)	2.4%	44.5%
	6: Ligases (1566)	1.7%	6: Ligases (796)	1.8%	50.8%
	7: Translocases (416)	0.5%	7: Translocases (240)	0.5%	57.7%
Total	54204		27476		

In Table 7.1, column 2 represents the distribution of 54204 proteins out of 92099 proteins examined in seven different protein classes defined above. Column 3 shows the percentage distribution of those 54204 proteins in each protein class (percentage calculations were performed out of total 92099 proteins). 44163 out of 92099 proteins have one or more surface vicinal tyrosine pairs within 5 Å distance. Among those 44163 proteins, 27476 proteins belong to 7 protein classes defined above (column 4). Column 5 represents the percentage distribution of those 27476 proteins in each protein class (percentage calculations were performed out of

44163). Column 6 shows the percentage of each protein class having at least one vicinal Y-Y pair. For each class, the number of proteins having at least one vicinal Y-Y pair on the surface was divided by the total number of proteins in each protein class (e.g., for row 2, (number of surface vicinal Y-Y containing proteins that belong to hydrolases (10901) divided by the total number of proteins in the entire protein data bank that belong to hydrolases (19727)) x 100).

Table 7.2. Distribution of surface Y-Y containing proteins in different sub-categories.

Hydrolases	Transferases	Oxidoreductases
3.4: Acting on peptide bo ... (3979)	2.7: Transferring phospho ... (4176)	1.14: Acting on paired do ... (1083)
3.1: Acting on ester bonds (2706)	2.4: Glycosyltransferases (1036)	1.1: Acting on the CH-OH ... (754)
3.2: Glycosylases (2630)	2.3: Acyltransferases (1001)	1.11: Acting on a peroxid ... (312)
3.5: Acting on carbon-nit ... (973)	2.1: Transferring one-car ... (763)	1.2: Acting on the aldehy ... (287)
3.6: Acting on acid anhyd ... (597)	2.5: Transferring alkyl o ... (584)	1.3: Acting on the CH-CH ... (281)
2.7: Transferring phospho ... (173)	2.6: Transferring nitroge ... (243)	1.7: Acting on other nitr ... (264)
1.17: Acting on CH or CH(... (40)	3.1: Acting on ester bonds (145)	1.8: Acting on a sulfur g ... (260)
Other (354)	Other (527)	Other (1495)

Surface vicinal Y-Y pairs seem to be prevalent in all enzyme classes as shown in the last column of Table 7.1. Around 45 – 55 % of proteins in all enzyme classes have one or more vicinal Y-Y residues on the surface. Therefore, it was worth looking at sub-category distributions more in detail (Table 7.2.). Moreover, the actual number of surface vicinal Y-Y residues as well as the percentage distribution in different enzyme classes was examined.

Table 7.3. Surface vicinal W-W.

	Classification of 92099 proteins into 7 Different Categories	Percentage Distribution	Classification of 16798 Surface Vicinal W-W Containing Proteins into 7 Different Categories	Percentage Distribution	Percentage of Surface Vicinal W-W Containing Proteins in Each Category Relative to the Total Protein Number in Each Category
	3: Hydrolases (19727)	21.4%	3: Hydrolases (4516)	26.9%	22.9%
	2: Transferases (15678)	17.0%	2: Transferases (2640)	15.7%	16.8%
	1: Oxidoreductases (9984)	10.8%	1: Oxidoreductases (2579)	15.4%	25.8%
	4: Lyases (4490)	4.9%	4: Lyases (839)	5.0%	18.7%
	5: Isomerases (2343)	2.5%	5: Isomerases (302)	1.8%	12.9%
	6: Ligases (1566)	1.7%	6: Ligases (266)	1.6%	17.0%
	7: Translocases (416)	0.5%	7: Translocases (78)	0.5%	18.8%
Total	54204		11220		

The data in Table 7.3 are organized in the same way as in Table 7.1. It can be seen that surface vicinal W-W residues are most prevalent in hydrolases and oxidoreductases. As shown in the last column of Table 7.3, 22.9% of all hydrolases and 25.8% of all oxidoreductases have one or more vicinal W-W residues on the surface.

Table 7.4. Surface vicinal Y-W.

	Classification of 92099 proteins into 7 Different Categories	Percentage Distribution	Classification of 40674 Surface Vicinal Y-W Containing Proteins into 7 Different Categories	Percentage Distribution	Percentage of Surface Vicinal Y-W Containing Proteins in Each Category Relative to the Total Protein Number in Each Category
	3: Hydrolases (19727)	21.4%	3: Hydrolases (11476)	28.2%	58.2%
	2: Transferases (15678)	17.0%	2: Transferases (7070)	17.4%	45.1%
	1: Oxidoreductases (9984)	10.8%	1: Oxidoreductases (4354)	10.7%	43.6%
	4: Lyases (4490)	4.9%	4: Lyases (1494)	3.7%	33.3%
	5: Isomerases (2343)	2.5%	5: Isomerases (769)	1.9%	32.8%
	6: Ligases (1566)	1.7%	6: Ligases (639)	1.6%	40.8%
	7: Translocases (416)	0.5%	7: Translocases (193)	0.5%	46.4%
Total	54204		25995		

The data in Table 7.4 are organized in the same way as in Table 7.1. Surface vicinal Y-W residues are most prevalent in hydrolases, but prevalent in all the other enzyme classes as well.

As shown in the last column of Table 7.4, 58.2% of all hydrolases and around 30 - 45% of proteins in other enzyme classes have one or more vicinal Y-W residues on the surface.

Table 7.5. Surface vicinal H-H.

	Classification of 92099 proteins into 7 Different Categories	Percentage Distribution	Classification of 22348 Surface Vicinal H-H Containing Proteins into 7 Different Categories	Percentage Distribution	Percentage of Surface Vicinal H-H Containing Proteins in Each Category Relative to the Total Protein Number in Each Category
	3: Hydrolases (19727)	21.4%	3: Hydrolases (4521)	20.2%	22.9%
	2: Transferases (15678)	17.0%	2: Transferases (4430)	19.8%	28.3%
	1: Oxidoreductases (9984)	10.8%	1: Oxidoreductases (3124)	14.0%	31.3%
	4: Lyases (4490)	4.9%	4: Lyases (1304)	5.8%	29.0%
	5: Isomerases (2343)	2.5%	5: Isomerases (617)	2.8%	26.3%
	6: Ligases (1566)	1.7%	6: Ligases (389)	1.7%	24.8%
	7: Translocases (416)	0.5%	7: Translocases (168)	0.8%	40.4%
Total	54204		14553		

The data in Table 7.5 are organized in the same way as in Table 7.1. Surface vicinal H-H residues are most prevalent in translocases, but prevalent in all the other enzyme classes as well. As shown in the last column of Table 7.5, 40.4% of all translocases and around 20 – 30% of proteins in other enzyme classes have one or more vicinal H-H residues on the surface.

Table 7.6. Surface vicinal F-F.

	Classification of 92099 proteins into 7 Different Categories	Percentage Distribution	Classification of 47755 Surface Vicinal F-F Containing Proteins into 7 Different Categories	Percentage Distribution	Percentage of Surface Vicinal F-F Containing Proteins in Each Category Relative to the Total Protein Number in Each Category
	3: Hydrolases (19727)	21.4%	3: Hydrolases (9619)	20.1%	48.8%
	2: Transferases (15678)	17.0%	2: Transferases (8923)	18.7%	56.9%
	1: Oxidoreductases (9984)	10.8%	1: Oxidoreductases (5760)	12.1%	57.7%
	4: Lyases (4490)	4.9%	4: Lyases (2559)	5.4%	57.0%
	5: Isomerases (2343)	2.5%	5: Isomerases (1244)	2.6%	53.1%
	6: Ligases (1566)	1.7%	6: Ligases (886)	1.9%	56.6%
	7: Translocases (416)	0.5%	7: Translocases (300)	0.6%	72.1%
Total	54204		29291		

The data in Table 7.6 are organized in the same way as in Table 7.1. Surface vicinal F-F residues are most prevalent in translocases, but are prevalent in all the other enzyme classes as well. As shown in the last column of Table 7.6, 72.1% of all translocases and around 49 - 57% of proteins in other enzyme classes have one or more vicinal F-F residues on the surface.

Table 7.7. Summary.

Protein Classification	Frequency Distribution (%)				
	Y-Y	W-W	Y-W	H-H	F-F
Hydrolases	55.3	22.9	58.2	22.9	48.8
Transferases	50.4	16.8	45.1	28.3	56.9
Oxidoreductases	46.2	25.8	43.6	31.3	57.7
Lyases	43.9	18.7	33.3	29.0	57.0
Isomerases	44.5	12.9	32.8	26.3	53.1
Ligases	50.8	17.0	40.8	24.8	56.6
Translocases	57.7	18.8	46.4	40.4	72.1

7.3. Expected and Observed Amino Acid Distribution

Table 7.8. Expected and observed amino acid distribution.

	Surface	Total	Surface Percentage		Expected Frequency (%)
Y	1187717	1698832	69.9%	Y	2.91
W	468204	873002	53.6%	W	1.50
H	887210	1421707	62.4%	F	2.33
F	1032966	2465044	41.9%	H	2.87

The numbers of each single aromatic amino acid distributed both on the surface and in the interior protein were examined to account for the relative frequencies of each amino acid

occurring on the surface. The total number of each amino acid in column 3 of Table 7.8 indicates the sum of surface and interior residues. (e.g., 69.9 % of all tyrosines observed in the list of 92099 proteins exist on the surface.) It can be seen from Table 7.8 that all four amino acids examined (Y, W, H, F) occur prevalently on the surface. The expected frequency in the last column of Table 7.8. was calculated for each amino acid by considering the sum of the expected frequency of each codon making up each amino acid. The expected frequency of each codon was calculated by taking into account the frequencies of DNA bases in nature: 22.0% Uracil, 30.3% Adenine, 21.7% Cytosine, and 26.1% Guanine.

Table 7.9. PDB frequency vs. Total frequency.

Surface Amino Acid Pair (5A)	PDB Frequency	PDB Frequency (%)	Total #X-X Frequency
Y-Y	44163	48.0	131012.0
Y-W	40674	44.2	40674.0
W-W	16798	18.2	16798.0
H-H	22348	24.3	22348.0
F-F	47755	51.9	47755.0

The PDB frequency shown in column 2 of Table 7.9 indicates the number of proteins among 92099 proteins that have one or more of each vicinal aromatic amino acid pair on the surface. The last column in Table 7.9 shows the actual total number of each surface vicinal aromatic amino acid pair. The mismatch between the numbers in column 2 and column 4 only for the Y-Y pair indicates that there are many proteins that have multiple surface vicinal Y-Y residues. On the other hand, it is clear that other vicinal aromatic amino acid pairs exhibit the characteristics of only a single occurrence per each protein.

Table 7.10. Percentage of aromatic residue pairs.

% Y-Y/Surface Y	22.1%	% Y-Y/Total Y	15.4%
% W-W/Surface W	7.2%	% W-W/Total W	3.8%
% H-H/Surface H	5.0%	% H-H/Total H	3.1%
% F-F/Surface F	9.2%	% F-F/Total F	3.9%

The total numbers of vicinal aromatic residues including the ones in the interior structure of the proteins were examined. 22.1 % of all surface tyrosines exist as a pair within 5 Å distance, which is almost 3-4 times more than it is for other aromatic residue pairs. 15.4 % of all tyrosines found in 92099 proteins with known structures exist as a pair within 5 Å distance, which is almost 4-5 times more than it is for other aromatic residue pairs (Table 7.10). This indicates that the probability of tyrosines existing on the surface as a pair within 5 Å distance is much higher than for other aromatic amino acids.

Table 7.11. Distribution of different numbers of Y-Y pairs in proteins.

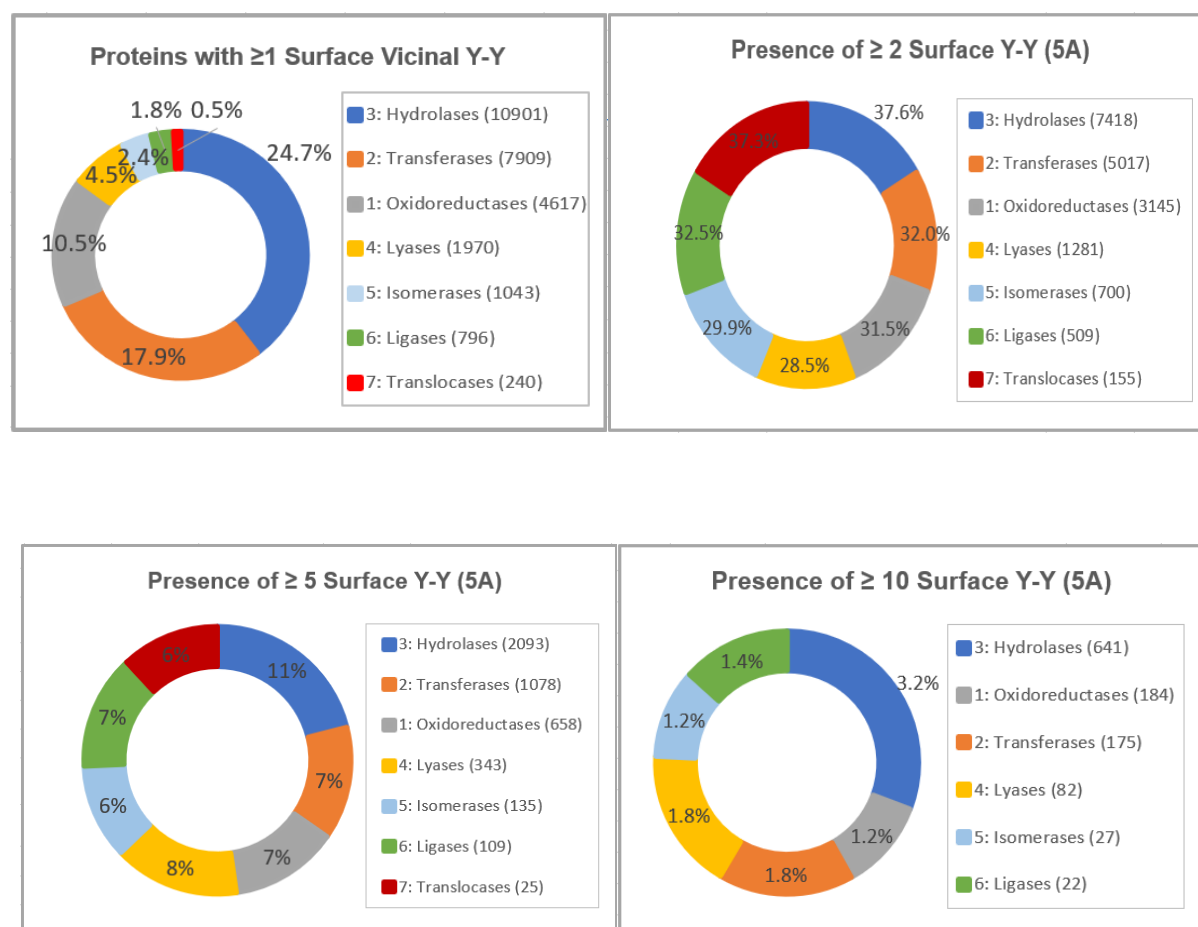
Number of Surface Y-Y	Number of Proteins
77	1
≥ 50	6
≥ 40	9
≥ 30	29
≥ 25	60
≥ 20	155
≥ 15	422
≥ 10	1707
≥ 5	7241
≥ 2	29117

The actual number of surface vicinal Y-Y pairs was examined more in detail. The highest number of surface vicinal Y-Y pairs from a single protein was found to be 77. As can be seen from Table 7.11 that six proteins among 92099 proteins examined have more than 50 pairs of vicinal Y-Y on the surface, and 9 proteins have more than 40 pairs of vicinal Y-Y on the surface.

Table 7.12. Identification of proteins with the highest number of surface Y-Y pairs.

	$40 \leq \text{Surface Y-Y} \leq 50$	$51 \leq \text{Surface Y-Y} \leq 60$	$61 \leq \text{Surface Y-Y} \leq 70$	$71 \leq \text{Surface Y-Y} \leq 80$
PDB	1IJG	1JOB		4GPK
	2GFB	2R9G		4HMK
	4APZ	4GM2		4XGZ

PDB codes of proteins having $40 \leq$ surface Y-Y pairs are shown in Table 7.12.



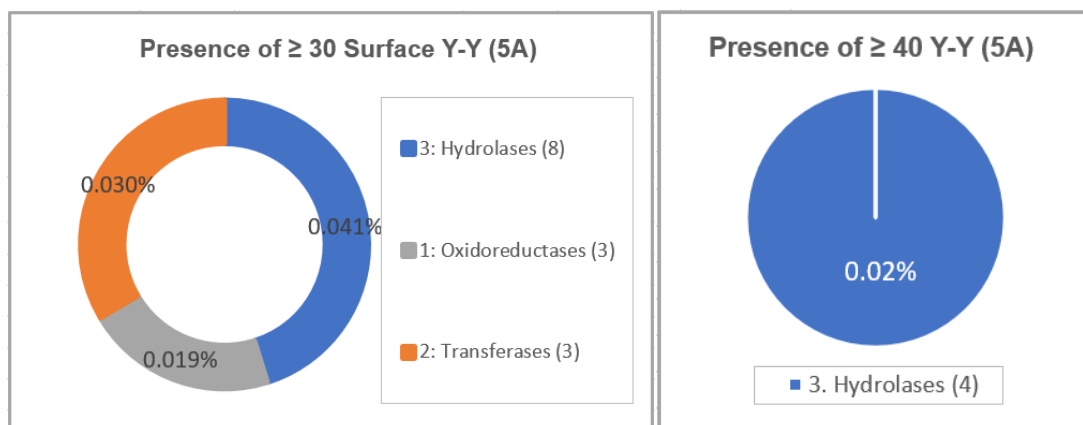


Figure 7.2. Distribution of proteins with different numbers of surface vicinal Y-Y in different protein classes.

It can be seen from Figure 7.2 above that the occurrence of multiple surface Y-Y pairs with higher frequencies is more prevalent in hydrolases. As the number of surface vicinal Y-Y pairs in a protein increases, the higher proportion of those proteins are found to be hydrolases.

Table 7.13. Percentage of surface vicinal aromatic amino acid pairs.

	Total X-X Frequency	Surface X-X Frequency	Percentage of Surface X-X
Y-Y	248061	131012	52.8%
W-W	120405	16798	14.0%
F-F	785668	47755	6.1%
H-H	159271	22348	14.0%

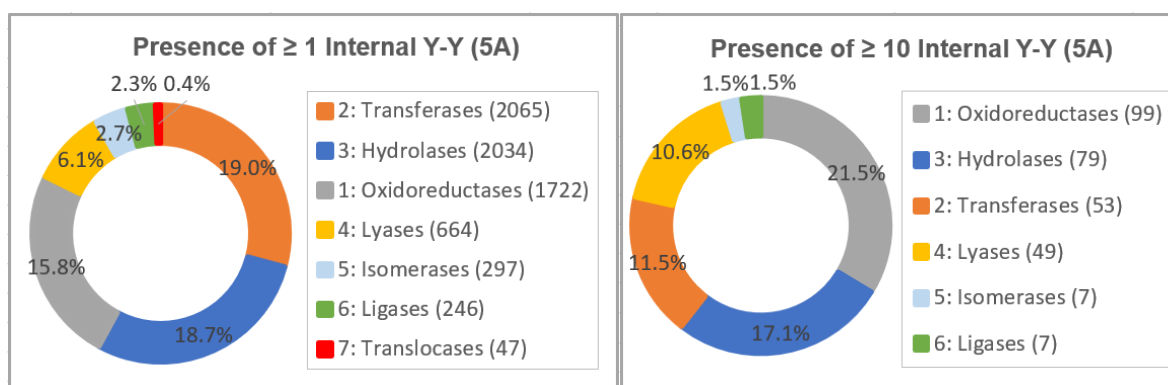
The number of each vicinal aromatic amino acid pair (Y-Y, W-W, F-F, H-H) was examined not only for the ones on the surface but also for the entire protein structure including the ones in the interior compartment. It is shown in Table 7.13, more than 50% of the vicinal Y-Y pairs exist on the surface whereas only around 6-14% of other vicinal aromatic amino acid pairs exist on the surface. Therefore, it is of interest to investigate the possibility of a biological significance of surface vicinal tyrosines.

7.4. Distribution of Internal Vicinal Y-Y in Different Enzyme Classes

To elucidate the roles of surface vicinal Y-Y pairs, distribution of internal vicinal Y-Y pairs in different enzyme classes was also examined to see if there are any differences with the observations made for the surface vicinal Y-Y pairs.

Table 7.14. Number of proteins having different numbers of internal Y-Y without any surface Y-Y [Column 1: (Total # of All Y-Y Pairs searched) – (# of Surface Y-Y Pairs)].

Number of Internal Y-Y For Proteins with No Surface Y-Y	Number of Proteins
72	1
≥ 50	3
≥ 40	10
≥ 30	13
≥ 20	87
≥ 10	461
≥ 1	10866



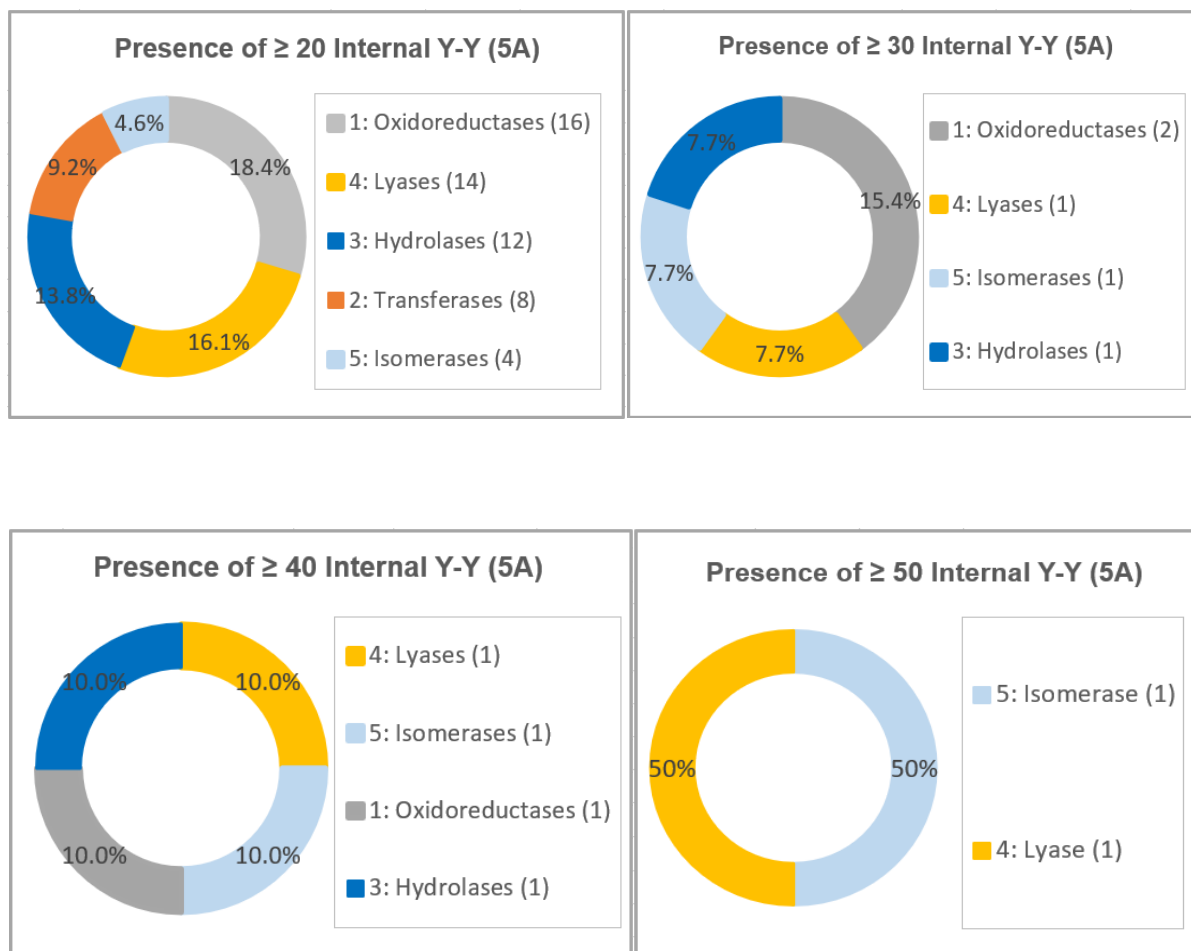


Figure 7.3. Distribution of different numbers of internal Y-Y in different enzyme classes.

It can be seen from Figure 7.3 that the occurrence of multiple internal Y-Y pairs is more prevalent in lyases, oxidoreductases, and isomerases. (As the number of internal vicinal Y-Y pairs in a protein increases, the higher proportion of those proteins are found to be isomerases, lyases, and oxidoreductases.) This trend is the opposite from the observations made for multiple surface vicinal Y-Y pairs which occur mostly in hydrolases and transferases (Figure 7.2).

7.5. More Refined List of PDB Codes

To minimize the redundancy problem of having too many similar structures, search results are to be filtered in such a way that multiple structures whose sequences have the specified level

(e.g., 90% or 95%) of sequence identity will be represented as a single structure. This way, overcounting of duplicates such as mutant structures, protein fragments or the structures with different redox states of the active sites can be prevented. It is also necessary to consider the structures of biological relevance (biological assemblies) which will be discussed in part 7.6.

7.6. Biological Assemblies

Biological assemblies are the functional forms of the enzymes. The list of PDBs for biological assemblies were downloaded from the RCSB PDB website, and among different biological assemblies existing for each asymmetric unit of the same protein crystal structure, the first biological assembly archived as pdb1 was used for this study. Around 32000 biological assemblies have been searched to explore the amino acid propensities of existing as a pair or in a cluster. For the X-X pairs, the cutoff distance between the two amino acid residues was set to 5Å, and the cluster was defined as more than two identical amino acids located in close proximity (5Å) to each other.

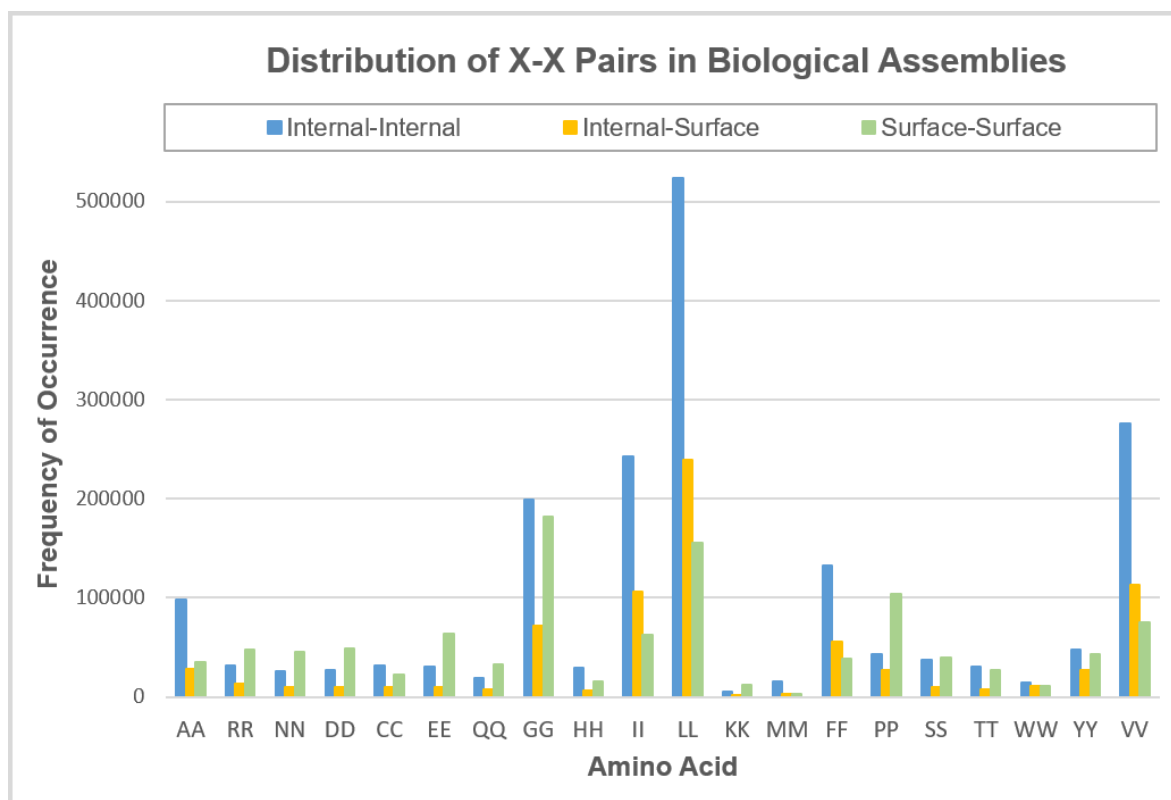


Figure 7.4. Distribution of X-X amino acid pairs in biological assemblies.

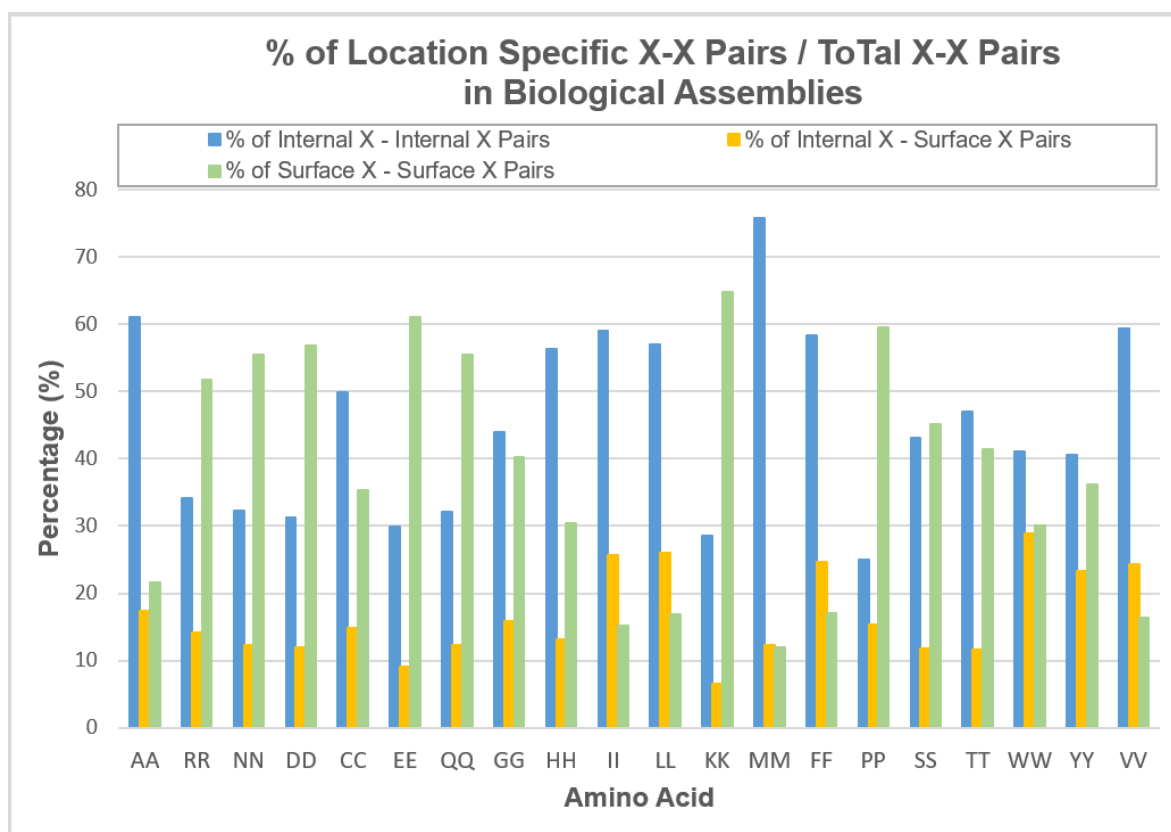


Figure 7.5. Relative numbers in % of location specific vicinal X-X in biological assemblies.

Figure 7.4 shows the frequency of each amino acid pair (X-X) in different locations of 32045 biological assemblies with different PDBs, and Figure 7.5 shows the relative numbers in % of those X-X pairs found in different locations. These two figures represent the general distribution of each amino acid pairs in different biological assemblies.

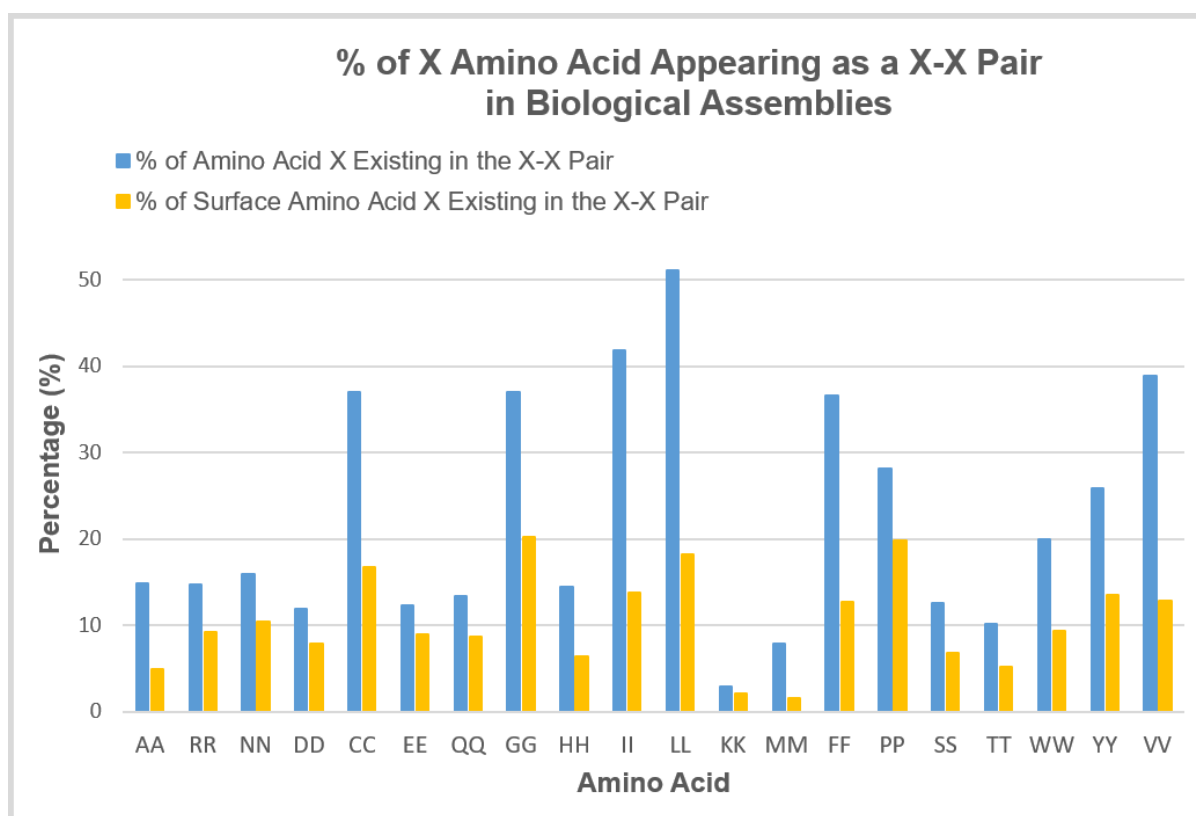


Figure 7.6. Percentage of amino acid X appearing as a vicinal X-X in biological assemblies.

Since the natural frequency of each amino acid occurring in proteins vary greatly (Figure 7.4) and some amino acids are involved in multiple pairs forming a cluster, the number of a unique set of amino acid residues forming a pair within a protein (and also specifically on the surface) was examined for each amino acid (X). In this search, it was made sure that each amino acid residue was counted only once to figure out the actual percentage of amino acid residues

appearing as pairs. Compared to the previous search which included non-biological assemblies (Part 7.2 – 7.5), it can be seen that the trend of higher occurrence frequency for Y-Y pairs disappears (Figure 7.6).

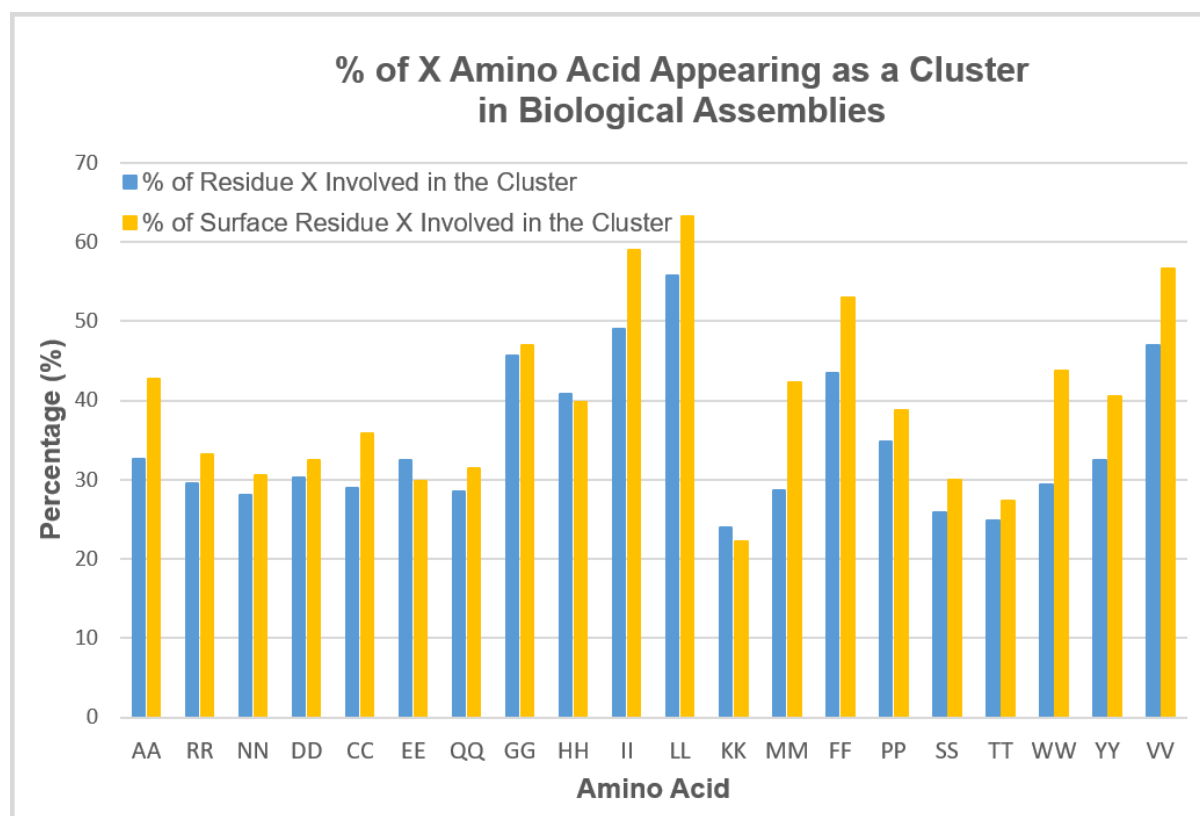


Figure 7.7. Percentage of amino acid X appearing as a cluster in biological assemblies.

When the residues existing as a cluster (with more than two identical amino acids separated by $5 \geq \text{Å}$ from each other) were examined, the trend was very similar to the pairs (Figure 7.6 and 7.7) except that cysteines and methionines exhibit opposite trends. Cysteines have a greater tendency to exist as pairs than to form clusters, presumably due to the biological necessity for disulfide backbone formation with Cys residues. Methionines, on the other hand, have a greater tendency to exist as clusters with more than two residues in close proximity than as pairs, since they are frequently found in helical loops on the protein surface.

7.7. Conclusion

Amino acids exist as pairs and clusters in proteins. To elucidate biological roles and the potential significance of these pairs and clusters, the occurrence frequencies, location and their distribution in different protein classes were examined. It is interesting that the prevalence of surface Y-Y pairs observed when all PDB structures available from RCSB PDB website were examined is no longer a prominent feature observed in biological assemblies. It implies that surface Y-Y pairs are indeed frequently observed in proteins that have been extensively studied and characterized by crystallography, but with the refined set of structures with the redundancy eliminated, it does not seem like proteins have evolved to have more surface Y-Y pairs for particular functional purposes.

Although surface Y-Y pairs are commonly observed in proteins with known crystal structures, their biological significance in the functional forms of proteins does not seem prominent. Thus, it would be interesting to perform more investigations by narrowing down the search categories to specific types of proteins (e.g., iron proteins, copper proteins etc.) to be able to draw more precise conclusions about their functional roles.

7.8. Appendices

The MATLAB search code was modified from an original W-W and Y-Y pair search codes wrote by Dr. Jay Winkler to search for all residue pairs as well as the residue clusters (>2 residues together).

MATLAB Function Code

```
Data = dir('*.*txt');
for i = 1:length(Data)
    Data(i).Data = readcell(Data(i).name);
end

bioassembly = Data.Data(:,1);
bioassembly_pdb = char(bioassembly);

%for i = 1:236172
for i = 1:236172
    if bioassembly_pdb(i,9)=='1'

biological_assembly_JS_allAA_struct('C:\Users\Jieun Shin\Desktop\bioinf', bioassembly_pdb(i, 1:4),
0, 5)
    end
end
```

MATLAB Search Code

```
function [AA_pairs, chains, models, resd_tot] =
biological_assembly_JS_allAA_struct(pname,pdb_id,
cutoff, cutoff2)
%
%f
AA_pairs=zeros(20,6);    % OUTPUT RESULT
%
try
    %
```

```

    savefile=['C:\Users\Jieun
Shin\Desktop\bioinf\',pdb_id, '.pdb1.gz'];

getfile=['https://files.rcsb.org/download/',pdb_id
, '.pdb1.gz'];
    outfile=websave(savefile,getfile);
    outfile_unzip=char(gunzip(outfile));
    if ~isempty(outfile_unzip)
        pdbstruct=pdbread(outfile_unzip);
        if ~isempty(pdbstruct)
            delete(outfile)
            delete(outfile_unzip)
        end
    else
        pdbstruct=getpdb(pdb_id);
    end

%     %savefile=['C:\Users\Jieun
Shin\Desktop\bioinf\',pdb_id, '.pdb1.gz'];
%     savefile='C:\Users\Jieun
Shin\Desktop\bioinf\';
%     %getpdb('6OWE', 'ToFile', '6OWE.pdb');
%     ftpobj = ftp('ftp.wwpdb.org');
%
mget(ftpobj,['pub/pdb/data/biounit/coordinates/all
/', pdb_id, '.pdb1.gz'], savefile);
%     %outfile = fileread(savefile);
%     outfile_unzip = char(gunzip([savefile,
'pub/pdb/data/biounit/coordinates/all/', pdb_id,
'.pdb1.gz']));
%
%     %getfile=['https://files.rcsb.org/download/'
,pdb_id, '.pdb1.gz'];
%     %outfile=websave(savefile,getfile);
%     %outfile=websave(savefile,getpdb);
%     %outfile_unzip=gunzip(outfile);
%     %outfile_unzip=char(gunzip(outfile));
%     if ~isempty(outfile_unzip)
%         pdbstruct=pdbread(outfile_unzip);
%         if ~isempty(pdbstruct)

```

```

%           %delete(outfile)
%           delete(outfile_unzip)
%       end
%   else
%       pdbstruct=getpdb(pdb_id);
%   end
catch
    fprintf(' pdbstruct failure')
    AA_pairs(1,1) = -5;    % -5 is failure code for
getting pdbstruct from pdb
                                % will try 3 times
before giving up
    fclose('all');
    return
end
%
% Find the number of models
%
models=length(pdbstruct.Model);
%
% Find the number and name of the chains
%
chain_str=[];
chain_str=[chain_str,pdbstruct.Sequence.ChainID]';
%character array of chain names
num_chains=length(chain_str);
%
resd_tot=0;
for ij=1:num_chains

resd_tot=resd_tot+length(strfind(pdbstruct.Sequence
e(ij).Sequence,'A'));

resd_tot=resd_tot+length(strfind(pdbstruct.Sequence
e(ij).Sequence,'R'));

resd_tot=resd_tot+length(strfind(pdbstruct.Sequence
e(ij).Sequence,'N'));

resd_tot=resd_tot+length(strfind(pdbstruct.Sequence

```

```
e(ij).Sequence, 'D'));  
  
resd_tot=resd_tot+length(strfind(pdbstruct.Sequence  
e(ij).Sequence, 'C'));  
  
resd_tot=resd_tot+length(strfind(pdbstruct.Sequence  
e(ij).Sequence, 'E'));  
  
resd_tot=resd_tot+length(strfind(pdbstruct.Sequence  
e(ij).Sequence, 'Q'));  
  
resd_tot=resd_tot+length(strfind(pdbstruct.Sequence  
e(ij).Sequence, 'G'));  
  
resd_tot=resd_tot+length(strfind(pdbstruct.Sequence  
e(ij).Sequence, 'H'));  
  
resd_tot=resd_tot+length(strfind(pdbstruct.Sequence  
e(ij).Sequence, 'I'));  
  
resd_tot=resd_tot+length(strfind(pdbstruct.Sequence  
e(ij).Sequence, 'L'));  
  
resd_tot=resd_tot+length(strfind(pdbstruct.Sequence  
e(ij).Sequence, 'K'));  
  
resd_tot=resd_tot+length(strfind(pdbstruct.Sequence  
e(ij).Sequence, 'M'));  
  
resd_tot=resd_tot+length(strfind(pdbstruct.Sequence  
e(ij).Sequence, 'F'));  
  
resd_tot=resd_tot+length(strfind(pdbstruct.Sequence  
e(ij).Sequence, 'P'));  
  
resd_tot=resd_tot+length(strfind(pdbstruct.Sequence  
e(ij).Sequence, 'S'));  
  
resd_tot=resd_tot+length(strfind(pdbstruct.Sequence  
e(ij).Sequence, 'T'));
```

```

resd_tot=resd_tot+length(strfind(pdbstruct.Sequence
e(ij).Sequence, 'W'));

resd_tot=resd_tot+length(strfind(pdbstruct.Sequence
e(ij).Sequence, 'Y'));

resd_tot=resd_tot+length(strfind(pdbstruct.Sequence
e(ij).Sequence, 'V'));

end
%
%
resd_tot=resd_tot*models;
%
if resd_tot>5000
    fprintf('\r structure is too big: %i residues
\r',resd_tot);
    AA_pairs=zeros(20,6);
    %

filename=[pname, 'mat\r',pdb_id, '_too_big_', num2str(
resd_tot), '_residues.mat'];
    save(filename, 'pdb_id', 'cutoff', 'resd_tot');
%
    fclose('all');
    return
end
%
%
chains=[];
for ij=1:models
    chains=[chains, chain_str];
end
%
%
num_chains_tot=models*num_chains;
% number of chains
%
%
% A_total=zeros(20,num_chains_tot,models);

```



```

AA_total=zeros(20,num_chains_tot,models);
%total number of residues in each chain
AA=zeros(20,num_chains_tot,models);
%number of participating residues in each chain
%
%
close all;
close all hidden;
close all force;
%
% Cell array of redox chains names and indeces
of atoms
%
REDOX_CHAIN=cell(1,2);
%
% Get Chain|Residue List
%
%
rcount=0;
%
allX=[];
allY=[];
allZ=[];
%
for mod_cnt=1:models % begin the
model counting loop
%
%get all residue coordinates
%
    allX=[allX,pdbstruct.Model(mod_cnt).Atom.X];
    allY=[allY,pdbstruct.Model(mod_cnt).Atom.Y];
    allZ=[allZ,pdbstruct.Model(mod_cnt).Atom.Z];
%
end % end the model
counting loop
%
test_atom_out={};
%
for mod_cnt=1:models % begin the
model counting loop

```

```

%
atom_chain=[];
if ~isfield(pdbstruct.Model(mod_cnt), 'Atom')
    fprintf(' Field error')
    AA_pairs=zeros(20,5);      % OUTPUT RESULT
    fclose('all');
    return
end
atom_chain=[atom_chain;
pdbstruct.Model(mod_cnt).Atom.chainID]';      %char
array of chain ID's
%
atom_res_name={};
atom_res_name=[atom_res_name,pdbstruct.Model(mod_c
nt).Atom.resName];
atom_res_name=char(atom_res_name);
    %char array of residue names
%
atom_chn_res=cellstr([atom_chain,atom_res_name]);
    %cell array of ChainID|ResidueName strings
%atom_chn_res_num=strnum_eq(atom_chn_res);
%
%
% ***** Start finding all residues
*****
%
%
ala_chn_ind=cell(1,num_chains);
%cell array of ALA indeces for each chain
%
for ij=((mod_cnt-
1)*num_chains)+1:(mod_cnt)*num_chains
    %
    % find ALA residues
    %

ala_ind=0;      %
reset ALA index counter
%
```

```

chn_ind=find(strcmp([chains(ij),'ALA'],atom_chn_re
s)>0);
    %chn_ind=strfind(pdbstruct.Sequence(ij).Sequen
ce,'AA')';
%
    ala_res_nos_un=[];
    if ~isempty(chn_ind)
        res_nos=[];

res_nos=[res_nos,pdbstruct.Model(mod_cnt).Atom(chn
_ind).resSeq]'; %get residue numbers of
ChainID|ALA

ala_res_nos_un=unique(res_nos);
%find the unique residue numbers
    end
%
A_total(1,ij,mod_cnt)=length(ala_res_nos_un);

%     cmp_var = ala_res_nos_un(2:end);
%     cmp_var = [cmp_var; -1];
%     diff = cmp_var - ala_res_nos_un;
%     check_var = [find(diff==1),
find(diff==1)+1]';
%     check_var = reshape(check_var,
[size(check_var, 2)*2, 1]);
%     check_var_u = unique(check_var);
%     ala_res_nos_un =
ala_res_nos_un(check_var_u);
%

AA_total(1,ij,mod_cnt)=length(ala_res_nos_un);
%
    for
jk=1:length(ala_res_nos_un)
% test all of ALA residues

res_ind=find(res_nos==ala_res_nos_un(jk));
    %get the indeces of Chain|ALA(jk)
%

```

```

        if (length(res_ind)>4) &&
        (strcmp(pdbstruct.Model(mod_cnt).Atom(chn_ind(res_
ind(5))).AtomName, 'CB'))
            rcount=rcount+1;
            ala_ind=ala_ind+1;
            ala_chn_ind{1,ij}(ala_ind,1)=rcount;
            resd_name=['ALA-
',chains(ij),num2str(ala_res_nos_un(jk)),']'];
            %resd_name=['TYR-
',chains(ij),num2str(ala_res_nos_un(jk)),']'];
            AA(1,ij,mod_cnt)=AA(1,ij,mod_cnt)+1;
%
            tmpX=[];
            tmpY=[];
            tmpZ=[];

tmpX=[tmpX,pdbstruct.Model(mod_cnt).Atom(chn_ind(r
es_ind(5))).X];

tmpY=[tmpY,pdbstruct.Model(mod_cnt).Atom(chn_ind(r
es_ind(5))).Y];

tmpZ=[tmpZ,pdbstruct.Model(mod_cnt).Atom(chn_ind(r
es_ind(5))).Z];
            % check for surface atoms
            [ surf_tst ] =
surf_atom( [tmpX',tmpY',tmpZ'],[allX',allY',allZ'
]);
%
            if surf_tst==1
                resd_name=['*',resd_name];
            else
                resd_name=[' ',resd_name];
            end
            REDOX_CHAIN{1,1}{rcount,1}=resd_name;

REDOX_CHAIN{1,2}{rcount,1}=chn_ind(res_ind(5))';
        end
%

```

```

        end
    end
    %
    %
    arg_chn_ind=cell(1,num_chains);
    %cell array of ARG indeces for each chain
    %
    for ij=((mod_cnt-
    1)*num_chains)+1:(mod_cnt)*num_chains
        %
        %     find ARG residues
        %
        arg_ind=0; %
        reset ARG index counter
        %
        chn_ind=find(strcmp([chains(ij),'ARG'],atom_chn_re
        s)>0);
        %chn_ind=strfind(pdbstruct.Sequence(ij).Sequen
        ce,'AA')';
        %
        arg_res_nos_un=[];
        if ~isempty(chn_ind)
            res_nos=[];
        end
        res_nos=[res_nos,pdbstruct.Model(mod_cnt).Atom(chn
        _ind).resSeq]'; %get residue numbers of
        ChainID|ARG
        arg_res_nos_un=unique(res_nos);
        %find the unique residue numbers
        end
    %
    A_total(2,ij,mod_cnt)=length(arg_res_nos_un);
    %
    %     cmp_var = arg_res_nos_un(2:end);
    %     cmp_var = [cmp_var; -1];
    %     diff = cmp_var - arg_res_nos_un;
    %     check_var = [find(diff==1),

```

```

find(diff==1)+1]';
%   check_var = reshape(check_var,
[size(check_var, 2)*2, 1]);
%   check_var_u = unique(check_var);
%   arg_res_nos_un =
arg_res_nos_un(check_var_u);

AA_total(2,ij,mod_cnt)=length(arg_res_nos_un);
%
    for
jkl=1:length(arg_res_nos_un)
% test all of ARG residues

res_ind=find(res_nos==arg_res_nos_un(jkl));
%get the indeces of Chain|ARG(jkl)
%
    if (length(res_ind)>7) % 8th atom:'NE',
9th atom:'CZ', 10th atom:'NH1', 11th atom: 'NHZ')
        rcount=rcount+1;
        arg_ind=arg_ind+1;
        arg_chn_ind{1,ij}(arg_ind,1)=rcount;
        resd_name=['ARG-
',chains(ij),num2str(arg_res_nos_un(jkl)),''];
        %resd_name=['TYR-
',chains(ij),num2str(tyr_res_nos_un(jkl)),''];
        AA(2,ij,mod_cnt)=AA(2,ij,mod_cnt)+1;
%
        atms=min(11,length(res_ind));
        tmpX=[];
        tmpY=[];
        tmpZ=[];

tmpX=[tmpX,pdbstruct.Model(mod_cnt).Atom(chn_ind(r
es_ind(8:atms))).X];

tmpY=[tmpY,pdbstruct.Model(mod_cnt).Atom(chn_ind(r
es_ind(8:atms))).Y];

```



```

%
for ij=((mod_cnt-
1)*num_chains)+1:(mod_cnt)*num_chains
    %
    % find TRP residues
    %

trp_ind=0; %
reset TRP index counter
    %

chn_ind=find(strcmp([chains(ij),'TRP'],atom_chn_re
s)>0);
    %chn_ind=strfind(pdbstruct.Sequence(ij).Sequen
ce,'AA')';
%
    trp_res_nos_un=[];
    if ~isempty(chn_ind)
        res_nos=[];

res_nos=[res_nos,pdbstruct.Model(mod_cnt).Atom(chn
_ind).resSeq]'; %get residue numbers of
ChainID|TRP

trp_res_nos_un=unique(res_nos);
%find the unique residue numbers
    end

%
A_total(18,ij,mod_cnt)=length(trp_res_nos_un);
%
% cmp_var = trp_res_nos_un(2:end);
% cmp_var = [cmp_var; -1];
% diff = cmp_var - trp_res_nos_un;
% check_var = [find(diff==1),
find(diff==1)+1]';
% check_var = reshape(check_var,
[size(check_var, 2)*2, 1]);
% check_var_u = unique(check_var);
% trp_res_nos_un =
trp_res_nos_un(check_var_u);

```



```

AA_total(18,ij,mod_cnt)=length(trp_res_nos_un);
                                %
    for
    jk=1:length(trp_res_nos_un)
    % test all of TRP residues

    res_ind=find(res_nos==trp_res_nos_un(jk));
    %get the indeces of Chain|TRP(jk)
    %
        if (length(res_ind)>5)
            rcount=rcount+1;
            trp_ind=trp_ind+1;
            trp_chn_ind{1,ij}(trp_ind,1)=rcount;
            resd_name=['TRP-
',chains(ij),num2str(trp_res_nos_un(jk)),''];
            %resd_name=['TYR-
',chains(ij),num2str(tyr_res_nos_un(jk)),''];
            AA(18,ij,mod_cnt)=AA(18,ij,mod_cnt)+1;
        %

            atms=min(14,length(res_ind));
            tmpX=[];
            tmpY=[];
            tmpZ=[];

            tmpX=[tmpX,pdbstruct.Model(mod_cnt).Atom(chn_ind(r
es_ind(6:atms))).X];

            tmpY=[tmpY,pdbstruct.Model(mod_cnt).Atom(chn_ind(r
es_ind(6:atms))).Y];

            tmpZ=[tmpZ,pdbstruct.Model(mod_cnt).Atom(chn_ind(r
es_ind(6:atms))).Z];
            % check for surface atoms
            [ surf_tst ] =
surf_atom( [tmpX',tmpY',tmpZ'],[allX',allY',allZ']
);

```

```

%
    if surf_tst==1
        resd_name=['*',resd_name];
    else
        resd_name=[' ',resd_name];
    end
    REDOX_CHAIN{1,1}{rcount,1}=resd_name;

REDOX_CHAIN{1,2}{rcount,1}=chn_ind(res_ind(6:atms)
)';
    end
end
end
%
end
end
%
tyr_chn_ind=cell(1,num_chains);
%cell array of TYR indeces for each chain
%
for ij=((mod_cnt-
1)*num_chains)+1:(mod_cnt)*num_chains
    %
    % find TYR residues
    %

tyr_ind=0; %
reset TYR index counter
%

chn_ind=find(strcmp([chains(ij),'TYR'],atom_chn_re
s)>0);
    %chn_ind=strfind(pdbstruct.Sequence(ij).Sequen
ce,'AA')';
%
    tyr_res_nos_un=[];
    if ~isempty(chn_ind)
        res_nos=[];

```

```
res_nos=[res_nos,pdbstruct.Model(mod_cnt).Atom(chn
_ind).resSeq]'; %get residue numbers of
ChainID|TYR
```

```
tyr_res_nos_un=unique(res_nos);
%find the unique residue numbers
    end
%
A_total(19,ij,mod_cnt)=length(tyr_res_nos_un);
%
%     cmp_var = tyr_res_nos_un(2:end);
%     cmp_var = [cmp_var; -1];
%     diff = cmp_var - tyr_res_nos_un;
%     check_var = [find(diff==1),
find(diff==1)+1]';
%     check_var = reshape(check_var,
[size(check_var, 2)*2, 1]);
%     check_var_u = unique(check_var);
%     tyr_res_nos_un =
tyr_res_nos_un(check_var_u);
```

```
AA_total(19,ij,mod_cnt)=length(tyr_res_nos_un);
%
    for
jkl=1:length(tyr_res_nos_un)
% test all of TYR residues

res_ind=find(res_nos==tyr_res_nos_un(jkl));
    %get the indeces of Chain|TYR(jkl)
%
        if (length(res_ind)>5)
            rcount=rcount+1;
            tyr_ind=tyr_ind+1;
            tyr_chn_ind{1,ij}(tyr_ind,1)=rcount;
            resd_name=['TYR-
',chains(ij),num2str(tyr_res_nos_un(jkl)),']'];
            %resd_name=['TYR-
',chains(ij),num2str(tyr_res_nos_un(jkl)),']'];
```

```

AA(19,ij,mod_cnt)=AA(19,ij,mod_cnt)+1;
%

atms=min(12,length(res_ind));
tmpX=[];
tmpY=[];
tmpZ=[];

tmpX=[tmpX,pdbstruct.Model(mod_cnt).Atom(chn_ind(r
es_ind(6:atms))).X];

tmpY=[tmpY,pdbstruct.Model(mod_cnt).Atom(chn_ind(r
es_ind(6:atms))).Y];

tmpZ=[tmpZ,pdbstruct.Model(mod_cnt).Atom(chn_ind(r
es_ind(6:atms))).Z];
    % check for surface atoms
    [ surf_tst ] =
surf_atom( [tmpX',tmpY',tmpZ'],[allX',allY',allZ']
);
    %

    if surf_tst==1
        resd_name=['[*',resd_name];
    else
        resd_name=['[',resd_name];
    end
    REDOX_CHAIN{1,1}{rcount,1}=resd_name;

REDOX_CHAIN{1,2}{rcount,1}=chn_ind(res_ind(6:atms)
)';
    end
%
end
end
%
%
val_chn_ind=cell(1,num_chains);
%cell array of VAL indeces for each chain

```

```

%
for ij=((mod_cnt-
1)*num_chains)+1:(mod_cnt)*num_chains
    %
    % find VAL residues
    %

val_ind=0; %
reset VAL index counter
    %

chn_ind=find(strcmp([chains(ij),'VAL'],atom_chn_re
s)>0);
    %chn_ind=strfind(pdbstruct.Sequence(ij).Sequen
ce,'AA')';
%
    val_res_nos_un=[];
    if ~isempty(chn_ind)
        res_nos=[];

res_nos=[res_nos,pdbstruct.Model(mod_cnt).Atom(chn
_ind).resSeq]'; %get residue numbers of
ChainID|VAL

val_res_nos_un=unique(res_nos);
%find the unique residue numbers
    end

%
A_total(20,ij,mod_cnt)=length(val_res_nos_un);
%
%
% cmp_var = val_res_nos_un(2:end);
% cmp_var = [cmp_var; -1];
% diff = cmp_var - val_res_nos_un;
% check_var = [find(diff==1),
find(diff==1)+1]';
% check_var = reshape(check_var,
[size(check_var, 2)*2, 1]);
% check_var_u = unique(check_var);
% val_res_nos_un =

```

```

val_res_nos_un(check_var_u);

AA_total(20,ij,mod_cnt)=length(val_res_nos_un);
%
    for
jk=1:length(val_res_nos_un)
% test all of VAL residues

res_ind=find(res_nos==val_res_nos_un(jk));
%get the indeces of Chain|VAL(jk)
%
    if (length(res_ind)>5)
        rcount=rcount+1;
        val_ind=val_ind+1;
        val_chn_ind{1,ij}(val_ind,1)=rcount;
        resd_name=['VAL-
',chains(ij),num2str(val_res_nos_un(jk)),''];
        %resd_name=['TYR-
',chains(ij),num2str(tyr_res_nos_un(jk)),''];
        AA(20,ij,mod_cnt)=AA(20,ij,mod_cnt)+1;
%

        atms=min(7,length(res_ind));
        tmpX=[];
        tmpY=[];
        tmpZ=[];

tmpX=[tmpX,pdbstruct.Model(mod_cnt).Atom(chn_ind(r
es_ind(6:atms))).X];

tmpY=[tmpY,pdbstruct.Model(mod_cnt).Atom(chn_ind(r
es_ind(6:atms))).Y];

tmpZ=[tmpZ,pdbstruct.Model(mod_cnt).Atom(chn_ind(r
es_ind(6:atms))).Z];
        % check for surface atoms
        [ surf_tst ] =
surf_atom( [tmpX',tmpY',tmpZ'],[allX',allY',allZ']

```

```

);
    %
    if surf_tst==1
        resd_name=['[*',resd_name];
    else
        resd_name=['[' ,resd_name];
    end
    REDOX_CHAIN{1,1}{rcount,1}=resd_name;

REDOX_CHAIN{1,2}{rcount,1}=chn_ind(res_ind(6:atms)
)';
    end
%
end
end
%
end %end the model counting loop
%
% Now generate the distance matrix, DIST;
%
% new algorithm
%
DIST=zeros( sum(sum(sum(AA))),
sum(sum(sum(AA))) );
%
%
jk_cnt=0;
%
for mod_cnt1=1:models
%
%
    for jk=1:sum(sum(AA(:, :, mod_cnt1)))
        jk_cnt=jk_cnt+1;
        kl_cnt=0;
        atms1=length(REDOX_CHAIN{1,2}{jk_cnt});
        %
test_X=[pdbstruct.Model(mod_cnt1).Atom(REDOX_CHAIN

```

```

{1,2}{jk_cnt}).X];

test_Y=[pdbstruct.Model(mod_cnt1).Atom(REDOX_CHAIN
{1,2}{jk_cnt}).Y];

test_Z=[pdbstruct.Model(mod_cnt1).Atom(REDOX_CHAIN
{1,2}{jk_cnt}).Z];
    %
    for mod_cnt2=1:models
        %
        for kl=1:sum(sum(AA(:, :, mod_cnt2)))
            kl_cnt=kl_cnt+1;
        %
atms2=length(REDOX_CHAIN{1,2}{kl_cnt});
        %
            Amat_X=((test_X'*ones(1,atms2))-
(ones(atms1,1)*[pdbstruct.Model(mod_cnt2).Atom(RED
OX_CHAIN{1,2}{kl_cnt}).X])).^2;
            Amat_Y=((test_Y'*ones(1,atms2))-
(ones(atms1,1)*[pdbstruct.Model(mod_cnt2).Atom(RED
OX_CHAIN{1,2}{kl_cnt}).Y])).^2;
            Amat_Z=((test_Z'*ones(1,atms2))-
(ones(atms1,1)*[pdbstruct.Model(mod_cnt2).Atom(RED
OX_CHAIN{1,2}{kl_cnt}).Z])).^2;
        %
            Amat_XYZ=sqrt(Amat_X+Amat_Y+Amat_Z);
            dist=min(min(Amat_XYZ));
        %
            DIST(jk_cnt,kl_cnt)=min(min(dist));
        %
        end
    %
    end
end
end
end
%
end          %end the model counting loop
%
%
```



```

tmp=ones( sum(sum(sum(AA))), sum(sum(sum(AA))) );
%V=1:sum(sum(sum(YWMFHCTS))+1:(sum(sum(sum(YWMFHC
TS))))^2;
%tmp(V)=0;
tmp=tril(tmp,-1);
DIST=DIST./tmp;      % put Inf along and above the
main diagonal
indy= DIST>cuttoff2;
    % put Inf at all places where distance > cutoff
DIST(indy)=Inf;
indy= DIST==cuttoff;
DIST(indy)=Inf;
indy= DIST<cuttoff;
DIST(indy)=Inf;
%
%
if isempty(REDOX_CHAIN{1})
    fprintf(' No residues')
    AA_pairs=zeros(20,5);      % OUTPUT RESULT
    fclose('all');
    return
end
%
K=strfind(REDOX_CHAIN{1}, 'ALA'); % compare all
residues to ALA
A_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDOX_CHAIN{1}, 'ARG'); % compare all
residues to ARG
R_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDOX_CHAIN{1}, 'ASN'); % compare all
residues to ASN
N_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDOX_CHAIN{1}, 'ASP'); % compare all

```

```
residues to ASP
D_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDUX_CHAIN{1}, 'CYS'); % compare all
residues to CYS
C_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDUX_CHAIN{1}, 'GLU'); % compare all
residues to GLU
E_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDUX_CHAIN{1}, 'GLN'); % compare all
residues to GLN
Q_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDUX_CHAIN{1}, 'GLY'); % compare all
residues to GLY
G_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDUX_CHAIN{1}, 'HIS'); % compare all
residues to HIS
H_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDUX_CHAIN{1}, 'ILE'); % compare all
residues to Ile
I_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDUX_CHAIN{1}, 'LEU'); % compare all
residues to Leu
L_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDUX_CHAIN{1}, 'LYS'); % compare all
```

```
residues to Lys
K_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDUX_CHAIN{1}, 'MET'); % compare all
residues to Met
M_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDUX_CHAIN{1}, 'PHE'); % compare all
residues to Phe
F_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDUX_CHAIN{1}, 'PRO'); % compare all
residues to PRO
P_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDUX_CHAIN{1}, 'SER'); % compare all
residues to SER
S_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDUX_CHAIN{1}, 'THR'); % compare all
residues to THR
T_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDUX_CHAIN{1}, 'TRP'); % compare all
residues to TRP
W_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDUX_CHAIN{1}, 'TYR'); % compare all
residues to TYR
Y_ind=find(cellfun(@isempty,K)==0);
%
%
K=strfind(REDUX_CHAIN{1}, 'VAL'); % compare all
```

residues to VAL

```
V_ind=find(cellfun(@isempty,K)==0);
```

```
%
```

```
%
```

```
DISTA=DIST(A_ind,A_ind)./tril(ones(length(A_ind)),  
-1);
```

```
DISTR=DIST(R_ind,R_ind)./tril(ones(length(R_ind)),  
-1);
```

```
DISTN=DIST(N_ind,N_ind)./tril(ones(length(N_ind)),  
-1);
```

```
DISTD=DIST(D_ind,D_ind)./tril(ones(length(D_ind)),  
-1);
```

```
DISTC=DIST(C_ind,C_ind)./tril(ones(length(C_ind)),  
-1);
```

```
DISTE=DIST(E_ind,E_ind)./tril(ones(length(E_ind)),  
-1);
```

```
DISTQ=DIST(Q_ind,Q_ind)./tril(ones(length(Q_ind)),  
-1);
```

```
DISTG=DIST(G_ind,G_ind)./tril(ones(length(G_ind)),  
-1);
```

```
DISTH=DIST(H_ind,H_ind)./tril(ones(length(H_ind)),  
-1);
```

```
DISTI=DIST(I_ind,I_ind)./tril(ones(length(I_ind)),  
-1);
```

```
DISTL=DIST(L_ind,L_ind)./tril(ones(length(L_ind)),  
-1);
```

```
DISTK=DIST(K_ind,K_ind)./tril(ones(length(K_ind)),  
-1);
```

```
DISTM=DIST(M_ind,M_ind)./tril(ones(length(M_ind)),  
-1);
```

```
DISTF=DIST(F_ind,F_ind)./tril(ones(length(F_ind)),  
-1);
```

```
DISTP=DIST(P_ind,P_ind)./tril(ones(length(P_ind)),  
-1);
```

```
DISTS=DIST(S_ind,S_ind)./tril(ones(length(S_ind)),  
-1);
```

```
DISTT=DIST(T_ind,T_ind)./tril(ones(length(T_ind)),  
-1);
```

```
DISTW=DIST(W_ind,W_ind)./tril(ones(length(W_ind)),
```

```

-1);
DISTY=DIST(Y_ind,Y_ind)./tril(ones(length(Y_ind)),
-1);
DISTV=DIST(V_ind,V_ind)./tril(ones(length(V_ind)),
-1);

tmp=sum(sum(AA_total,3),2);
tmp=[tmp;zeros(0,1)];
AA_pairs(:,6)=tmp;      % put total number of each
residue in seq pairs in column 6

% tmp=sum(sum(A_total,3),2);
% tmp=[tmp;zeros(0,1)];
% AA_pairs(:,5)=tmp;      % put total number of
each residue in column 5

AA_string={};
str_cnt=0;
%
str_cnt=str_cnt+1;
AA_string{str_cnt,1}=[pdb_id,':
',num2str(cutoff),' ? cutoff distance'];
tmp='  ';
str_cnt=str_cnt+1;
AA_string{str_cnt,1}=tmp;
%
%
[I,J]=find(isfinite(DISTA)&(DISTA>0));
tmp='ALA-ALA: ';
str_cnt=str_cnt+1;
str_cnt2=1;
AA_string{str_cnt,1}=tmp;
if ~isempty(I)
    for lpY=1:length(I)
        tmp=[REDOX_CHAIN{1,1}{A_ind(I(lpY)),1},'-
',REDOX_CHAIN{1,1}{A_ind(J(lpY)),1},':
',num2str(DISTA(I(lpY),J(lpY))), ' ?'];
%         if
REDOX_CHAIN{1,1}{A_ind(I(lpY)),1}(2)=='*'
%             chain_name =

```

```

REDOX_CHAIN{1,1}{A_ind(I(lpY)),1}(7);
%         chain_num =
REDOX_CHAIN{1,1}{A_ind(I(lpY)),1}(8:end-1);
%         else
%         chain_name =
REDOX_CHAIN{1,1}{A_ind(I(lpY)),1}(6);
%         chain_num =
REDOX_CHAIN{1,1}{A_ind(I(lpY)),1}(7:end-1);
%         end
%         chain_num = str2num(chain_num);
%
%         if
REDOX_CHAIN{1,1}{A_ind(J(lpY)),1}(2)=='*'
%         chain_name2 =
REDOX_CHAIN{1,1}{A_ind(J(lpY)),1}(7);
%         chain_num2 =
REDOX_CHAIN{1,1}{A_ind(J(lpY)),1}(8:end-1);
%         else
%         chain_name2 =
REDOX_CHAIN{1,1}{A_ind(J(lpY)),1}(6);
%         chain_num2 =
REDOX_CHAIN{1,1}{A_ind(J(lpY)),1}(7:end-1);
%         end
%         chain_num2 = str2num(chain_num2);
%         if (chain_name == chain_name2) &&
(abs(chain_num2 - chain_num) == 1)

        str_cnt=str_cnt+1;
        AA_string{str_cnt,1}=tmp;

        tmp=[REDOX_CHAIN{1,1}{A_ind(I(lpY)),1}];
        A_string{str_cnt2,1}=tmp;
        str_cnt2=str_cnt2+1;
        tmp=[REDOX_CHAIN{1,1}{A_ind(J(lpY)),1}];
        A_string{str_cnt2,1}=tmp;
        str_cnt2=str_cnt2+1;

indy1=strfind(REDOX_CHAIN{1}{A_ind(I(lpY)),1}, '*')
;

```

```

indy2=strfind(REDUX_CHAIN{1}{A_ind(J(lpY)),1},'*')
;
    if ~isempty(indy1)&&~isempty(indy2)
        AA_pairs(1,3)=AA_pairs(1,3)+1;
    elseif ~isempty(indy1)||~isempty(indy2)
        AA_pairs(1,2)=AA_pairs(1,2)+1;
    else
        AA_pairs(1,1)=AA_pairs(1,1)+1;
    end
end
%     end
end
%
end

% A_string_value=strfind(A_string(4,1),'ALA');
if contains(AA_string{str_cnt,1},'ALA-ALA:')==1
    AA_pairs(1,4)=0;
elseif contains(A_string{1,1},'ALA')==1

A_number_as_pair=unique(cellfun(@num2str,A_string,
'uni',false));
    AA_pairs(1,4)=length(A_number_as_pair);
    A_surf=count(A_number_as_pair, '*');
    AA_pairs(1,5)=sum(A_surf);
end

%
A_number_as_pair=cell2table(A_number_as_pair);
%
AA_pairs(1,4)=length(unique(A_number_as_pair,
'rows'));

tmp=' ';
str_cnt=str_cnt+1;
AA_string{str_cnt,1}=tmp;
%
%
[I,J]=find(isfinite(DISTR)&(DISTR>0));
tmp='ARG-ARG: ';

```

```

str_cnt2=1;
str_cnt=str_cnt+1;
AA_string{str_cnt,1}=tmp;
if ~isempty(I)
    for lpY=1:length(I)
        tmp=[REDOX_CHAIN{1,1}{R_ind(I(lpY)),1},'-
',REDOX_CHAIN{1,1}{R_ind(J(lpY)),1},':
',num2str(DISTR(I(lpY),J(lpY))), '?'];
        %         if
REDOX_CHAIN{1,1}{R_ind(I(lpY)),1}(2)=='*'
        %             chain_name =
REDOX_CHAIN{1,1}{R_ind(I(lpY)),1}(7);
        %             chain_num =
REDOX_CHAIN{1,1}{R_ind(I(lpY)),1}(8:end-1);
        %         else
        %             chain_name =
REDOX_CHAIN{1,1}{R_ind(I(lpY)),1}(6);
        %             chain_num =
REDOX_CHAIN{1,1}{R_ind(I(lpY)),1}(7:end-1);
        %         end
        %         chain_num = str2num(chain_num);
        %
        %         if
REDOX_CHAIN{1,1}{R_ind(J(lpY)),1}(2)=='*'
        %             chain_name2 =
REDOX_CHAIN{1,1}{R_ind(J(lpY)),1}(7);
        %             chain_num2 =
REDOX_CHAIN{1,1}{R_ind(J(lpY)),1}(8:end-1);
        %         else
        %             chain_name2 =
REDOX_CHAIN{1,1}{R_ind(J(lpY)),1}(6);
        %             chain_num2 =
REDOX_CHAIN{1,1}{R_ind(J(lpY)),1}(7:end-1);
        %         end
        %         chain_num2 = str2num(chain_num2);
        %         if (chain_name == chain_name2) &&
(abs(chain_num2 - chain_num) == 1)

            str_cnt=str_cnt+1;
            AA_string{str_cnt,1}=tmp;

```



```

        tmp=[REDOX_CHAIN{1,1}{R_ind(I(lpY)),1}];
        R_string{str_cnt2,1}=tmp;
        str_cnt2=str_cnt2+1;
        tmp=[REDOX_CHAIN{1,1}{R_ind(J(lpY)),1}];
        R_string{str_cnt2,1}=tmp;
        str_cnt2=str_cnt2+1;
    %
    indy1=strfind(REDOX_CHAIN{1}{R_ind(I(lpY)),1},'*')
    ;

    indy2=strfind(REDOX_CHAIN{1}{R_ind(J(lpY)),1},'*')
    ;
        if ~isempty(indy1)&&~isempty(indy2)
            AA_pairs(2,3)=AA_pairs(2,3)+1;
        elseif ~isempty(indy1)||~isempty(indy2)
            AA_pairs(2,2)=AA_pairs(2,2)+1;
        else
            AA_pairs(2,1)=AA_pairs(2,1)+1;
        end
    %     end
    end
    %
end

if contains(AA_string{str_cnt,1},'ARG-ARG:')==1
    AA_pairs(2,4)=0;
elseif contains(R_string{1,1},'ARG')==1

R_number_as_pair=unique(cellfun(@num2str,R_string,
'uni',false));
    AA_pairs(2,4)=length(R_number_as_pair);
    R_surf=count(R_number_as_pair, '*');
    AA_pairs(2,5)=sum(R_surf);
end

tmp=' ';
str_cnt=str_cnt+1;
AA_string{str_cnt,1}=tmp;

```

```

%
% end
%

.

.

.

.

.

[I,J]=find(isfinite(DISTW) & (DISTW>0));
tmp='TRP-TRP: ';
str_cnt2=1;
str_cnt=str_cnt+1;
AA_string{str_cnt,1}=tmp;
if ~isempty(I)
    for lpY=1:length(I)
        tmp=[REDOX_CHAIN{1,1}{W_ind(I(lpY)),1}, '-
',REDOX_CHAIN{1,1}{W_ind(J(lpY)),1}, ':
',num2str(DISTW(I(lpY),J(lpY))), ' ?'];
%         if
REDOX_CHAIN{1,1}{W_ind(I(lpY)),1}(2)=='*'
%             chain_name =
REDOX_CHAIN{1,1}{W_ind(I(lpY)),1}(7);
%             chain_num =
REDOX_CHAIN{1,1}{W_ind(I(lpY)),1}(8:end-1);
%         else
%             chain_name =
REDOX_CHAIN{1,1}{W_ind(I(lpY)),1}(6);
%             chain_num =
REDOX_CHAIN{1,1}{W_ind(I(lpY)),1}(7:end-1);
%         end
%         chain_num = str2num(chain_num);
%
%         if

```

```

REDOX_CHAIN{1,1}{W_ind(J(lpY)),1}(2)=='*'
%         chain_name2 =
REDOX_CHAIN{1,1}{W_ind(J(lpY)),1}(7);
%         chain_num2 =
REDOX_CHAIN{1,1}{W_ind(J(lpY)),1}(8:end-1);
%         else
%         chain_name2 =
REDOX_CHAIN{1,1}{W_ind(J(lpY)),1}(6);
%         chain_num2 =
REDOX_CHAIN{1,1}{W_ind(J(lpY)),1}(7:end-1);
%         end
%         chain_num2 = str2num(chain_num2);
%         if (chain_name == chain_name2) &&
(abs(chain_num2 - chain_num) == 1)

        str_cnt=str_cnt+1;
        AA_string{str_cnt,1}=tmp;

        tmp=[REDOX_CHAIN{1,1}{W_ind(I(lpY)),1}];
        W_string{str_cnt2,1}=tmp;
        str_cnt2=str_cnt2+1;
        tmp=[REDOX_CHAIN{1,1}{W_ind(J(lpY)),1}];
        W_string{str_cnt2,1}=tmp;
        str_cnt2=str_cnt2+1;
%
%
indy1=strfind(REDOX_CHAIN{1}{W_ind(I(lpY)),1}, '*')
;

indy2=strfind(REDOX_CHAIN{1}{W_ind(J(lpY)),1}, '*')
;

        if ~isempty(indy1) && ~isempty(indy2)
            AA_pairs(18,3)=AA_pairs(18,3)+1;
        elseif ~isempty(indy1) || ~isempty(indy2)
            AA_pairs(18,2)=AA_pairs(18,2)+1;
        else
            AA_pairs(18,1)=AA_pairs(18,1)+1;
        end
%         end
end
end

```

```

%
end

if contains(AA_string{str_cnt,1}, 'TRP-TRP:')==1
    AA_pairs(18,4)=0;
elseif contains(W_string{1,1}, 'TRP')==1

W_number_as_pair=unique(cellfun(@num2str,W_string,
'uni',false));
    AA_pairs(18,4)=length(W_number_as_pair);
    W_surf=count(W_number_as_pair, '*');
    AA_pairs(18,5)=sum(W_surf);
end

tmp=' ';
str_cnt=str_cnt+1;
AA_string{str_cnt,1}=tmp;
%
% end
%

[I,J]=find(isfinite(DISTY) & (DISTY>0));
tmp='TYR-TYR: ';
str_cnt2=1;
str_cnt=str_cnt+1;
AA_string{str_cnt,1}=tmp;
if ~isempty(I)
    for lpY=1:length(I)
        tmp=[REDOX_CHAIN{1,1}{Y_ind(I(lpY)),1}, '-
',REDOX_CHAIN{1,1}{Y_ind(J(lpY)),1}, ':
',num2str(DISTY(I(lpY),J(lpY))), ' ?'];
%         if
REDOX_CHAIN{1,1}{Y_ind(I(lpY)),1}(2)=='*'
%             chain_name =
REDOX_CHAIN{1,1}{Y_ind(I(lpY)),1}(7);
%             chain_num =
REDOX_CHAIN{1,1}{Y_ind(I(lpY)),1}(8:end-1);
%         else
%             chain_name =
REDOX_CHAIN{1,1}{Y_ind(I(lpY)),1}(6);

```



```

        AA_pairs(19,2)=AA_pairs(19,2)+1;
    else
        AA_pairs(19,1)=AA_pairs(19,1)+1;
    end
%     end
end
%
end

if contains(AA_string{str_cnt,1}, 'TYR-TYR:')==1
    AA_pairs(19,4)=0;
elseif contains(Y_string{1,1}, 'TYR')==1

Y_number_as_pair=unique(cellfun(@num2str,Y_string,
'uni',false));
    AA_pairs(19,4)=length(Y_number_as_pair);
    Y_surf=count(Y_number_as_pair, '*');
    AA_pairs(19,5)=sum(Y_surf);
end

tmp=' ';
str_cnt=str_cnt+1;
AA_string{str_cnt,1}=tmp;
%
% end
%

[I,J]=find(isfinite(DISTV) & (DISTV>0));
tmp='VAL-VAL: ';
str_cnt2=1;
str_cnt=str_cnt+1;
AA_string{str_cnt,1}=tmp;
if ~isempty(I)
    for lpY=1:length(I)
        tmp=[REDOX_CHAIN{1,1}{V_ind(I(lpY)),1}, '-
',REDOX_CHAIN{1,1}{V_ind(J(lpY)),1}, ':
',num2str(DISTV(I(lpY),J(lpY))), ' ?'];
%         if
REDOX_CHAIN{1,1}{V_ind(I(lpY)),1}(2)=='*'
%             chain_name =

```

```

REDOX_CHAIN{1,1}{V_ind(I(lpY)),1}(7);
%         chain_num =
REDOX_CHAIN{1,1}{V_ind(I(lpY)),1}(8:end-1);
%         else
%         chain_name =
REDOX_CHAIN{1,1}{V_ind(I(lpY)),1}(6);
%         chain_num =
REDOX_CHAIN{1,1}{V_ind(I(lpY)),1}(7:end-1);
%         end
%         chain_num = str2num(chain_num);
%
%         if
REDOX_CHAIN{1,1}{V_ind(J(lpY)),1}(2)=='*'
%         chain_name2 =
REDOX_CHAIN{1,1}{V_ind(J(lpY)),1}(7);
%         chain_num2 =
REDOX_CHAIN{1,1}{V_ind(J(lpY)),1}(8:end-1);
%         else
%         chain_name2 =
REDOX_CHAIN{1,1}{V_ind(J(lpY)),1}(6);
%         chain_num2 =
REDOX_CHAIN{1,1}{V_ind(J(lpY)),1}(7:end-1);
%         end
%         chain_num2 = str2num(chain_num2);
%         if (chain_name == chain_name2) &&
(abs(chain_num2 - chain_num) == 1)

        str_cnt=str_cnt+1;
        AA_string{str_cnt,1}=tmp;

        tmp=[REDOX_CHAIN{1,1}{V_ind(I(lpY)),1}];
        V_string{str_cnt2,1}=tmp;
        str_cnt2=str_cnt2+1;
        tmp=[REDOX_CHAIN{1,1}{V_ind(J(lpY)),1}];
        V_string{str_cnt2,1}=tmp;
        str_cnt2=str_cnt2+1;
%
%
indy1=strfind(REDOX_CHAIN{1}{V_ind(I(lpY)),1}, '*')
;

```

```

indy2=strfind(REDUX_CHAIN{1}{V_ind(J(lpY)),1},'*')
;
    if ~isempty(indy1)&&~isempty(indy2)
        AA_pairs(20,3)=AA_pairs(20,3)+1;
    elseif ~isempty(indy1)||~isempty(indy2)
        AA_pairs(20,2)=AA_pairs(20,2)+1;
    else
        AA_pairs(20,1)=AA_pairs(20,1)+1;
    end
end
%     end
end

if contains(AA_string{str_cnt,1},'VAL-VAL:')==1
    AA_pairs(20,4)=0;
elseif contains(V_string{1,1},'VAL')==1

V_number_as_pair=unique(cellfun(@num2str,V_string,
'uni',false));
    AA_pairs(20,4)=length(V_number_as_pair);
    V_surf=count(V_number_as_pair, '*');
    AA_pairs(20,5)=sum(V_surf);
end

%
% end
% tmp=' ';
% str_cnt=str_cnt+1;
% AA_string{str_cnt,1}=tmp;
%
% end
%
%
end
%
%
filename=[pname,'mat\',pdb_id,'_AA_pairs_struct_0_
5.mat'];
save(filename,'pdb_id','cutoff','AA_pairs','AA_str
ing');

```



```
%  
delete('C:\Users\Jieun  
Shin\Desktop\bioinf\pub\pdb\data\biunit\coordinat  
es\all\*.gz')  
end
```