UNIVERSITY OF CALIFORNIA SAN DIEGO

NMR Applications to Study Natural Product Biosynthesis and Biodegradation

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Chemistry

by

Brianna Nicole Kalaj

Committee in charge:

Professor Michael D. Burkart, Chair
Professor Tadeusz Molinksi
Professor Bradley Moore
Professor Joseph Noel
Professor Valerie Schmidt

2024

The Dissertation of Brianna Nicole Kalaj is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

DEDICATION

To my parents, who pushed me to prioritize my education and have supported me through any journey I look to adventure upon. I love you and appreciate your advice and love more than you can ever know.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SCHEMES

LIST OF TABLES

ACKNOWLEDGEMENTS

First, I would like to acknowledge my advisor, Professor Michael D. Burkart for his guidance throughout my graduate studies. From the beginning of this journey, your mentorship and support through anything I have encountered has been truly appreciated. You build a lab that is not only filled with highly motivated and collaborative scientists but also a family in many aspects. From group retreats to holiday celebrations, we appreciate your ability to see us as members of the Burkart Lab family and genuinely acknowledge everyone as individuals. Graduate school would not have been nearly as enjoyable of an experience with any other mentor or lab, thank you. I would also like to thank Dr. James J La Clair for all of his help and support throughout graduate school as well. Your help in learning to be a good scientist as well as how to present and write professionally has been invaluable, thank you.

I would like to thank my committee members for pushing me to be a better scientist through the questions, discussions, and feedback during my exams. I would also like to thank all of those who were collaborators through my various projects, especially Dr. Qing Zhe Ni, Dr. Ryan Simkovsky, and Dr. Lalit Deshmukh. Your mentorship and teaching in each respective field as well as guidance on projects is greatly appreciated, the papers we have developed would not be what they are without our discussions.

I would like to thank those who encouraged me to pursue chemistry graduate school from the time I gained an interest in it and helped lead me toward pursing my PhD. Professor Scott Eagon, my mentor during my research at Cal Poly SLO, thank you for introducing me to research and its joys. My family at Promega Biosciences including but not limited to my Dr. Cesear Corona, David Good, and all of the Process Development and Research & Development

team scientists. Thank you all for showing me what scientists can be and opening my eyes to the different aspects of biochemical research.

I would like to thank Burkart Lab members both past and present, you all are truly more than colleagues but friends as well. Tons of memories were made and laughs were shared at many Burkart lab events and I will cherish them always. I also greatly appreciate the support and guidance from various postdocs and labmates including Katia Charov, Kelsey Krug, Tony D. Davis, and Graham Heberlig over the course of various projects, successes and failures, laughable moments and frustrations.

I would like to thank my cohort friends Judith Flores, Kelly Hunter, Stephanie McCabe, Raymond Berkeley, Josh Corpuz, and Satyam Khanal for the many memories we made during graduate school together. From first year memories when we thought we were stressed, to years into our PhDs when we realized what stress really was, you were all there to commiserate with and bring happiness to the day to day. A special thank you to Rebecca Re, a best friend and bouncing board for anything science, I cannot express how much you have meant to me throughout this adventure. I would not have been able to get through this program without you just a room or call away for all of my questions, complaints, worries, and joys, both lab and life related. I will always cherish our memories we have made together during this time and cannot wait for all the new ones we will be there to see each other through in life.

I would especially like to thank my family for their years of support and love throughout this journey. To my grandparents, aunts, uncles and cousins who always ask how I am doing and make sure I know they are proud of me, your blind support of what I am doing and ability to connect me to real life again when I come home as been tremendously appreciated. To my second family, my in-laws, Margerita and Leonard Kalaj, and my newest siblings, thank you

for being my winter wonderland home, always filling me with the best Albanian foods, and showing me love as if I am one of your own. To my wonderful sisters, Erika Mouton and Angela Sierra, I always appreciate you bringing me back down to earth when I need it. Though we may be far in distance you are both always in my heart and I love you so much. And the biggest thank you to my parents, Michael and Samantha Sierra, I can not begin to express how grateful I am for your love and support in everything I set out to adventure upon. You have taught me through everything you do in life that with hard work and determination I can achieve whatever I set my mind to. Your emphasis on education and my sisters and I achieving our own successes in life has been what has led me to where I am and will always carry with me, I love you more than words can express.

And last but certainly never least, thank you to my husband, Mark Kalaj. You are my home, my best friend, the love of my life, and meeting you in this program was the best thing to ever happen to me. You are an amazing scientist, friend, and husband, and an inspiration to me to always be a better person in everything I do. Thank you for being my sounding board and biggest supporter in all things during graduate school and always picking me up when I am down. I can't wait to spend the rest of my life with you and am excited for all of our new adventures we will embark on as a family. Shpirti jem, ju du fort Zemer.

Chapter 1, in full, is a reprint of the material as it appears, Ni, Qing Zhe; Sierra, Brianna N., La Clair, James J., Burkart, Michael D. "Chemoenzymatic elaboration of the Raper-Mason pathway unravels the structural diversity within eumelanin pigments." *Chem. Sci.* **2020**, *11*, 7836-7841. The dissertation author was the secondary investigator and author of this paper.

Chapter 2, in full, is a reprint of the material as it appears, Kalaj, Brianna N.; Ni, Qing Zhe.; La Clair, James J.; Burkart, Michael D. "Chemoenzymatic isolation and characterization

of high purity mammalian melanin." *ChemBioChem*, **2022**, *23*, e202200021. The dissertation author was the co-primary investigator and author of this paper.

Chapter 3, in full, is currently being prepared for submission for publication of the material as Kalaj, Brianna N.; Tessman, Marissa; Simkovsky, Ryan; Mayfield, Stephen P.; Burkart, Michael D. "Investigating methylene diphenyl diamine alteration and uptake by soil and compost organisms." The dissertation author was the primary researcher and author of this material.

Chapter 4, in full, is currently being prepared for submission for publication of the material as Kalaj, Brianna N.; LaClair, James J.; Shen, Yang.; Schweiters, Charles, D.; Deshmukh, Lalit; Burkart, Michael D. "Quantitative characterization of chain-flipping of acyl carrier protein of *Escherichia coli* using chemical exchange NMR." The dissertation author was the primary researcher and author of this material.

VITA

2017 Bachelor of Science, California Polytechnic State University - San Luis Obispo

2019 Master of Science, University of California San Diego

2024 Doctor of Philosophy, University of California San Diego

PUBLICATIONS

Kalaj, B., La Clair, J.J., Shen, Y., Schweiters, C.D., Deshmukh, L., Burkart, M.D. Quantitative characterization of chain-flipping of acyl carrier protein of *Escherichia coli* using chemical exchange NMR. (To be submitted)

Kalaj, B., Tessman, M., Simkovsky, R., Mayfield, S.P., Burkart, M.D. Investigating methylene diphenyl diamine alteration and uptake by soil and compost organisms. (To be submitted)

Viswanathan, N.K., Chirgwin, M.E., Gibbs, J., **Kalaj, B.**, Durham, S., Tran, J.K., Gomez, M., Lazaro, H., Chen, M., Mansfield, C.R., Derbyshire, E.R., Eagon, S. *Bioorg. Med. Chem. Lett*. **2023**, *92*, 129410. Synthesis and activity of β-carboline antimalarials targeting the *Plasmodium falciparum* heat shock 90 protein.

Kalaj, B., Zhe Ni, Q., La Clair, J.J., Burkart, M.D. *ChemBioChem.* **2022**, *23*, e202200021. Chemoenzymatic isolation and characterization of high purity mammalian melanin.

Zhe Ni, Q., **Sierra, B**., La Clair, J.J., Burkart, M.D. *Chem. Sci.* **2020**, *11*, 7836-7841. Chemoenzymatic elaboration of the Raper-Mason pathway unravels the structural diversity within eumelanin pigments.

Durham, S., **Sierra, B**., Gomez, M.J., Tran, J.K, Anderson, M.O., Whittington-Davis, N.A., Eagon, S. *Tet. Lett.* **2017**, *58*, 2747-2750. Synthesis of β-carbolines via a silver-mediated oxidation of tetrahydro-β-carbolines.

ABSTRACT OF THE DISSERTATION

NMR Applications to Study Natural Product Biosynthesis and Biodegradation

by

Brianna Nicole Kalaj

Doctor of Philosophy in Chemistry

University of California San Diego, 2024

Professor Michael D. Burkart, Chair

Nuclear magnetic resonance (NMR) spectroscopy is a powerful analytical technique and can be used to analyze anything from small molecules to large polymers and proteins. This non-destructive tool is useful in many fields of research from basic science to applied work. In

this thesis, NMR is applied in a multitude of ways including analyzing small molecule structure, backbone fingerprinting of polymeric material, and protein ligand interactions and dynamics.

In the first two chapters of this thesis, amorphous melanin polymers are analyzed by solid-state NMR cross polarization ^{13}C experiments. Through this technique we identify types and distributions of carbon within a sample and correlate proposed starting material with fingerprints of complex, chemoenzymatically generated melanins. We also use solid-state NMR to estimate purity of extracted melanin samples from marine melanin species by identification of components of marine tissue material.

The third chapter of this thesis investigates how a predicted toxic metabolite, methylene diphenyl diamine, is biodegraded over time by microbial organisms and communities by analyzing the nitrogen state changes. Using an ^{15}N labeled small molecule and ^{15}N INEPT experiments, nitrogen environment changes are observed for monitoring a how this molecule is metabolized after feeding studies to microbial communities and thus identifying new insights into the biodegradation process. In this work NMR and LCMS are combined toward a possible explanation of how methylene diphenyl diamine is altered after potential release from bio-based polyurethane biodegradation.

In the final chapter of this thesis, NMR is used to monitoring carrier protein-ligand dynamics by employing a site specific ^{15}N labeled probe. Synthetic organic chemistry is utilized to develop a versatile ^{15}N tool which can be applied to various carrier protein dependent systems through a simple modification. NMR's capability to monitor dynamic changes of a system over time is used to trace the chemical shift change of a minor state of a ligand in a carrier protein dependent system, allowing for more concrete insight into sequestration and chain flipping

dynamics. ^{15}N CEST experiments are used to extract data about ligand dynamics which mimic the native substrates of fatty acid biosynthesis, a model carrier protein dependent system.

INTRODUCTION

Nuclear magnetic resonance (NMR) spectroscopy is a powerful technique for analyzing structures of molecules, the nature of bonds, and environments of atoms. NMR can be used to analyze anything from small molecules to large polymers and proteins, and to study samples in liquid, semi-solid, or solid state.[1,2] This analysis tool is also non-destructive to samples, making it useful in many fields of research from basic science to applied work in medicine and agriculture.[3,4,5] NMR is a versatile technique to many fields as it can require little sample preparation, dilute samples, does not require chemical derivation, and is quantifiable. However, the strength in this tool and what makes it useful in a wide range of problem solving techniques, is its constant utility through new innovations.[6,7]

After stable isotopes first were isolated and applied to chemical compounds in the 1930s, particularly those of ^{13}C, ^{15}N, ^{18}O, and ^{2}H, they have been used as tracers in a wide range of applications from studying biological systems to forensics.[8,9,10,11] These "*chemically and functionally identical*" versions of elements, in which a different number of neutrons are in the nucleus, occupy the same position on the periodic table yet are analytically distinguishable.[12] When these labels are introduced to a molecule or system, NMR sensitivity can be increased as the natural abundance of these elements with unpaired nuclear spins become more prominent in the sample. Nuclear magnetic resonance works through certain nuclei possessing spin, or angular momentum, generating a magnetic moment.[13] The nuclei able to generate magnetic moments contain unpaired nuclear spins and are therefore able to be visualized by NMR with their ability to resonate at specific frequencies characteristic to their environment.

Proton and carbon have a sufficient natural abundance of their respective NMR sensitive nuclei, ^{1}H and ^{13}C, as they are 99.98% and 1.11% respectively, leading to easily obtainable NMR

spectra.[13] However, with nitrogen the natural abundance of ^{15}N, the NMR detectable nuclei, is only 0.365% leading to the inability to obtain decent signal to noise easily with natural samples and the need for stable isotope incorporation.[13] Stable isotopes of ^{15}N are incorporated into small molecules and biological systems for NMR as they not only cover a wide range of chemical shifts but are also sensitive to changes in hydrogen bonding and protonation along with other changes in their environment and structure.[14]

In this thesis, NMR and stable isotopes have been used in a multitude of ways including analyzing small molecule structure, backbone fingerprinting of polymeric material, and protein ligand interactions and dynamics. In the first two chapters of this thesis, insoluble melanin polymers are analyzed by solid-state NMR cross polarization ^{13}C experiments. This is utilized as a tool to generate standard sample fingerprints of proposed small molecule starting blocks followed by comparison with complex, chemoenzymatically generated polymeric melanin materials. Through this technique we can identify types and distributions of carbon within a sample and correlate what starting materials were incorporated from the metabolic pool of specialized enzymes. We can also use solid-state NMR to estimate purity of extracted melanin samples from marine melanin species by comparison to chemoenzymatically generated melanins and identification of components of marine tissue material.

In the third chapter of this thesis, NMR's use as an everyday tool for synthetic chemists looking to quantify and identify the structure of small molecules using basic 1D techniques is taken advantage of. Herein, this is modified to look at how a predicted toxic metabolite is biodegraded over time by microbial organisms and communities by analyzing the nitrogen state changes.[15] In being able to trace nitrogen through different forms using an ^{15}N labeled small molecule and ^{15}N INEPT experiments, a simple output is observed for monitoring a change in how this molecule is

metabolized and thus identifying new insights into the process. In this work NMR is combined with LCMS to work toward a possible biodegradation path of a small molecule monomer predicted to be released from bio-based polyurethane biodegradation.

Arguably the most impactful use of NMR is in biological systems using solution state protein NMR to monitor protein-protein and protein-ligand interactions and dynamics.[7,8] In the seminal work of this thesis, NMR is used to monitoring protein-ligand dynamics by utilizing a synthesized site specific ^{15}N labeled probe. Synthetic organic chemistry techniques are deployed to develop an isotopically labeled tool which can be applied to various carrier protein dependent systems in the exchange of chains identities. Taking advantage of NMR's capability to monitor not only a single state of interaction, but the dynamic changes of a system over time has been used in carrier protein dependent systems to study protein-protein interactions with ^{15}N protein backbones before.[16,17] However, this probe can be used to trace the chemical shift change of a minor state of a ligand in a carrier protein dependent system, allowing for more concrete insight into chain flipping dynamics. Utilizing ^{15}N CEST experiments which can monitor exchanges on a timescale of 20 to 200 s^{-1} and minor states as low as 1.5 % we are able to extract data about ligand dynamics which mimic the native substrates of fatty acid biosynthesis, a model carrier protein dependent system.[18]

References

1. P. H. Wiloughby, M. J. Jansma, T. R. Hoye, A guide to small-molecule structure assignment through computation of (^{1}H and ^{13}C) NMR chemical shifts. *Nat. Protoc.* **2014**, *9*, 643-660.

2. D. D. Laws, H. M. L. Bitter, A. Jerschow, Solid-State NMR Spectroscopic Methods in Chemistry. *Angew. Chemie. Int. Ed.* **2002**, *41*, 3096-3129.

3. A. J. Simpson, M. J. Simpson, R. Soong, Nuclear Magnetic Resonance Spectroscopy and Its Key Role in Environmental Research. *Environ. Sci. Technol.* **2012**, *46*, 11488-11496.

4. P. A. Bottomley, NMR in medicine, *Computerized Radiology*, **1984**, *8*, 57-77.

5. L. A. Colnago, Z. Weisman, G. Pages, M. Musse, T. Monaretto, C. W. Windt, C. Rondeau-Mouro, *J. Magn. Reson.* **2021,** *323*, 106899.

6. P. J. Barrett, J. Chen, M. K. Cho, J. H. Kim, Z. Lu, S. Mathew, D. Peng, Y. Song, W. D. Van Horn, T. Zhuang, F. D. Sonnichsen, C. R. Sanders, The Quiet Renaissance of Protein Nuclear Magnetic Resonance. *Biochem.* **2013**, *52*, 1303-1320.

7. A. Mittermaier, L. E. Kay, New Tools Provide Insights in NMR Studies of Protein Dynamics. *Science*, **2006**, *312*, 224-229.

8. D. D. Boehr, H. J. Dyson, P. E. Wright, An NMR Perspective on Enzyme Dynamics. *Chem. Rev.* **2006**, *106*, 3055-3079.

9. E. R. P. Zuiderweg, Mapping Protein-Protein Interactions in Solution by NMR Spectroscopy. *Biochem*. **2002**, *41*, 1-7.

10. R. P. Philp, The emergence of stable isotopes in environmental and forensic geochemistry studies: a review. *Environ. Chem. Lett.* **2007**, *5*, 57-66.

11. T. E. Cerling, J. E. Barnette, G. J. Bowen, L. A. Chesson, J. R. Ehleringer, C. H. Remien, P. Shea, B. J. Tripple, J. B. West, Forensic Stable Isotope Biogeochemistry. *Annu. Rev. Earth Planet.* **2016**, *44*, 175-206.

12. D. J. Wilkinson, Historical and contemporary stable isotope tracer approaches to studying mammalian protein metabolism. *Mass Spectrom. Rev.* **2018**, *37*, 57-80.

13. F. A. Bovey, P. A. Mirau, H. S. Gutowsky, *Nuclear Magnetic Resonance Spectroscopy.* Elsevier, 1988.

14. J. R. Hanson, Labelling Compounds with the Stable Isotopes of Nitrogen and Oxygen. In *The Organic Chemistry of Isotopic Labelling*, RSC. Publishing, 2011; pp 110-121

15. J. D. Neufeld, M. Wagner, J. C. Murrell, Who eats what, where and when? Isotope labelling experiments are coming of age. *ISME J.* **2007**, *1*, 103-110.

16. C. Nyguen, R.W. Haushalter, D.J. Lee, P.R.L. Marwick, J. Bruegger, G. Caldara-Festin, K. Finzel, D.R. Jackson, F. Ishikawa, B. O'Dowd, J.A. McCammon, S.J. Opella, S.C. Tsai, M.D. Burkart, Trapping the dynamic acyl carrier protein in fatty acid biosynthesis. *Nature*, **2014**, *505*, 427-431.

17. M. J. Smith, C. B. Marshall, F. X. Theillet, A. Binolfi, P. Selenko, M. Ikura, Real-time NMR monitoring of biological activities in complex physiological environments. *Curr. Opin. Struct. Biol.* **2015**, *32*, 39-47.

18. P. Vallurupalli, G. Bouvignes, L. E. Kay, Studying "invisible" excited protein states in slow exchange with a major state conformation. *J Am. Chem. Soc.*, **2012**, *134*, 8148-8161.

Chapter 1 Chemoenzymatic Elaboration of the Raper-Mason Pathway Unravels the Structural Diversity Within Eumelanin Pigments

1.1 Introduction

Although the heterogeneous biopolymer melanin[1] is ubiquitous in life, there remains a lack in understanding of its intricate molecular structure.[2] This structure diverges between three main types of melanin. Eumelanin,[3] the most common form, is a black/brown pigment derived from tyrosine and is associated with melanoma.[4] Pheomelanin,[5] the reddish cysteine derivative can be found in hair and skin. Neuromelanin, is a dark pigment located in the mammalian brain, which studies have shown to be linked to Parkinson's disease.[6] Melanin also has a broad electromagnetic absorption range and plays a role in structural colour.[7] These properties are all incredibly useful to harness and emulate in the synthesis of new materials. Although the different melanin types share similar properties,[8] they are believed to vary significantly in their molecular composition.[9,10] Despite melanin's ubiquity, significant structural uncertainties still remain,[11] and a major challenge lies in the unique connectivity and isomeric potential in these polymers. As illustrated by **1**-**4** (Fig. 1.1), eumelanin is composed of multiple linkages of discrete monomers. Using a minimal 'dimeric' unit, the structural diversity in eumelanin can include disparity in both regioselectivity and connectivity. The complexity becomes even greater in pheomelanin **5** (Fig. 1.1) which contains multiple structurally discrete units.

The current understanding of melanin[12] is based on the generally accepted biosynthetic pathway for polymerization, the Raper-Mason pathway (RMP),[13-14] that was first proposed 70 years ago. The pathway involves a stepwise progression from L-Tyr to dihydroxyindole (DHI) or dihydroxyindole. carboxylic acid (DHICA), as shown in Scheme 1. These units then undergo an oxidation to generate IQCA or IQ, the immediate precursors for polymerization. Our findings here show that the general understanding and interpretation of the RMP is incorrect. In addition to the

various structural complexities, melanogenesis processes are regulated *in vivo* through a series of enzymes, including tyrosinases[15] and decarboxylases,[16] that unite the oxidative and decarboxylative steps shown in Scheme 1.1. However, our findings suggest that each RMP step need not reach completion prior to polymerization, and heterogeneity may be generated by the relative abundance of pathway intermediates. Here, we explore the link between tyrosinase substrate selectivity and structural output by exploring RMP intermediate identity and structural output in melanisation.

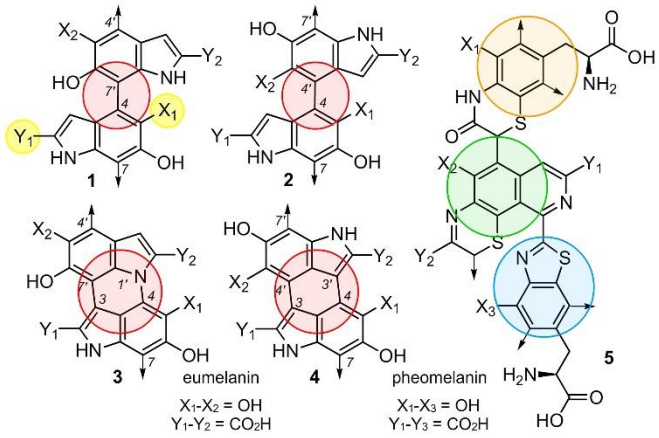
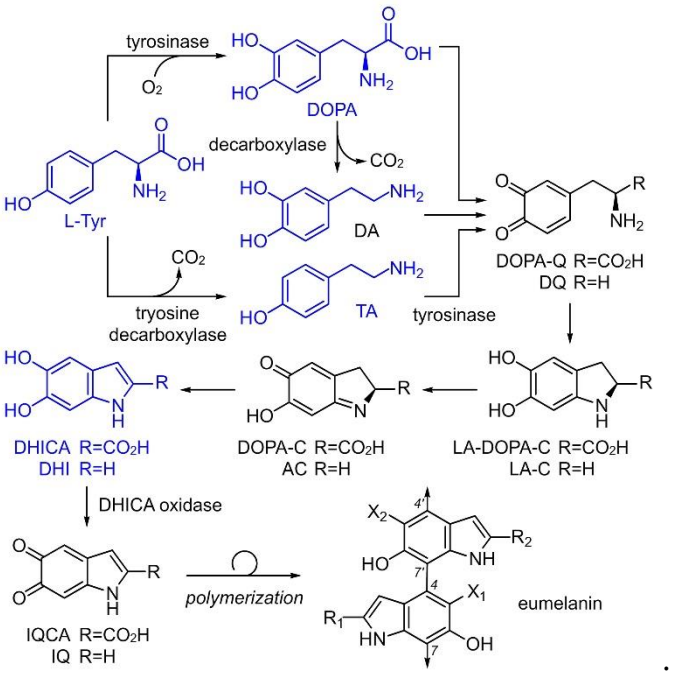


Figure 1.1: Eumelanin presents multiple levels of structural diversity. The first, **1**, arises through different levels of oxidation represented by X1 and Y1 (yellow). The second arises in the monomers connectivity, as noted by the comparison of **1** to **2** (red). The third is the degree of monomer connectivity, e.g. mono-linked **1**/**2** to di-linked **3**/**4**. Further complexity can come from different functional linkages, as shown in pheomelanin **5** (orange, green and blue).

To date, our structural understanding of melanin arises from pioneering solid-state NMR (ssNMR) work on fungal melanin by Stark and Casadevall.[17,18] These datasets have enabled a hypothesis toward the linkages between DHI and DHICA. Typically, melanin ssNMR spectra are assigned using chemical shift predictions.[19] While this method was vital to clarifying the key functionality, the lack of expanded experimental data to support these predictions reduces the accuracy of their assignments.

We began by exploring the ability of the well-characterized mushroom tyrosinase from *Agaricus bisporus*[20] for polymerization of different RMP intermediates. Unlike previous studies where the polymerized pigments are treated with 6 M HCl and boiled, which can destroy essential structural features, our pigments were carefully washed with neutral H_2O. The tools of ssNMR and dynamic nuclear polarization (DNP)[21,22] were applied to evaluate the four critical monomers tyrosine (L-Tyr), tyramine (TA), dopamine (DA),[23] and L-DOPA (Supporting Figs. S1.1-S1.7) of the RMP (Scheme 1.1) in their monomeric and polymeric forms.

In this study we used both ^{13}C chemical shift predictions generated by ChemNMR (Supporting Fig. S1.6) and ssNMR spectra on the monomers to assign the structures of each polymer (Fig. 1.2-1.4). Likewise, corresponding scanning electron microscopy (SEM) images of these pigments provide insight into the different morphologies that can possibly arise from these nanoparticles. From this, one can appreciate the vast differences between melanin pigments at the nanoscale and how it can attribute to different biological functions.[24]



Scheme 1.1: The Raper-Mason pathway (RMP) proposes that melanin originates from L-Tyr being oxidized to melanin via L-DOPA, which could be followed by a decarboxylation to dopamine (DA). L-Tyr can also be decarboxylated to tyramine (TA). These substrates are further oxidized to form transient intermediates: dopaquinone (DOPA-Q), dopamine-o-quinone (DQ), leukodopachrome (LA-DOPA-C), leukoaminochrome (LA-C), dopachrome (DOPA-C), and aminochrome (AC), leading to DHI and DHICA units, the proposed building blocks of eumelanin. Prior to polymerization, DHI and DHICA are further oxidized to indole-5,6-quinone (IQ) and indole-5,6-quinone-2-carboxylic acid (IQCA).

L-Tyr-melanin. Starting with L-Tyr (Fig. 1.2a), a SEM image shows that a black pigment containing rods imbedded within a solid matrix was obtained after treating L-Tyr with tyrosinase. The ssNMR ^{13}C spectrum of the L-Tyr was evaluated against the L-Tyr-melanin. Interestingly, spectral overlays indicate that each of the ^{13}C peaks in the L-Tyr monomer (grey, Fig. 1.2a) were observed in the L-Tyr-melanin product (black, Fig. 1.2a). With optimized magic-angle spinning NMR cross-polarization experimental conditions for both carbonyl and aliphatics, the lower aliphatic α,β peaks at δ 38 and 53 ppm suggest a con-version of the alkyl carbons into an aromatic environment. Meanwhile, a set of peaks with low intensity was detected at δ 105 and 145 ppm, indicative of DHI and DHICA (Scheme 1.1). The three peaks (δ 160, 135 and 42 ppm) within the

^{15}N spectrum suggested N atoms commonly associated with pyrroles, indoles, and amines, respectively.

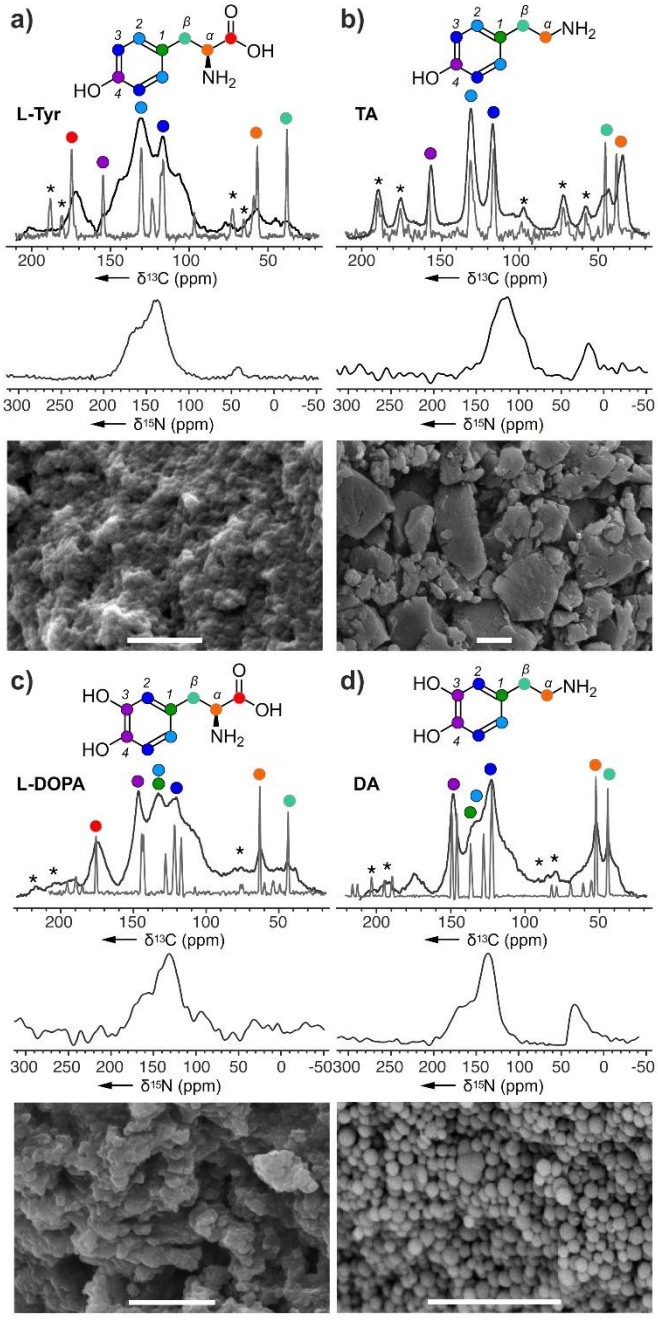


Figure 1.2: Melanin polymers were obtained from the action of *A. bisporus* tyrosinase on **a)** L-Tyr, **b)** TA, **c)** L-DOPA, and **d)** DA. ssNMR spectra of each melanised sample are overlaid with its original monomer, and these peaks were used for assignments. Only the ^{15}N spectra were collected using DNP ssNMR with an enhancement ~20. Peak assignments are provided in the

Supporting Figs. S1.1-S1.7. * denotes the spinning side bands. Bars in the SEM images represent 1 μm.

TA-melanin. The same method was then applied to tyramine (TA), another RMP intermediate suggested to require processing before polymerization. The SEM image of TA-melanin (Fig. 1.2b) illustrates a polymer lacking nanostructural features. ssNMR analyses revealed a clear correlation between TA-monomer and TA-melanin (Supporting Fig. S1.2). As evident in Fig. 1.2b, the carbon atoms underwent only a modest chemical environment change upon polymerization. The spectral data from TA-melanin was in stark contrast with L-Tyr-melanin, whose broad peaks suggested it contained a nonuniform melanin composition through processes such as incomplete decarboxylation, resulting in mixtures of decarboxylated (R_1,R_2=H) and carboxylated (R_1,R_2=CO_2H) polymers. TA-melanin has an upfield ^{15}N chemical shift of δ 155 ppm resembling peptide linkages. These observations reveal that the action of the mushroom tyrosinase on TA was not restricted to comply with the stepwise nature of the RMP (Scheme 1) but was able to process an expanded monomeric scope.

L-DOPA and DA melanin. To complete the RMP intermediate set, L-DOPA-melanin and DA-melanin spectra were evaluated (Fig. 1.2c-1.2d, Supporting Fig. S1.3-S1.4). As shown in Fig. 1.2c, the polymerization of L-DOPA returned a polymer comprised of 100±30 nm sized spheres. Here, the presence of an additional hydroxyl group with respect to L-Tyr (Fig. 1.2a) resulted in a nanostructural change from a material with imbedded surfaces to more geometrically defined nanometer-sized spheres in L-DOPA-melanin. The ssNMR spectrum from L-DOPA-melanin (Fig. 1.3c), like L-Tyr-melanin (Fig. 1.3a), contained both the aromatic and aliphatic peaks of its monomer. Similarly, the presence of the carbonyl at δ 173 ppm, along with α and β carbons at δ 65 and 45 ppm, in L-DOPA-melanin confirmed incomplete decarboxylation and indole ring formation. Meanwhile, the intensity of the aliphatic carbons at δ 53 and 47 ppm from DA-melanin

were the strongest of all three polymers, suggesting an increased level of non-indole cyclised materials. This was further supported by the presence of a δ 45 ppm peak in the ^{15}N spectrum. The dopamine-formed particles produced visible 100±10 nm spheres by SEM imaging.

All four spectral sets (Fig. 1.2) showed that chemical shift values of the melanised particles were similar to their monomeric precursors with the addition of the cyclised DHI or DHICA units. The fact that all four monomers melanised shows the broad selectiveness of the enzyme and the structural diversity arising from the metabolites. This alters the mechanistic model for melanin formation and impacts the resulting nanoparticle structure. SEM images of the four melanin products with different nanoparticle morphologies illustrates each metabolite plays a role in influencing or altering the functional capabilities of melanin. The ssNMR spectra also shows that the polymer is mostly populated by the monomer itself, but it includes a small amount of cyclised DHI or DHICA. Among them, TA-melanin has the most resolved aromatic region in its ^{13}C spectrum and the most upfield ^{15}N chemical shift at δ 115 ppm which could arise from other types of linkages. All of the melanin polymers in Fig. 1.2 retained primary amine residues at δ 35 ppm or 25 ppm.

Hydroxyindole-melanins. Next, we explored the ability of the *A. bisporus* tyrosinase to polymerize indole derivatives of DHI or DHICA (Scheme 1.1). The goal of this study was to demonstrate the ability of tyrosinase to act on non-natural substrates and further probe the substrate-tolerance of this pathway. Using conditions established in the previous study, the tyrosinase was presented with 4 different hydroxyindole (HI) monomers: 4-HI, 5-HI, 6-HI and 7-HI, and the resulting melanins were analysed via ssNMR and SEM (Supporting Figs. S1.8-S1.11). Comparable to the acyclic precursors (Fig. 1.2), the ^{13}C spectra of HI-melanins contained similar chemical shift values to their corresponding monomers. Based on the peak positions, 4-HI-melanin

contained either a C5 or C7 downfield shift from δ 105 to 115 ppm, supporting a 5-5 linkage (Fig. 1.3a). In contrast, the increase in the number of peaks in 5-HI-melanin (Fig. 1.3b), would result from a 2-4 heterolinkage between the monomer units, which can be explained by a loss of symmetry and a duplication of its 7 carbons.

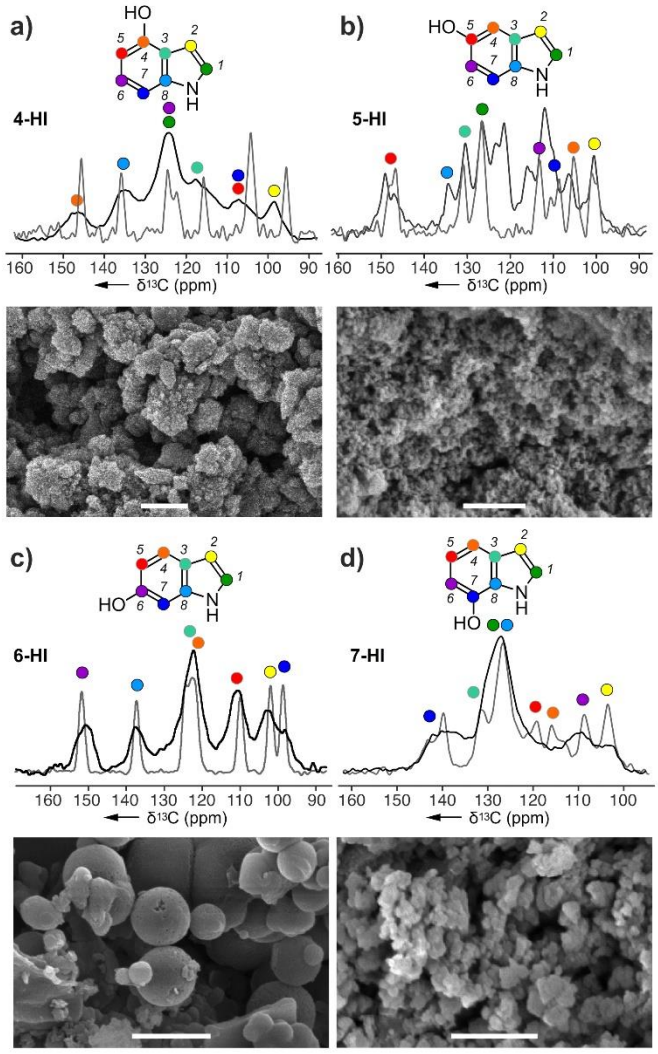


Figure 1.3: Melanin particles obtained from the action of A. bisporus tyrosinase on a) 4-HI, b) 5-HI, c) 6-HI, and d) 7-HI. ssNMR spectra of each melanised sample are overlaid with its original monomer, and these peaks were used for assignments. Bars in the SEM images represent 1 µm. Peak assignments are provided in Supporting Figs. S1.8-S1.11.

Each of the carbons observed in the ^{13}C-NMR spectrum of 6-HI-melanin (Fig. 1.3c) were identified in the 6-HI monomer, except for C3, which was weakly observed. 6-HI-melanin exhibited the most resolved spectra, indicating a more homogenous environment. Symmetry studies suggest 6-HI-melanin consists of a 2-2 linkage. Unfortunately, the overlap of three residues within the central peak at δ 126 ppm in the spectrum from 7-HI-melanin complicated the elucidation of its linkage (Fig. 1.3d).

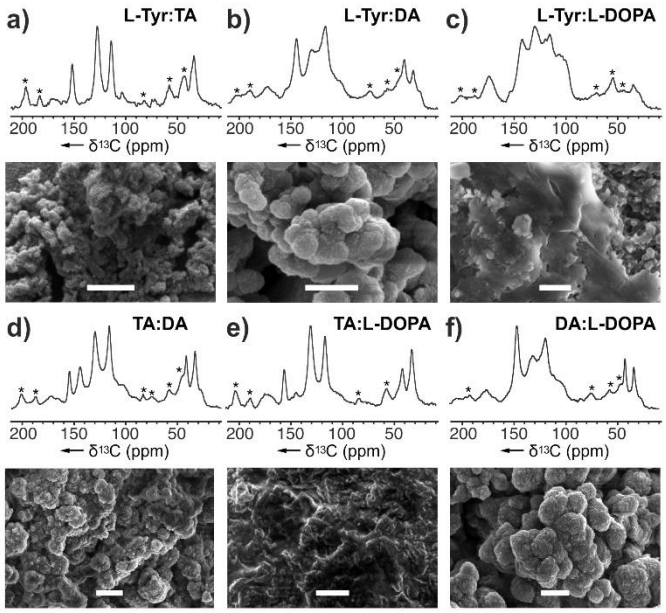


Figure 1.4: Melanin produced by copolymerizing two monomers: a) L-Tyr:TA melanin generated from a mixture of L-Tyr and TA, b) L-Tyr:L-DA melanin, c) L-Tyr:L-DOPA melanin, d) TA:DA melanin, e) TA:L-DOPA melanin, and f) DA:L-DOPA melanin using A. bisporus tyrosinase. Comparison of the NMR spectra of these mixtures with that of pure monomers (Fig. 1.2) was used to explore the monomer selectivity of this enzyme. Peak and monomer assignments are provided in Supporting Fig. S1.12-S1.13 as well as the experimental conditions of copolymerization procedure. * denotes the spinning side bands. Bars in the SEM images represent 1 μm.

4-HI-melanin (Fig. 1.3a) and 7-HI-melanin (Fig. 1.3d) SEM images resembled L-DOPA-melanin with clusters of nanoparticle sizes of 100±25 nm. SEM imaging revealed 5-HI-melanin contained dispersed particles with a high size uniformity of 75±15 nm while 6-HI-melanin formed spherical particles with a large size distribution, 350-900 nm (Fig. 1.3c). Though the proposed

natural intermediate, DHI, (Scheme 1.1) contains oxidation at C5 and C6, the substrate tolerance of the tyrosinase enabled it to polymerize materials with oxidations at non-natural positions.

Copolymerized melanin blends. To investigate its selectivity, the tyrosinase was used to copolymerize various combinations of monomers. The resulting melanin products along with their monomer counterparts were analysed by ssNMR and SEM (Fig. 1.4 and Supporting Figs. S1.12-1.13). Among the six copolymers, the ssNMR spectra suggested the resulting melanin was not of 1:1 alternating pattern; rather, TA took precedence over the others. Though the overall lineshape follows TA, it is clear that portions of the peaks are derived from the other monomer. Additionally, a set of peaks with distinct DHI or DHICA chemical shifts appeared at low intensity. Furthermore, we observed that DA took precedence over L-DOPA, which took precedence over L-Tyr. Overall, this suggests the polymers generated from the tyrosinase can incorporate monomers at different stages of the RMP and does not require a stepwise conversion as shown in Scheme 1.1. Instead, the *A. bisporus* tyrosinase can polymerize a 'snapshot' of the monomeric metabolic pool within its local environment. This observation begins to explain the complexity seen within melanin polymers and provides an important advance in understanding the structural diversity accessible in eumelanin.

The SEM images from these blends (Fig. 1.4) show particles of various morphologies. In particular, the ones containing DA yield products increased sphericity and higher ordered structures. While DA-melanised nanoparticles are around 100 nm, once it is copolymerized with other monomers, the resulting nanoparticles increase in size. We observed 100 nm particles molded into ovoid objects ranging from 500 nm to 1 μm.

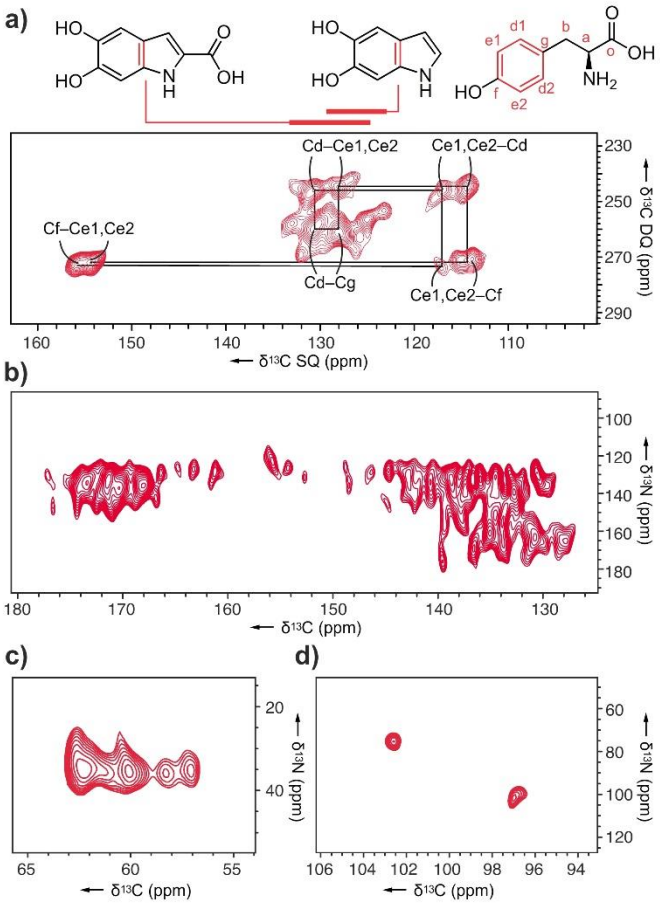


Figure 1.5: DNP ssNMR analysis of U-[^{13}C,^{15}N]-L-Tyr-melanin. **a)** 2D ^{13}C-^{13}C SPC5 spectrum showing correlations from the tyrosine ring, DHI, and DHICA units**. b)** 2D ^{13}C-^{15}N TEDOR spectra acquired with mixing times of 3.2 ms showing pyrrole correlations show only aromatic contacts and no carboxylated motifs. **c)** The aliphatic region shows four distinct Ca-N environments. d) C4-N correlations from DHICA units were detected.

Next, we explored melanin prepared from U-[^{13}C,^{15}N]-L-Tyr melanin using multidimensional DNP NMR (Fig. 1.5). Correlations observed from the 2D ^{13}C-^{13}C PDSD experiments[25] indicated high populations of intramolecular contacts (Supporting Fig. S1.14). However, this doesn't exclude the possibility of intermolecular interactions between tyrosine units, in the case of pi-pi stacking between monomers.[26] The 2D ^{13}C-^{13}C SPC5 experiment[27] (Fig. 1.5a) shows covalent bonds from both the L-Tyr ring and DHI/DHICA units. This further validates the highly populated presence of L-Tyr monomer in this multicomponent melanised polymer. The

unique chemical shift values of the tyrosine aromatic ring are clearly assigned, starting with Cf (Fig. 1.5a) attached to the hydroxyl group at δ 145 ppm to the double resonances of Ce1, Ce2 at δ 115 ppm to another set of double resonances Cd1,Cd2 at δ 130 ppm, and lastly to Cg at δ 130 ppm.

In the ^{13}C-^{15}N TEDOR spectra[28] (Figs. 1.5b-1.5d), nitrogen atoms from hydroxyindole units correlate with both aromatic and carbonyl carbons, suggesting DHI and DHICA units. Pyrrole correlations show only aromatic contacts and no carboxylated motifs, which could result from Baeyer-Villiger type reactions originating from indole metabolites of the RMP. The primary amine ^{15}N is correlated to four distinct ^{13}C atoms, with the two major peaks originating from L-Tyr.

1.2 Conclusions

This study suggests the classical RMP doesn't fully-encompass substrate scope in melanin biosynthesis, as enzymatic polymerization can occur with many of its monomeric intermediates (blue structures, Scheme 1.1). Here, comparative ssNMR evaluations of monomers (Figs. 1.2-1.3) provide a useful tool that enables a nuanced understanding of the structural features of each melanin polymer. While copolymerization studies (Fig. 1.4) indicate some selectivity, the serial nature of the RMP can be misleading. These studies suggest that natural melanin structure likely relies on the local metabolic flux[29] to guide the structural composition of the resulting polymers. Overall, tandem SEM and ssNMR analyses enable the development of a structure-function correlation for melanin pigments. Demonstrated here for the *A. bisporus* tyrosinase, this chemoenzymatic approach offers a general tool to explore the structural selectivity and tolerance of melanin-producing enzymes. Efforts are now underway to apply this union between ssNMR and electron microscopy to further interrogate the diversity of natural melanin.

1.3 Supporting Information

A. General methods. Chemical reagents and enzymes were purchased from Worthington, Acros, Fluka, Sigma-Aldrich, or TCI. Deuterated NMR solvents were purchased from Cambridge Isotope Laboratories. All reactions were performed with stirring from Teflon coated stir bars using an IKAMAG RCT-basic mechanical stirrer (IKA GmbH). 1H NMR spectra were recorded on 750 MHz Magnex Scientific magnet and Bruker spectrometer. Chemical shifts for 13C NMR were referenced to the reported values of adamantine at 38.5 and 29.5 ppm. Chemical shift δ values for 1H and 13C spectra are reported in parts per million (ppm) relative to these referenced values. ChemDraw was utilized for all chemical shift predictions. All 13C NMR spectra were recorded with complete proton decoupling. FID files were processed using TopSpin 3.2 (Bruker) and MestreNova 6.0.2 (Mestrelab Research). Scanning Electron Microscopy (SEM) analyses were performed using a FEI Quanta FEG 250 microscope (ThermoFisher Scientific) using a 3-5 kV energy source and a spot size of 3 under high vacuum at a working distance of 9 mm. Spectral data and procedures are provided for all new compounds and copies of select spectra have been provided.

B. General procedure for melanin polymerization. All precursor monomers were dissolved in 5 mM phosphate-buffered saline pH 7.5. Powdered mushroom tyrosinase I.U.B.1.14.18.1 (Worthington Biochemical) formula weight 128 kDa was added to the solution and stirred at 500-600 rpm for 24 h to 48 h at 23 °C. Solution color progressed from clear to pale yellow, followed by orange/brown and led to the final black solution.

C. General procedure for melanin isolation. The black solution was centrifuged at 4,500 rpm in MultifugeX3R (ThermoFisher Scientific), the supernatant was discarded, and the pellet was resuspended in H2O (50 mL). The process was repeated 3-4 times until the supernatant became clear. The final pellet was then placed in a 1.9 mL tube (Eppendorf) to centrifuge at 13,000 rpm

PrismR (Carolina Biological Supply Company). The supernatant was discarded and sample was dried by evaporation or via speed vacuum to remove residual water.

D. Polymerization reactions from single monomers. The following sections provide the methods for the preparation of each of the melanin samples in Figures 2-3.

L-Tyr-melanin. The reaction was conducted using the general procedure for melanin polymerization (section B) using L-tyrosine (L-Tyr, 1.0 g, 5.5 mmol) and mushroom tyrosinase (5.0 mg, 39.0 nmol) in 1.1 L of 5 mM phosphate-buffered saline pH 7.5. The solution was isolated using the general procedure (section C) to yield 109 mg (11%) of L-Tyr-melanin.

TA-melanin. The reaction was conducted using the general procedure for melanin polymerization (section B) using tyramine (TA, 754.5 mg, 5.5 mmol) and mushroom tyrosinase (5.0 mg, 39.0 nmol) in 1.1 L of 5 mM phosphate-buffered saline pH 7.5. The solution was isolated using the general procedure (section C) to yield 73 mg (10%) of TA-melanin.

L-DOPA-melanin. The reaction was conducted using the general procedure for melanin polymerization (section B) using L-DOPA (L-DOPA, 1.0 g, 5.1 mmol) and mushroom tyrosinase (5.0 mg, 39.0 nmol) in 1.1 L of 5 mM phosphate-buffered saline pH 7.5. The solution was isolated using the general procedure (section C) to yield 612 mg (61%) of L-DOPA-melanin.

DA-melanin. The reaction was conducted using the general procedure for melanin (section B) using dopamine • HCl (DA, 1.0 g, 5.1 mmol) and mushroom tyrosinase (5.0 mg, 39.0 nmol) in 1.1 L of 5 mM phosphate-buffered saline pH 7.5. The solution was isolated using the general procedure (section C) to yield 465 mg (46%) of DA-melanin.

4-HI-melanin. The reaction was conducted using the general procedure for melanin polymerization (section B) using 4-hydroxyindole (4-HI, 45.0 mg, 0.34 mmol) and mushroom

tyrosinase (2.0 mg, 15.6 nmol) in 100 mL of 5 mM phosphate-buffered saline pH 7.5. The solution was isolated using the general procedure (section C) to yield 16 mg (36%) of 4-HI-melanin.

5-HI-melanin. The reaction was conducted using the general procedure for melanin polymerization (section B) using 5-hydroxyindole (5-HI, 53.0 mg, 0.40 mmol) and mushroom tyrosinase (2.0 mg, 15.6 nmol) in 100 mL of 5 mM phosphate-buffered saline pH 7.5. The solution was isolated using the general procedure (section C) to yield 8.4 mg (16%) of 5-HI-melanin.

6-HI-melanin. The reaction was conducted using the general procedure for melanin polymerization (section B) using 6-hydroxyindole (6-HI, 45.0 mg, 0.34 mmol) and mushroom tyrosinase (2.0 mg, 15.6 nmol) in 100 mL of 5 mM phosphate-buffered saline pH 7.5. The solution was isolated using the general procedure (section C) to yield 21 mg (47%) of 6-HI-melanin.

7-HI-melanin. The reaction was conducted using the general procedure for melanin polymerization (section B) using 7-hydroxyindole (7-HI, 45.0 mg, 0.34 mmol) and mushroom tyrosinase (2.0 mg, 15.6 nmol) in 100 mL of 5 mM phosphate-buffered saline pH 7.5. The solution was isolated using the general procedure (section C) to yield 17 mg (38%) of 7-HI-melanin.

E. Polymerization reactions from monomer mixtures. The following sections provide the methods for the preparation of each of the melanin samples in Figure 4.

Melanized L-Tyr:TA. The reaction was conducted using the general procedure for melanin polymerization (section B) with L-Tyr (100.0 mg, 0.55 mmol), TA (100.0 mg, 0.73 mmol), and mushroom tyrosinase (2.0 mg, 15.6 nmol) in 200 mL of 5 mM phosphate-buffered saline pH 7.5. The solution was isolated using the general procedure (section C) to yield 11 mg (11%) of a L-Tyr/TA- mixed melanin polymer.

Melanized L-Tyr:L-DOPA. The reaction was conducted using the general procedure for melanin polymerization (section B) with L-Tyr (100.0 mg, 0.55 mmol), L-DOPA (100 mg, 0.51

mmol), and mushroom tyrosinase (2.0 mg, 15.6 nmol) in 200 mL of 5 mM phosphate-buffered saline pH 7.5. The solution was isolated using the general procedure (section C) to yield 53 mg (27%) of a L-Tyr/L-DOPA-mixed melanin polymer.

Melanized L-Tyr:DA. The reaction was conducted using the general procedure for melanin polymerization (section B) with L-Tyr (100.0 mg, 0.55 mmol), dopamine • HCl (100.0 mg, 0.53 mmol), and mushroom tyrosinase (2.0 mg, 15.6 nmol) in 200 mL of 5 mM phosphate-buffered saline pH 7.5. The solution was isolated using the general procedure (section C) to yield 38 mg (19%) of a L-Tyr/DA-mixed melanin polymer.

Melanized TA:L-DOPA. The reaction was conducted using the general procedure for melanin polymerization (section B) with TA (100.0 mg, 0.73 mmol), L-DOPA (100.0 mg, 0.51 mmol), and mushroom tyrosinase (2.0 mg, 15.6 nmol) in 200 mL of 5 mM phosphate-buffered saline pH 7.5. The solution was isolated using the general procedure (section C) to yield 118 mg (59%) of a TA/L-DOPA-mixed melanin polymer.

Melanized TA:DA. The reaction was conducted using the general procedure for melanin polymerization (section B) with TA (100.0 mg, 0.55 mmol), dopamine • HCl (100.0 mg, 0.53 mmol), and mushroom tyrosinase (2.0 mg, 15.6 nmol) in 200 mL of 5 mM phosphate-buffered saline pH 7.5. The solution was isolated using the general procedure (section C) to yield 71 mg (36%) of a TA/DA-mixed melanin polymer.

Melanized L-DOPA:DA. The reaction was conducted using the general procedure for melanin polymerization (section B) with L-DOPA (100.0 mg, 0.51 mmol), dopamine • HCl (100.0 mg, 0.53 mmol), and mushroom tyrosinase (2.0 mg, 15.6 nmol) in 200 mL of 5 mM phosphate-buffered saline pH 7.5. The solution was isolated using the general procedure (section C) to yield 156 mg (78%) of a L-DOPA/DA-mixed melanin polymer.

F. Polymerization reactions from isotopically-labeled monomers. The following sections provide the methods for the preparation of each of the melanin samples in Figure 5.

U-13C,15N L-Tyr-melanin. The reaction was conducted using the general procedure for melanin polymerization (section B) with U-13C,15N -L-tyrosine (100.0 mg, 0.52 mmol) and mushroom tyrosinase (2.0 mg, 15.6 nmol) in 110 mL of 5 mM phosphate-buffered saline pH 7.5. The solution was isolated using the general procedure (section C) to yield ¬35 mg (35%) of U-13C,15N-Tyr-melanin.

G. Solid state NMR (ss-NMR) spectroscopy. Dried natural abundance melanin sample (30 mg) was packed in 3.2 mm Bruker rotor. 1D 13C CP spectra of the precursor monomers were acquired with 32 to 512 scans and a recycle delay of 3 s. Samples required 9k to 64k scans and were acquired with a recycle delay of 3 s. The CP contain times were optimized to 2 ms, with a 1H decoupling of 83 kHz and a spinning frequency of 13.5 kHz. Data was acquired on a 750 1H MHz spectrometer (Bruker). Spinning sidebands from the NMR are labeled in the spectra and they don't overlap with peaks of interest. When there is uncertainty, we ran the CP experiment at multiple spinning frequencies to be certain of our peaks.

Note: The intrinsic heterogeneous nature of melanin biopolymer results in broad linewidth in the NMR spectra, which can't be improved with higher field, spinning or temperature of the experiments. Cross polarization experiment was chosen over spectral editing or direct polarization experiments due do to the fact that it is not useful/applicable in what we want to achieve and also not feasible due to the long acquisition time of the experiments.

H. Dynamic nuclear polarization (DNP) solid-state NMR spectroscopy. Dried melanin sample (30mg) was mixed with 20 mM AMUPol polarizing agent in d8-glycerol/D2O/H2O (60/30/10 v/v). The resulting mixture was packed in a 3.2 mm sapphire rotor and all experiments

were acquired at 1H frequency of 600 MHz at 100 K. An enhancement factor of ~20 was obtained on the samples. The 1D 15N spectra of L-Tyr-melanin, TA-melanin, L-DOPA-melanin, and DA-melanin (Fig. 2) were acquired with 20k-36k scans. The 2D 13C-13C SPC5 spectrum of U-13C,15N-L-Tyr-melanin (Fig. 5) was acquired with a 3.5 s recycle delay, 512 scans per increment, and 64 t1 increments of 24 μs. The 2D 13C-15N TEDOR (Fig. 5) was acquired at 1.66 ms mixing time, 256 scans per increment, 64 t1 increments of 42 μs and a 3 s recycle delay. The 2D 13C-13C PDSD spectra (Supporting Fig. S13) was acquired at 10 ms, 50 ms, 100 ms, and 300 ms mixing times. Each PDSD was conducted with 256 scans, a 3.5 s recycle delay, and 128 t1 increments of 2 6μs.

I. Scanning electron microscopy (SEM). Dried melanin powder was placed on conductive carbon tape on an aluminium sample holder and coated using an iridium-sputter coating for 8 s. A FEI Quanta FEG 250 microscope was used for acquiring images using a 3-5 kV energy source with a spot size of 3 under high vacuum at a working distance of 9 mm.

J. Supporting Spectra. The following Figures provide supporting spectral analyses and assignments.

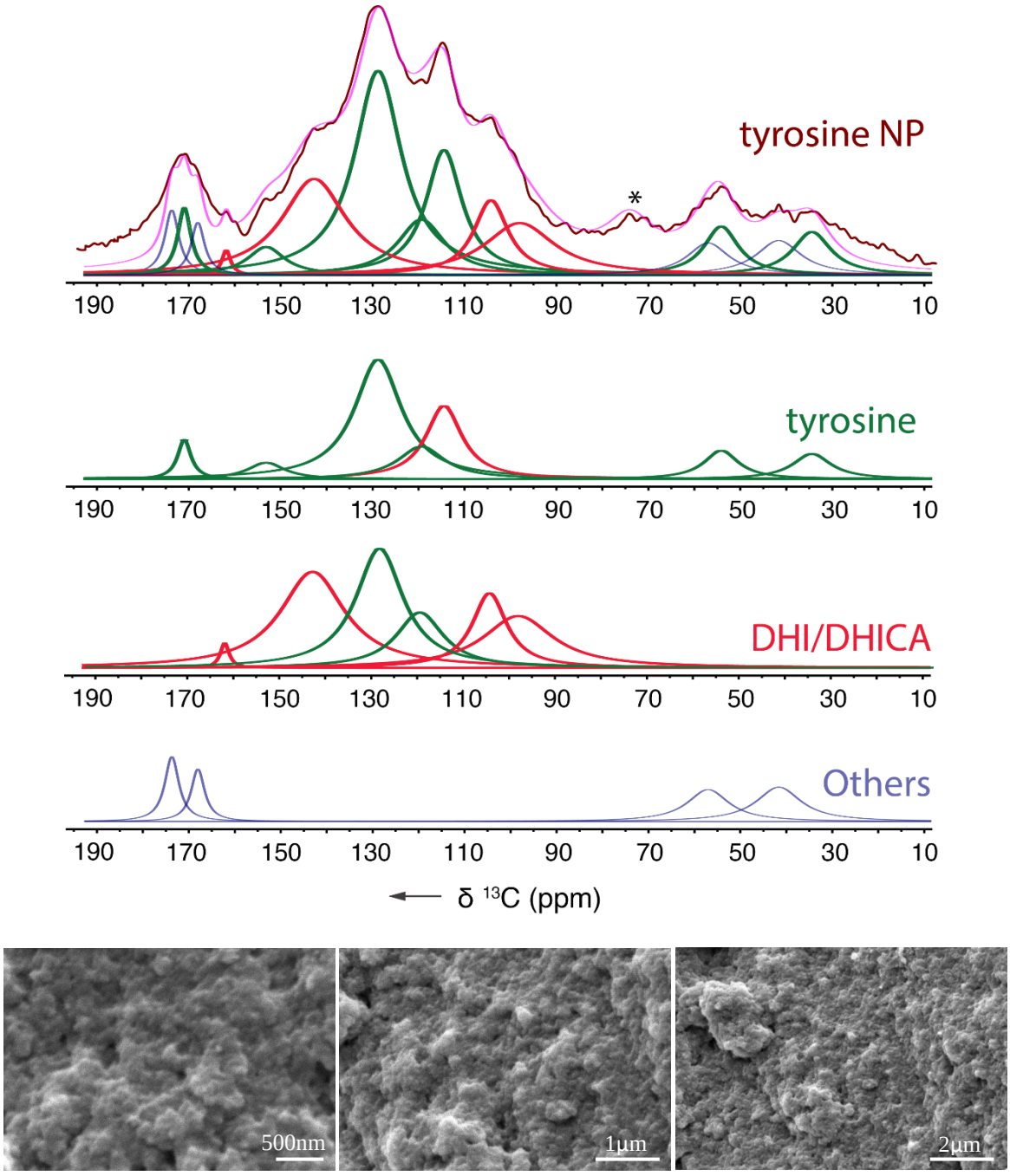


Figure S1.1: Spectral analysis of L-Tyr-melanin. (top) 1D ^{13}C spectrum of L-Tyr-melanin and deconvoluted peaks. There are at least three components that comprise L-Tyr-melanin. The first is peaks that follow pure tyrosine monomer units, second, are peaks containing chemical shift values that agree to those of DHI and DHICA units, and last, other components that have carbonyl and aliphatic moieties. (bottom) Select SEM images of L-Tyr-melanin at different magnifications.

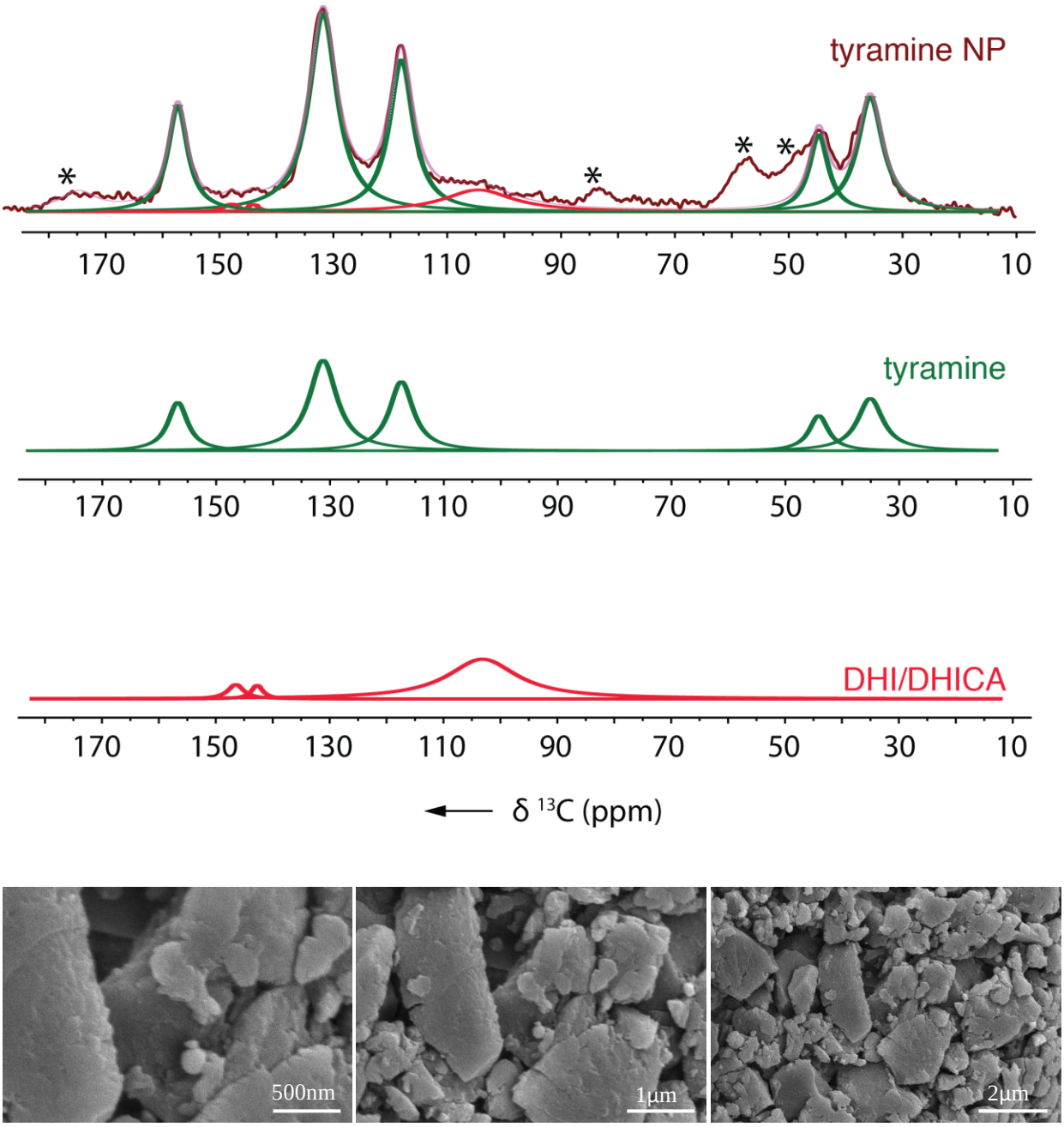


Figure S1.2: Spectral analysis of TA-melanin. (top) 1D ^{13}C of TA-melanin and its deconvoluted peaks. There are at least 2 components that comprise TA-melanin, peaks that follow pure tyramine monomer units and peaks containing chemical shift values that agree to those of DHI and DHICA units. (bottom) SEM images of TA-melanin at different magnifications.

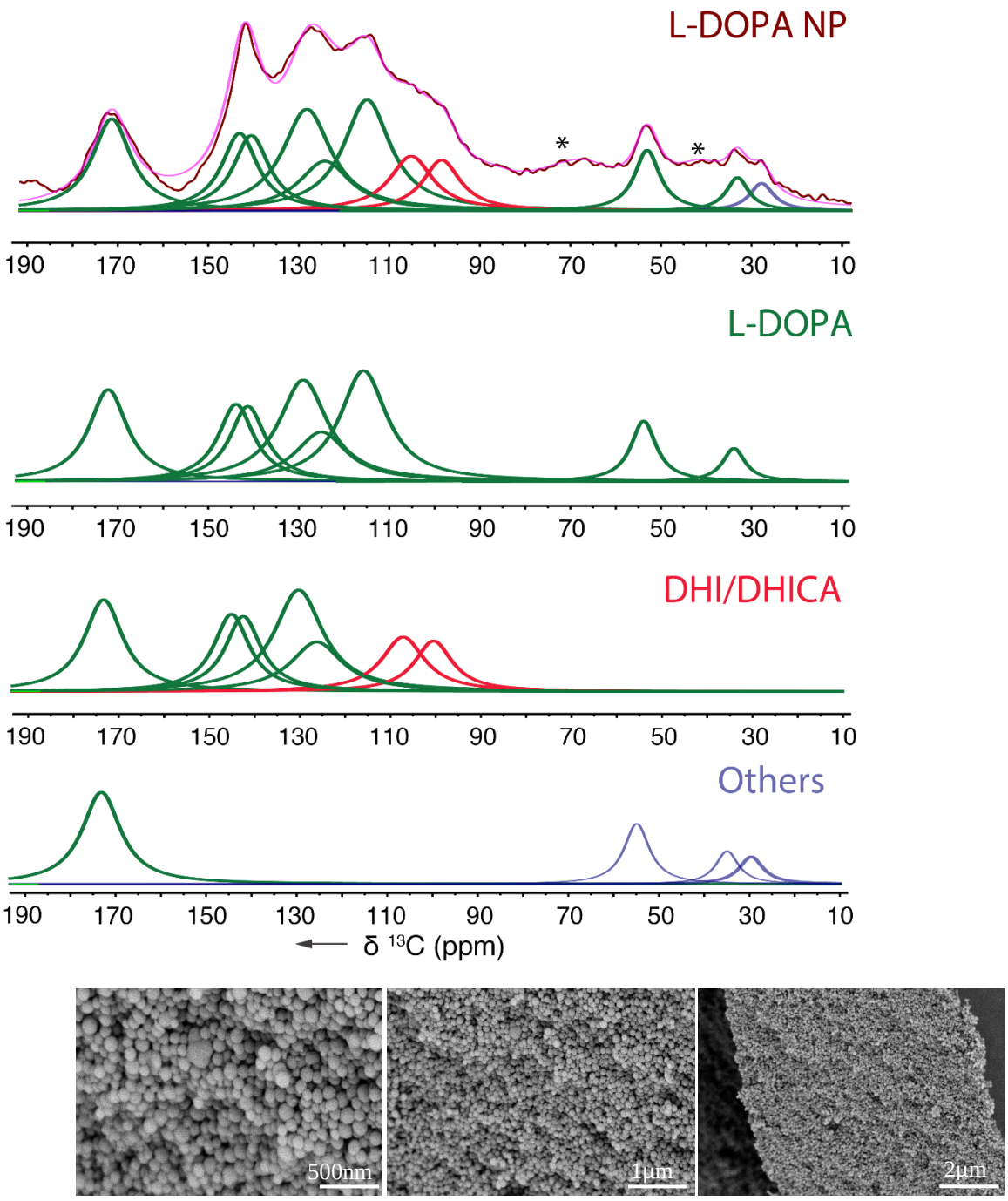


Figure S1.3: Spectral analysis of L-DOPA-melanin. (top) 1D ^{13}C of L-DOPA-melanin and its deconvoluted peaks. There are at least 3 components in L-DOPA-melanin, peaks that follow pure tyrosine monomer units, peaks containing chemical shift values that agree to those of DHI and DHICA units, and last, other components that have carbonyl and aliphatic moieties. (bottom) SEM images of L-DOPA-melanin at different magnifications.

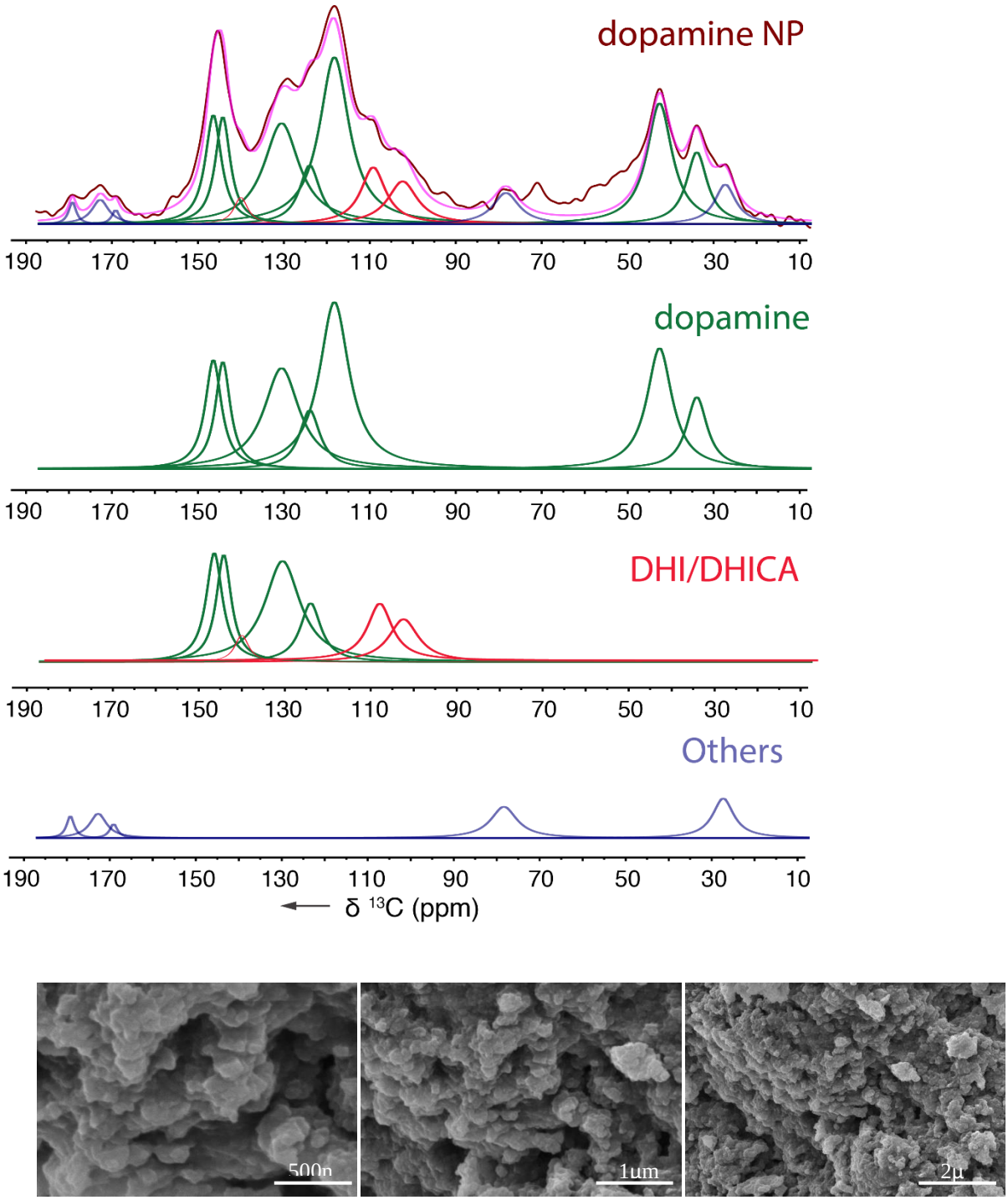


Figure S1.4: Spectral analysis of DA-melanin. (top) 1D ^{13}C of DA-melanin and its deconvoluted peaks. There are at least three components that comprise DA-melanin, peaks that follow pure tyrosine monomer units, peaks containing chemical shift values that agree to those of DHI and DHICA units, and last, other components with carbonyl and aliphatic moieties. (bottom) SEM images of DA-melanin at different magnifications.

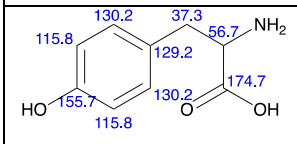
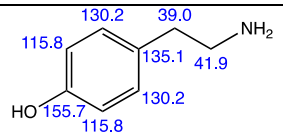
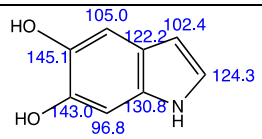
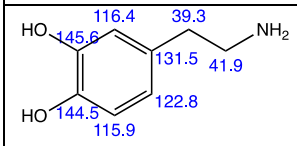
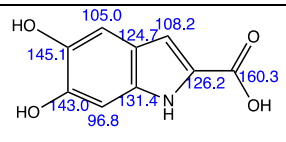
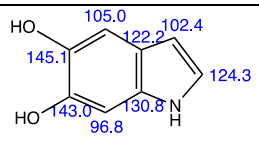
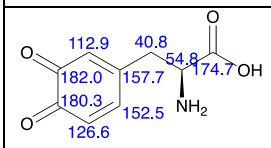
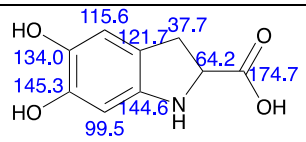
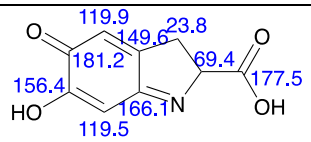
tyrosine	tyramine	L-dopa
		
dopamine	**DHICA**	**DHI**
		
dopaquinone	**leucochrome**	**dopachrome**
		

Figure S1.5: The ^{13}C chemical shift of every metabolite of the Raper-Mason pathway has been predicated by ChemDraw predictions and used to analyze our experimental data. Although the chemical values of some carbons on the aromatic rings of the metabolites are similar, each contains a set of unique chemical shift values that can be used to match with the experimental melanized chemical shift values.

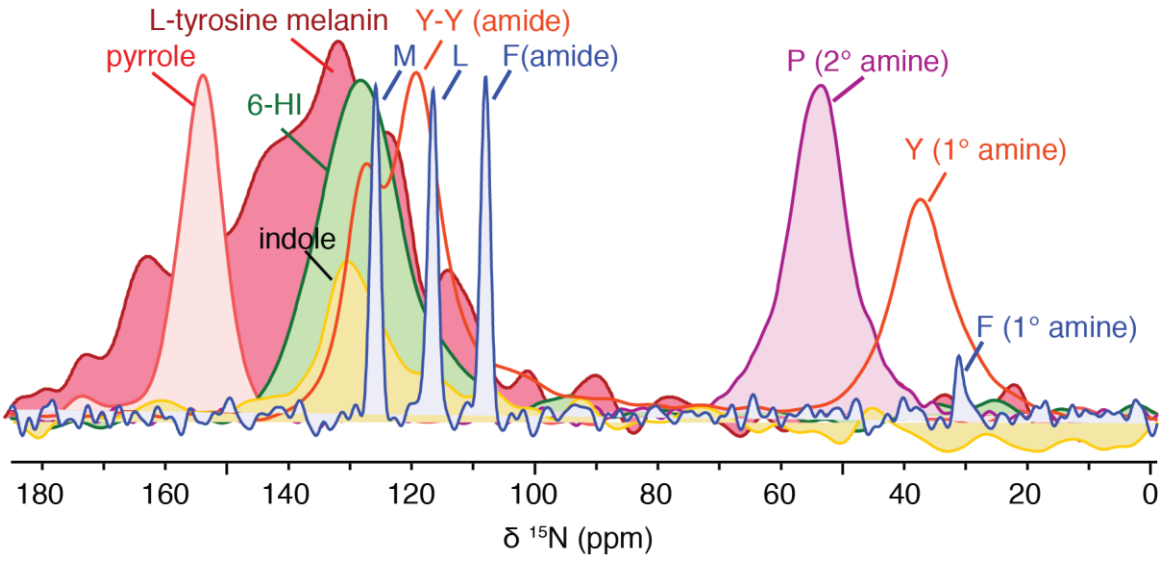


Figure S1.6: 1D ^{15}N CP DNP MAS NMR spectra of the nitrogen standards overlaid with melanized L-tyrosine. This is the method used in this manuscript to analyze the various possible nitrogen components in the melanized precursor monomers.

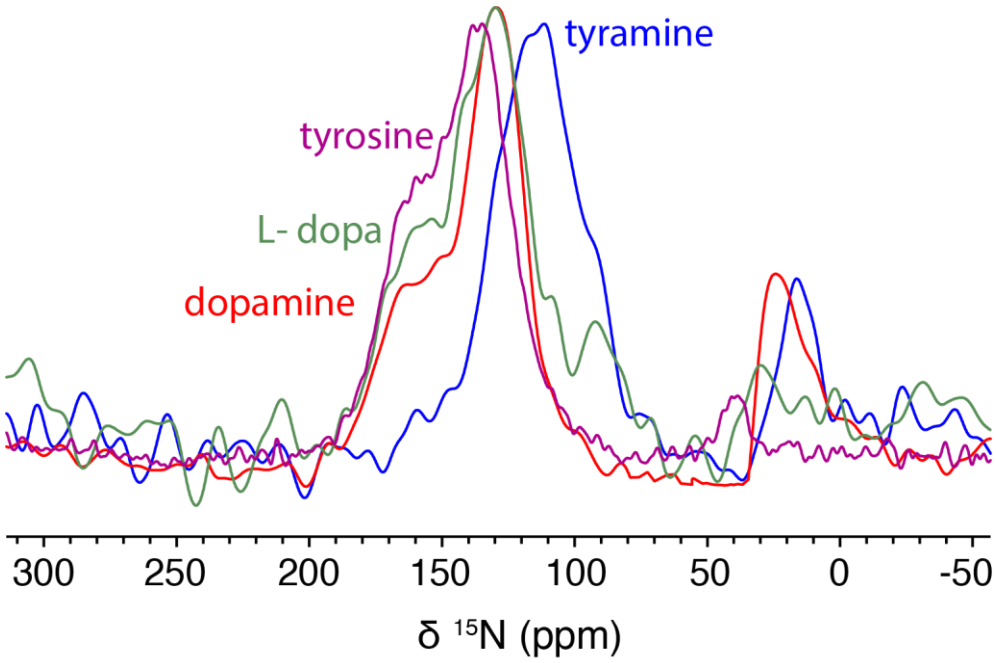
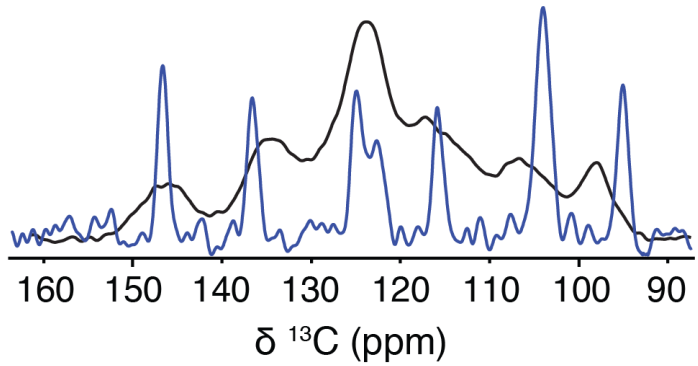
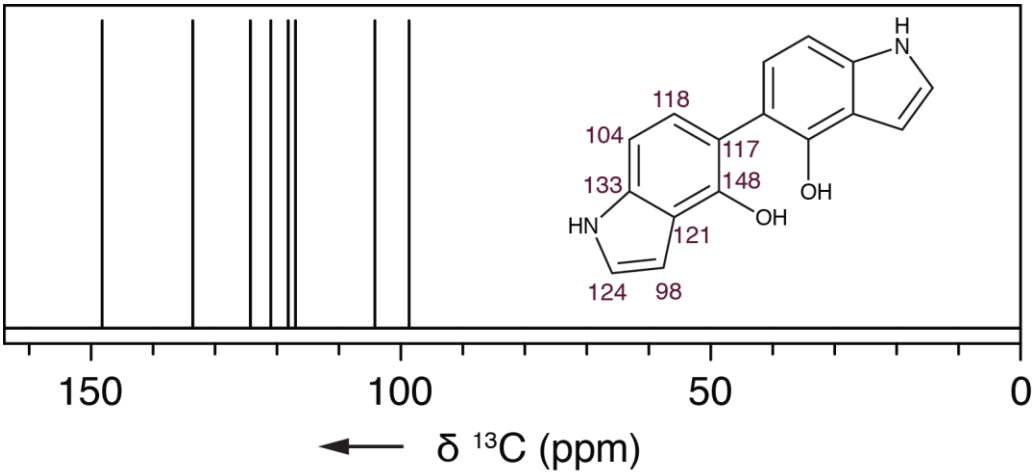


Figure S1.7: 1D ^{15}N CP DNP MAS NMR spectral overlay of L-Tyr-melanin, TA-melanin, L-DOPA-melanin, and DA-melanin. The peaks originate from pyrrole, indole, hydroxyindole, amide, and primary amine groups at ~δ 160 ppm, 135 ppm, 115 ppm and 22 ppm respectively. The chemical shifts of nitrogen standards are shown in Figure S1.5. Out of the four melanins, TA-melanin (blue) is the most upfield and is the only sample without a pyrrole shoulder peak at ~δ 160 ppm. All samples had a primary amine peak at ~δ 35 ± 10 ppm.



NMR predicted chemical shift value and linkage illustration (5-5)



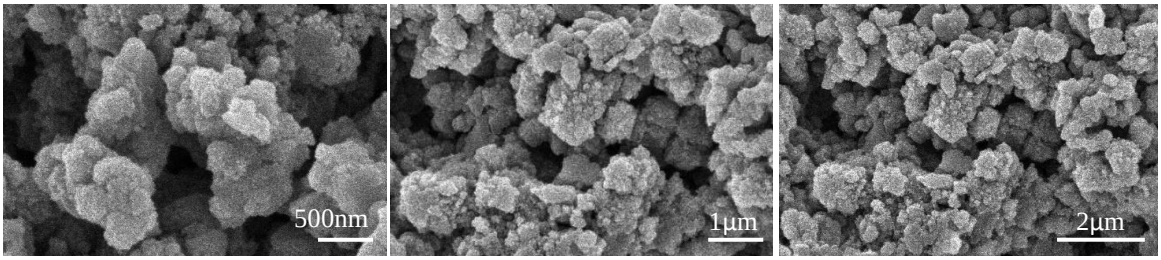
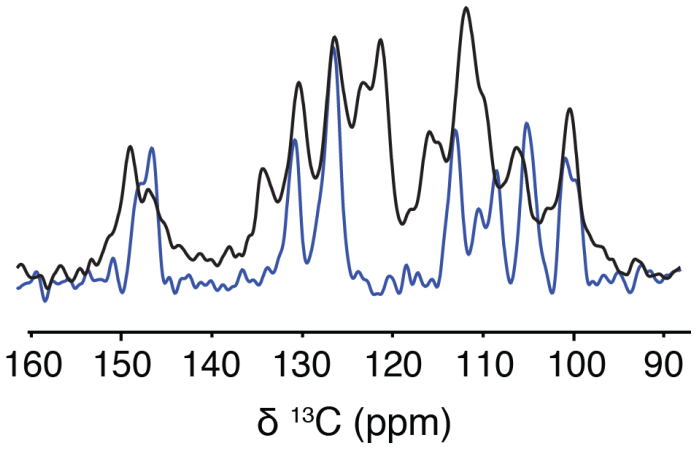
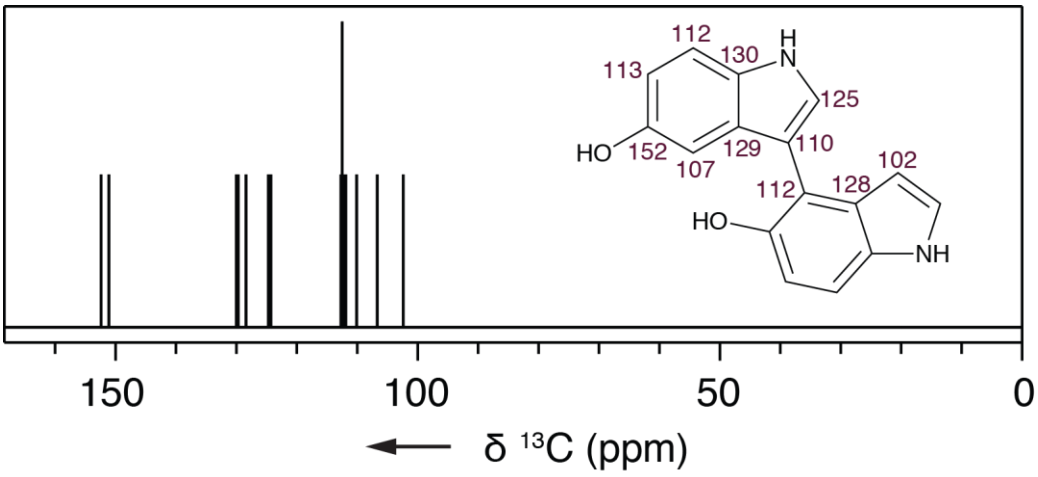


Figure S1.8: Spectral analysis of 4-HI-melanin. (top) 1D ^{13}C NMR of 4-HI-melanin (black) is overlaid with 4-HI monomer (blue). (middle) Using concepts of symmetric *versus* asymmetric linkage and chemical shift predictions via ChemDraw (Perkin-Elmer). All possible linkage combinations were drawn, and its chemical shifts were predicted. The one that matched the most with our experimental data was taken. We conclude a linkage of 5-5 between the 4-HI monomer. (bottom) SEM images of 4-HI-melanin at different magnifications.



NMR predicted chemical shift value and linkage illustration (2-4)



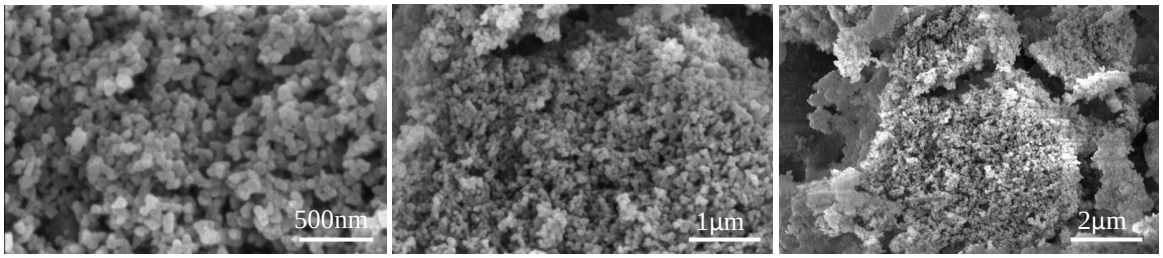
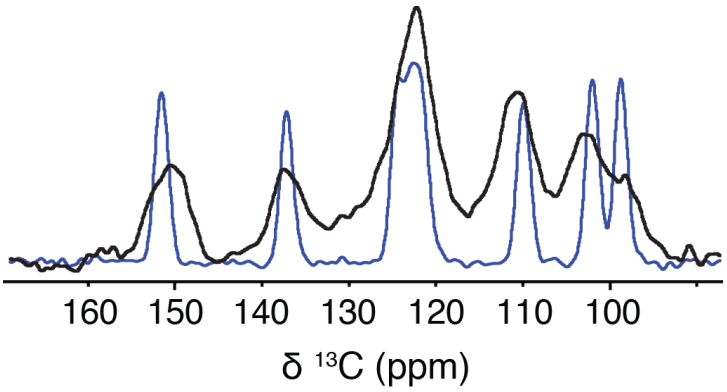


Figure S1.9: Spectral analysis of 5-HI-melanin. (top) 1D ^{13}C NMR of 5-HI-melanin (black) is overlaid with 5-HI monomer (blue). (middle) Using concepts of symmetric *versus* asymmetric linkage and chemical shift predictions via ChemDraw (Perkin-Elmer). All possible linkage combinations were drawn, and its chemical shifts were predicted. The one that matched the most with our experimental data was taken. We conclude a linkage of 2-4 between the 5-HI monomers. (bottom) SEM images of 5-HI-melanin at different magnifications.



NMR predicted chemical shift value and linkage illustration (4-4)

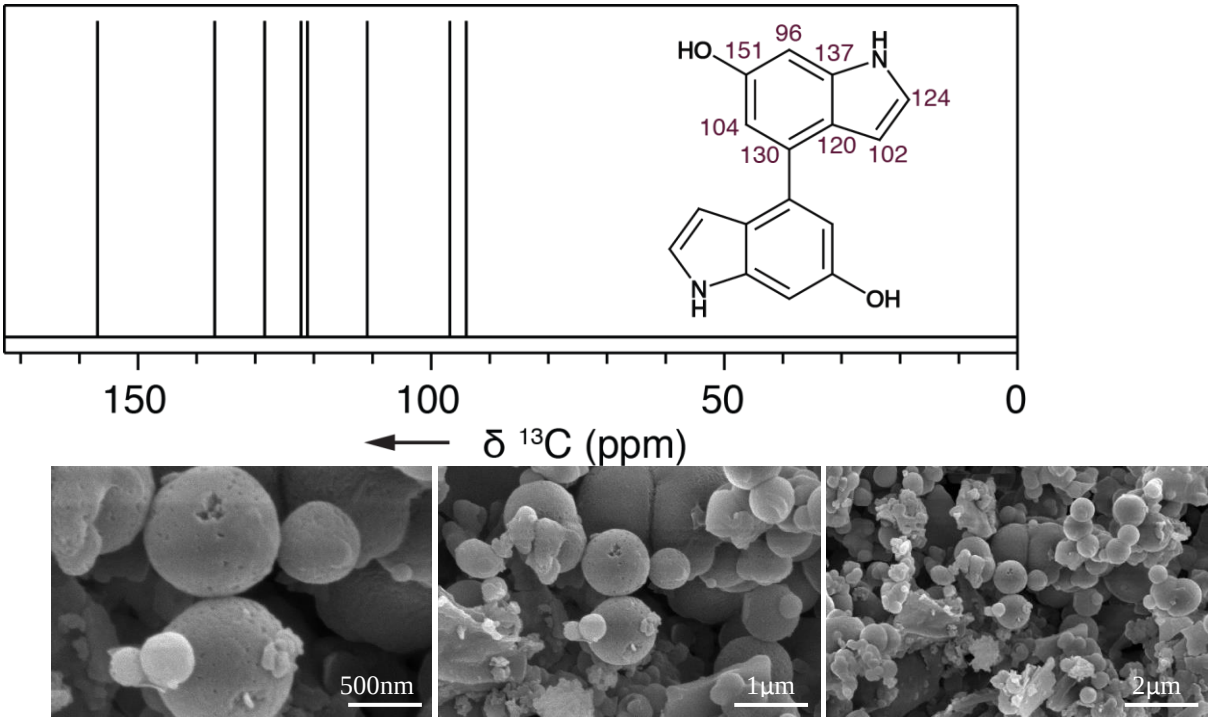


Figure S1.10: Spectral analysis of 6-HI-melanin. (top) 1D ^{13}C NMR of 6-HI-melanin (black) is overlaid with 6-HI monomer (blue). (middle) Using concepts of symmetric *versus* asymmetric linkage and chemical shift predictions via ChemDraw (Perkin-Elmer). All possible linkage combinations were drawn, and its chemical shifts were predicted. The one that matched the most with our experimental data was taken. We conclude a linkage of 4-4 between the 6-HI monomers. (bottom) SEM images of melanized 6-HI at different scales are shown.

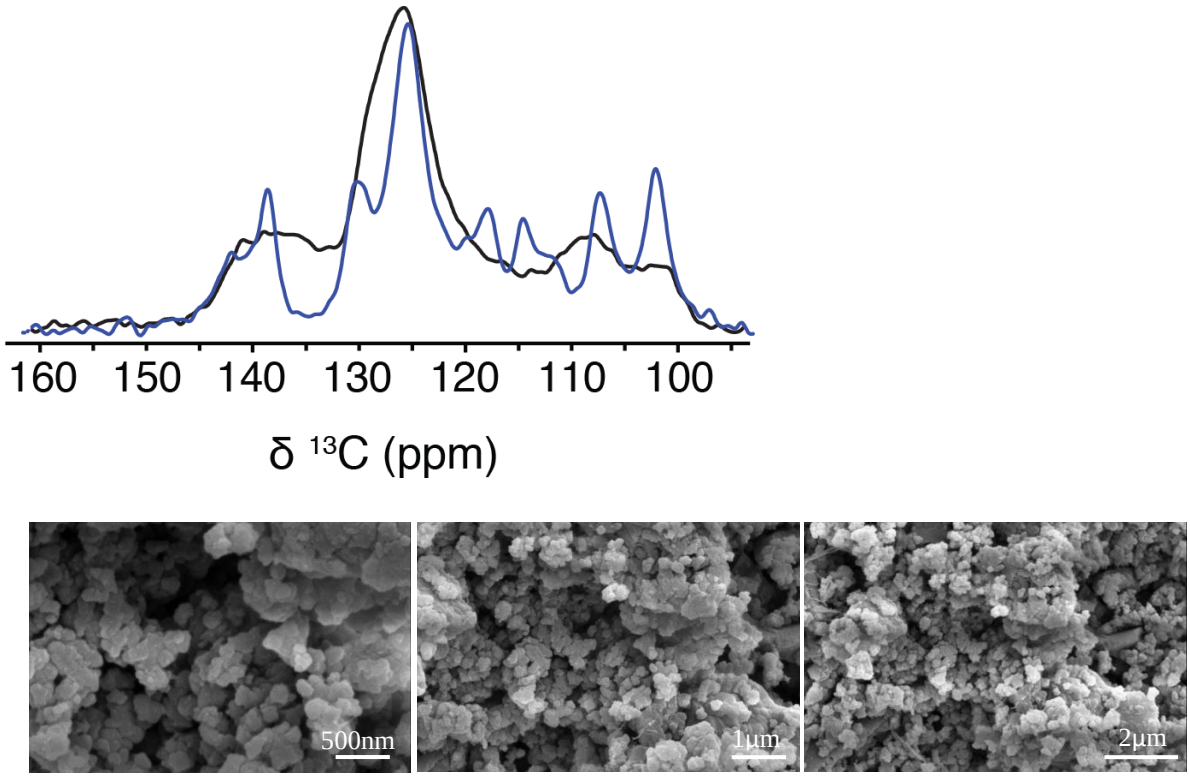


Figure S1.11: Spectral analysis of 7-HI-melanin. (top) 1D ^{13}C NMR of 7-HI-melanin (black) is overlaid with 7-HI precursor (blue). All possible linkage combinations were drawn, and its chemical shifts were predicted. Due to the broadness of the melanized 7-HI peaks, there are several possibilities of linkage between the 7-HI monomer groups. (bottom) SEM images of melanized 7-HI at different scales are shown.

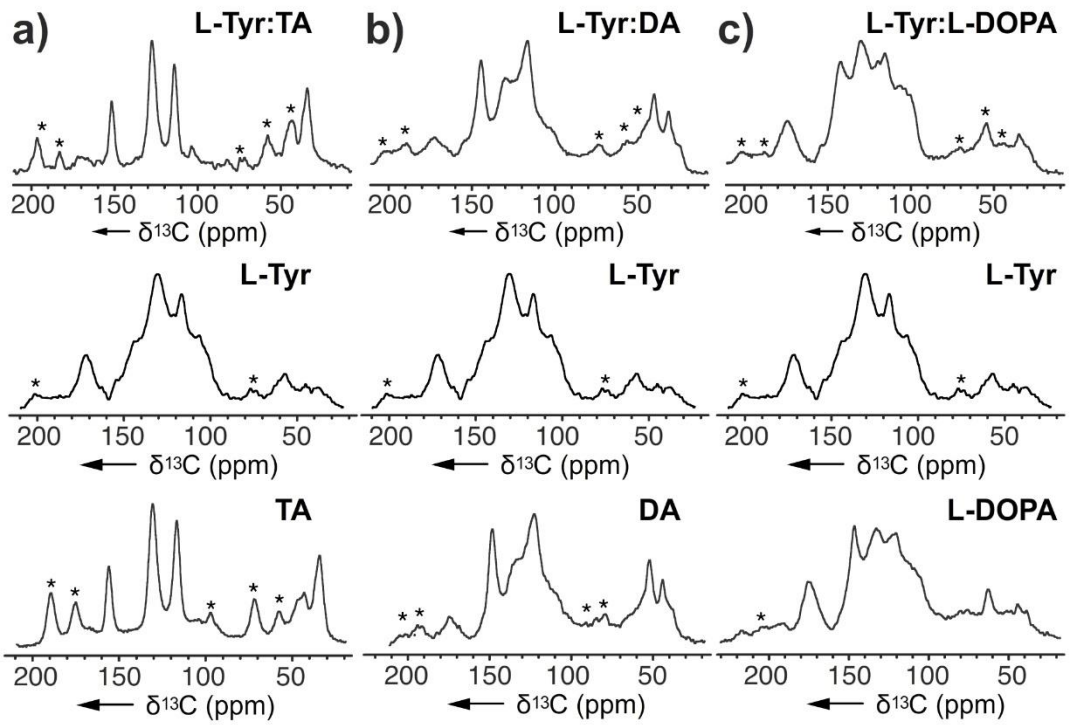


Figure S1.12: 1D ^{13}C NMR of melanin produced by copolymerizing two monomers: **a)** L-Tyr:TA melanin generated from a mixture of L-Tyr and TA, **b)** L-Tyr:L-DA melanin, **c)** L-Tyr:L-DOPA melanin using A. bisporus tyrosinase. Comparison of the NMR spectra of these mixtures to that of its pure monomers (shown below the copolymers) offers insight into the monomer selectivity of this enzyme. This figure is provided as supporting information for Fig. 1.4a-1.4c in the manuscript where panels a-c in Fig. S1.12 supports panel a-c in Fig. 1.4.

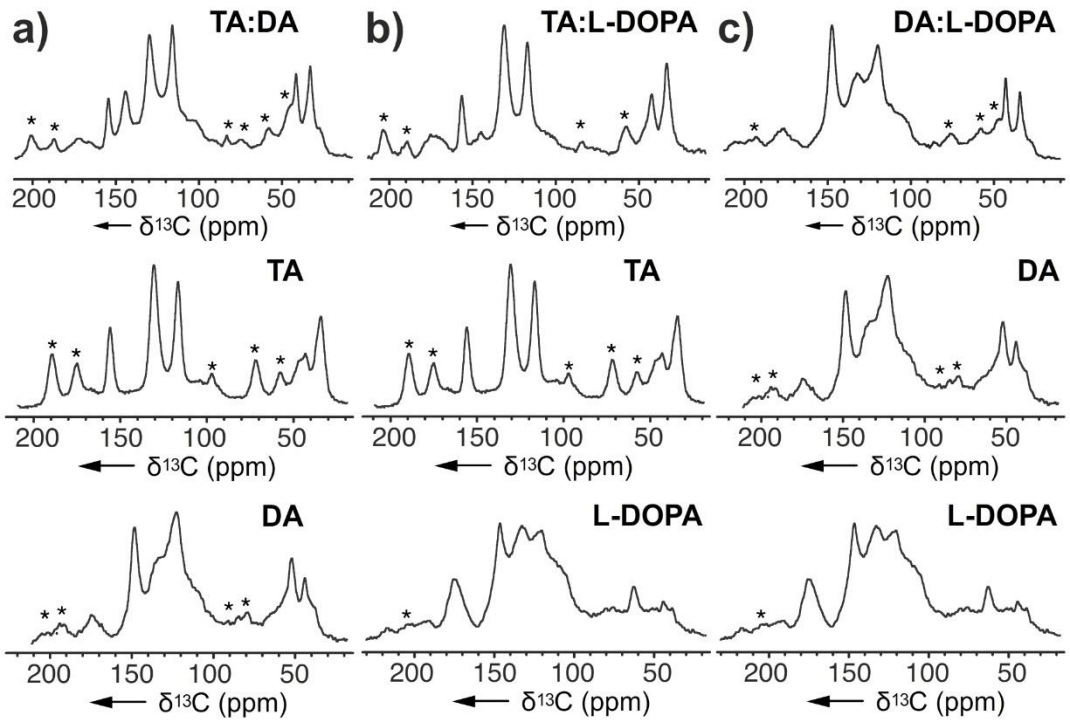


Figure S1.13: 1D ^{13}C NMR of melanin produced by copolymerizing two monomers: a) TA:DA melanin, b) TA:L-DOPA melanin, and c) DA:L-DOPA melanin using A. bisporus tyrosinase. Comparison of the NMR spectra of these mixtures with that of pure monomers (Fig. 2) was used to explore the monomer selectivity of this enzyme. This figure is provided as supporting information for Fig. 1.4d-1.4f in the manuscript where panels a-c in Fig. S1.13 supports panel d-f in Fig. 1.4.

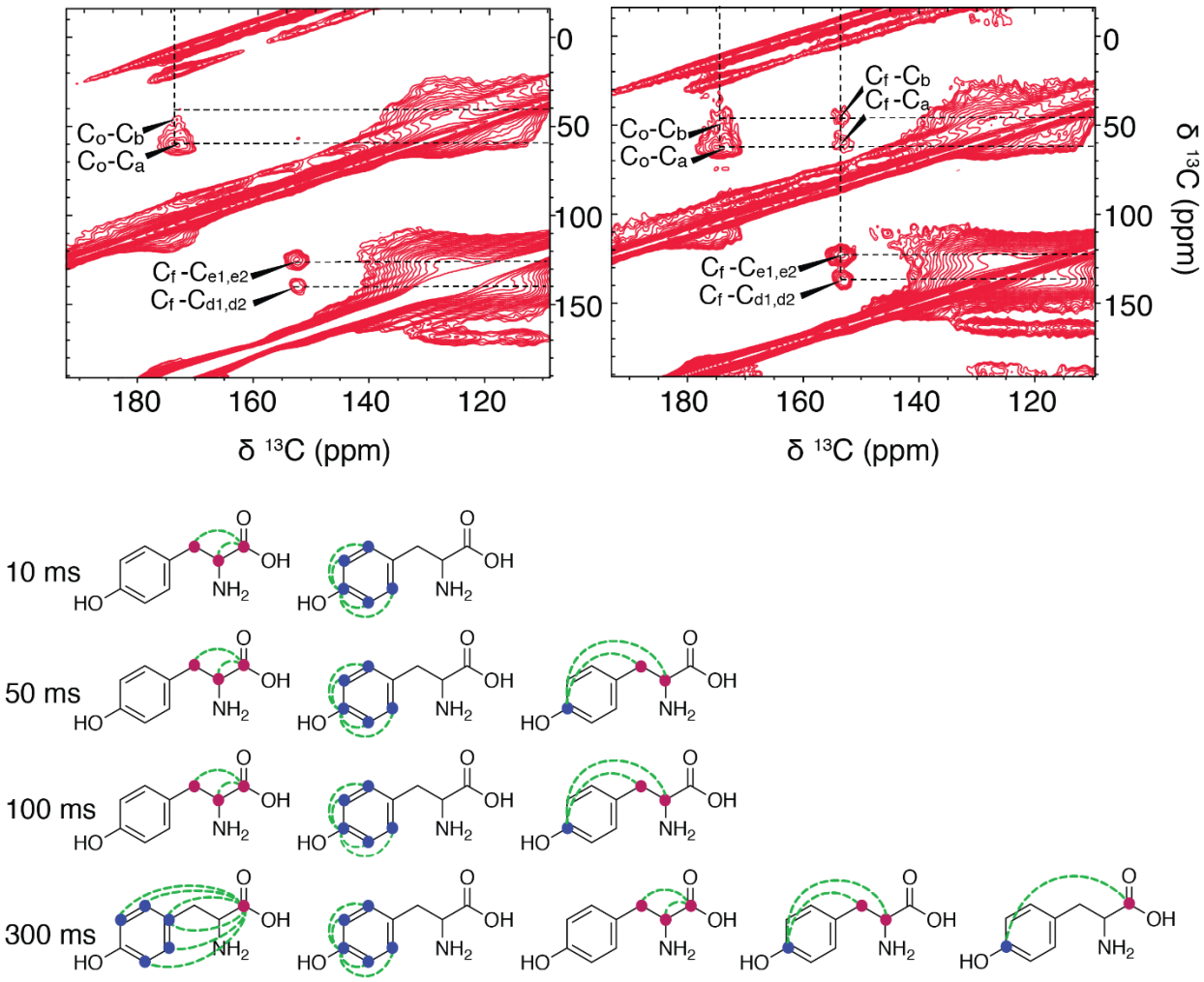


Figure S1.14: 2D NMR analyses of U-^{13}C,^{15}N-L-Tyr-melanin. (top) 2D ^{13}C-^{13}C PDSD spectra of U-^{13}C,^{15}N-L-Tyr-melanin showing the correlations between the atoms in the polymer at 10 ms (left) and 100 ms (right) mixing times. (bottom) Observed correlations are depicted on the tyrosine monomer as a function of mixing times. The order of appearance of the correlations is drawn according to the intensity of the contour peaks with respect to each other from strongest to weakest. The observed correlations are mostly due to intramolecular correlations instead of intermolecular correlations, with the exception of possible intermolecular interactions between stacked tyrosine units.

Note: The molecules that are of higher population are more likely to be detected by NMR. Not all molecules in a system are going to be detected by NMR especially given the broadness of the peaks, and PDSD doesn't guarantee all components of the sample will be detected, which is why we have combined our PDSD analysis with SPC5 data (Figure 1.5A). We tackled the C-C relationship through both through-space and through-bond experiments to obtain as much information as we can with the given technology. The DHI and DHICA units were detected in SPC5, which confirms the conclusion from Fig. S1.1.

K. Additional Discussion. The significance of many possibilities for structural diversity in eumelanin pigments explains why there are different variations of melanin in nature.

The sequential progression of intermediates in the Raper-Mason model is an oversimplification that has not been addressed in the literature thus far. As such, our findings open a new avenue for the exploration of the substantially nuanced Raper-Mason pathway that was previously based on simple sequential progression.

Other decarboxylase steps were mentioned in Scheme 1 to point out there are many other enzymes present *in vivo*, and which can each effect the output melanin.

In the past, very limited studies have been done on the selectivity of tyrosinase, for example, Nistora and coworkers[S1] use the tyrosinase as a biosensor in electrodes. In a second example, Florescu and coworkers[S2] focus on a tyrosinase-based biosensor developed to detect dopamine. None of the works focused on the enzyme's selectivity for the different metabolites in the Raper-Mason pathway, nor the selectivity for unnatural substrates that could be present as shown in Fig. 3 (4-HI, 5-HI, 6-HI, and 7-HI) and how all these alter the melanin's structure.

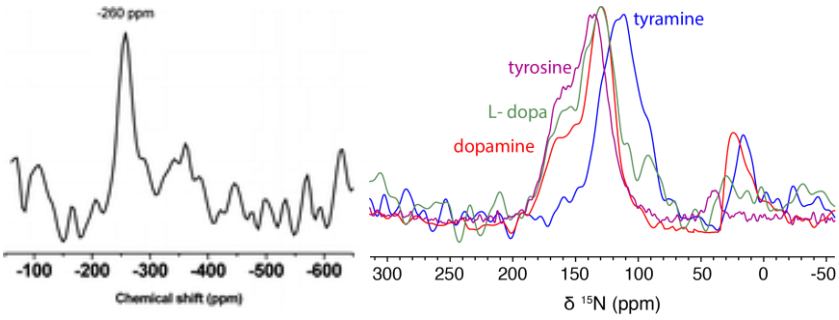


Figure S1.15: (left) Solid state ^{15}N NMR spectra of poly(dopamine) (400 MHz, spinning rate: 6 kHz).[S4] (right) The ^{15}N NMR spectrum of the four precursor melanins (Tyrosine NP, tyramine NP, L-dopa NP and dopamine NP) presented in this study acquired with DNP (600 MHz, spinning rate: 12.5 kHz).

It should be noted that the melanin polymers prepared in this study were prepared using different methods than those previously reported in the literature.[S3] Previous protocols involve washing the melanin polymer with harsh 6 M HCl after being polymerized, "The reaction mixture was then brought to pH 1 by addition of 6 M HCl and boiled for 20 min." This additional step destroys the overall pigment structure, while our approach of thoroughly washing the polymers with H_2O is noninvasive. This explains why our 2D NMR spectra and the conclusions drawn from this study are different from what has been published.

1.4 Acknowledgements

Chapter 1, in full, is a reprint of the material as it appears, Ni, Qing Zhe; Sierra, Brianna N., La Clair, James J., Burkart, Michael D. "Chemoenzymatic elaboration of the Raper-Mason

pathway unravels the structural diversity within eumelanin pigments." *Chem. Sci.* **2020**, *11*, 7836-7841. The dissertation author was the secondary investigator and author of this paper.

1.5 References

1. L. D'Alba, M. D. Shawkey, Melanosomes: Biogenesis, Properties, and Evolution of an Ancient Organelle. *Physiol. Rev.* **2019**, *99*, 1-19.

2. F. Solano, Melanin and Melanin-Related Polymers as Materials with Biomedical and Biotechnological Applications-Cuttlefish Ink and Mussel Foot Proteins as Inspired Biomolecules. *Int. J. Mol. Sci.* **2017**, *18*, E1561.

3. A. Büngeler, B. Hämisch, O. I. Strub, The Supramolecular Buildup of Eumelanin: Structures, Mechanisms, Controllability *Int. J. Mol. Sci.* **2017**, *18*, E1901.

4. R. Slominski, M. Zmijewski, A. Slominski. The role of melanin pigment in melanoma. *Exp. Dermatol.* **2015**, *4*, 258-259.

5. A. J. Thody, E. M. Higgins, K. Wakamatsu, S. Ito, S. A. Burchill, J. M. Marks, Pheomelanin as well as eumelanin is present in human epidermis. *J. Invest. Dermatol*. **1991**, *97*, 340-344.

6. R. L. Haining, C. Achat-Mendes, Neuromelanin, one of the most overlooked molecules in modern medicine, is not a spectator. *Neural Regen. Res*. **2017**, *12*, 372-375.

7. T. Iwasaki, Y. Tamai, M. Yamamoto, T. Taniguchi, K. Kishikawa, M. Kohri, Melanin Precursor Influence on Structural Colors from Artificial Melanin Particles: PolyDOPA, Polydopamine, and Polynorepinephrine. *Langmuir.* **2018**, *34*, 11814-11821.

8. B. Marchetti, T. N. Karsili, Theoretical insights into the photo-protective mechanisms of natural biological sunscreens: building blocks of eumelanin and pheomelanin. *Phys. Chem. Chem. Phys.* **2016**, *18*, 3644-3658.

9. M. J. Simpson, J. W. Wilson, M. A. Phipps, F. Robles, M. A. Selim, W. S. Warren, Nonlinear microscopy of eumelanin and pheomelanin with subcellular resolution. *J. Invest. Dermatol*. **2013**, *133*, 1822-1826.

10. H. C. Eisenman, A. Casadevall, A. Synthesis and assembly of fungal melanin. *Appl. Microbiol. Biotechnol*. **2012**, *93*, 931-940.

11. M. Sugumaran, Reactivities of Quinone Methides versus o-Quinones in Catecholamine Metabolism and Eumelanin Biosynthesis. *Int. J. Mol. Sci.* **2016**, *17*, E1576.

12. M. d'Ischia, A. Napolitano, A. Pezzella, P. Meredith, and T. Sarna, Chemical and Structural Diversity in Eumelanins: Unexplored Bio-Optoelectronic Materials. *Angew. Chem. Int. Ed*., **2009**, *48*, 3914-3921

13. H. S. Raper, The Aerobic Oxidases *Physiol. Rev*. **1928**, *8*, 245-282

14. H. S. Mason, The chemistry of melanin; mechanism of the oxidation of dihydroxyphenylalanine by tyrosinase. *J. Biol. Chem.* **1948**, *172*, 83-99.

15. G. M. Casanola-Martin, H. Le-Thi-Thu, Y. Marrero-Ponce, J. A. Castillo-Garit, F. Torrens, A. Rescigno, C. Abad, M. R. Khan, M. R. Tyrosinase enzyme: 1. An overview on a pharmacological target. *Curr. Top. Med. Chem*. **2014**, *14*, 1494-501.

16. M. Jiang, G. Xu, J. Ni, K. Zhang, J. Dong, R. Han, Y. Ni, Improving Soluble Expression of Tyrosine Decarboxylase from Lac-tobacillus brevis for Tyramine Synthesis with High Total Turnover Number. *Appl. Biochem. Biotechnol*. **2019,** *188*, 436-449.

17. S. Chatterjee, R, Prados-Rosales, S. Tan, V. C. Phan, C. Chrissian, B. Itin, H. Wang, A. Khajo, R. S. Magliozzo, A. Casadevall, R. E. Stark, The melanization road more traveled by: Precursor substrate effects on melanin synthesis in cell-free and fungal cell systems. *J. Biol. Chem.* **2018**, *293*, 20157-20168.

18. J. D. Nosanchuk, R. E. Stark, A. Casadevall, A. Fungal Melanin: What do We Know About Structure? *Front. Microbiol.* **2015**, *6*, 1463.

19. Meiler, J.; Maier, W.; Will, M.; Meusinger, R. Using neural networks for ^{13}C NMR chemical shift prediction-comparison with traditional methods. *J. Magn. Reson.* **2002**, *157*, 242-252.

20. S. Halaouli, M. Asther, J. C. Sigoillot, M. Hamdi, A. Lomascolo, Fungal tyrosinases: new prospects in molecular characteristics, bioengineering and biotechnological applications. *J. Appl. Microbiol.* **2006**, *100*, 219-232.

21. Q. Z. Ni, T. Daviso, T. V. Can, E. Markhasin, S. K. Jawla, T. M. Swager, R. J. Temkin, J. Herzfeld, R. G. Griffin, High frequency dynamic nuclear polarization. *Acc. Chem. Res*. **2013**, *46*, 1933-1941.

22. S. Lilly Thankamony, J. J. Wittmann, M. Kaushik, B. Corzilius, Dynamic nuclear polarization for sensitivity enhancement in modern solid-state NMR. *Prog. Nucl. Magn. Reson. Spectrosc*. **2017**, *102-103*, 120-195.

23. M. d'Ischia, A. Napolitano, V. Ball, CT. Chen, M. Buehler, Markus. Polydopamine and Eumelanin: From Structure–Property Relationships to a Unified Tailoring Strategy. *Acc. Chem. Res.* **2014**, *47*, 3541−3550

24. M. Gatoo, S. Naseem, M. Arfat, A. Dar, K. Qasim, S. Zubair. Physicochemical Properties of Nanomaterials: Implication in Associated Toxic Manifestations. *Biomed Res. Int.* **2014**, *498420,* 1-8

25. K. Takegoshi, S. Nakamura, T. Terao, ^{13}C-^{1}H dipolar-driven ^{13}C-^{13}C recoupling without ^{13}C rf irradiation in nuclear magnetic resonance of rotating solids. *J. Chem. Phys.* **2003**, *118*, 2325-2341.

26. Y. Li, J. Liu, Y. Wang, H. W. Chan, L. Wang, W. Chan, Mass Spectrometric and Spectrophotometric Analyses Reveal an Alternative Structure and a New Formation Mechanism for Melanin. *Anal. Chem*. **2015**, *87*, 7958-7963.

27. Y. K. Lee, N. D. Kurur, M. Helmle, O. G. Johannessen, N. C. Neilsen, M. H. Levitt, Efficient dipolar recoupling in the NMR of rotating solids. A sevenfold symmetric radiofrequency pulse sequence. *Chem. Phys. Lett*. **1995**, *242*, 304-309.

28. A. W. Hing, S. Vega, J. Schaefer, Transferred-echo double-resonance NMR. *J. Mag. Reson*. **1992**, *96,* 205-209.

29. L. P. Mekala, M. Mohammed, S. Chintalapati, V. R. Chintalapati, Stable Isotope-Assisted Metabolic Profiling Reveals Growth Mode Dependent Differential Metabolism and Multiple Catabolic Pathways of l-Phenylalanine in Rubrivivax benzoatilyticus JA2. *J. Proteome Res*. **2018**, *17,* 189-202.

S1. C. Nistora. J. Emnéus. L. Gorton, A. Ciucub, Improved stability and altered selectivity of tyrosinase based graphite electrodes for detection of phenolic compounds. *Anal. Chim. Acta.* **1999**, *387*, 309-326.

S2. M. Florescu, D. David, Tyrosinase-Based Biosensors for Selective Dopamine Detection. *Sensors (Basel),* **2017**, *17,* 1314

S3. S. Chatterjee, R. Prados-Rosales, S. Tan, V. C. Phan, C. Chrissian, B. Itin, H. Wang, A. Khajo, R. S. Magliozzo, A. Casadevall, R. E. Stark, The melanization road more traveled by: Precursor substrate effects on melanin synthesis in cell-free and fungal cell systems. *J. Biol. Chem.* **2018**, *293,* 20157-20168.

S4. D. R. Dreyer, D. Miller, B. Freeman, D. R. Paul, C. W. Bielawski, Elucidating the structure of poly(dopamine). *Langmuir,* **2002**, *28,* 6428-6435.

Chapter 2 Chemoenzymatic Isolation and Characterization of High Purity Mammalian Melanin

2.1 Introduction

Melanin is the generic name for a ubiquitous family of polymers found across a variety of organisms in nature.[1] Although mostly investigated in humans in relation to melanoma skin cancer, its remarkable stability has enabled the detection of melanin in dinosaurs, early birds, non-avian theropod and cephalopod fossils.[2] While recognized for its role in decoding evolution,[3] a detailed understanding of the role and function of its complex organization within tissues continues to challenge our perception of molecular and structural biology.[4] At the molecular level, melanin is a class of biological polymers derived from complex linkages between a small set of monomeric building blocks with eumelanin arising predominantly from 5,6-dihydroxyindole (DHI) and 5,6-dihydroxyindole-2-carboxylic acid (DHICA),[5] pheomelanin from benzothiazine and benzothiazole (both derived from 5-*S*-cysteinyl-DOPA),[6] neuromelanin from DHI,[7] allomelanin from 1,8-dihydroxynaphthalene,[8] and pyomelanin from homogentisate.[9]

The determination of the molecular structure of each of these five types of melanin polymers has faced significant obstacles.[10] First, melanin particles are insoluble solids with purity that is often difficult to evaluate by traditional methods such as solution-phase NMR. Second, common atomic-resolution approaches, such as single crystal or powder X-ray diffraction, have failed due to melanin's non-crystalline morphology. Alternatively, solid-state NMR (ss-NMR) studies have provided an initial understanding of the structural features.[11] However, these can be complicated to interpret due to contaminants and metabolites that obscure conclusive structural assignments. While access to synthetically and enzymatically-produced melanin can offer a viable approach to obtain melanin samples at high-purity, they fail to meet the structural complexity observed within nature.[12] In an effort to understand the structure-function relationships of natural

melanin, we began with the goal of isolating highly pure melanin samples from diverse mammalian tissues.[13] Here we report the development of a chemoenzymatic process which unites enzymes for tissue catabolism with solvent density gradients and tracks the process with fluorescent probes specific for biomolecular labeling. This methodology delivers a reproducible model for obtaining highly purified, isolated melanin that is further characterized using SEM and ss-NMR. These tools enable quantitative analysis of melanin content within tissues, which can be extended to other organisms.

Melanin particles have been conventionally isolated using combinations of strong acid (6 M HCl) or base (6 M NaOH) at elevated temperatures (60-100°C for 12-48 h) or by microwave assistance.[14] While effective at yielding black pigments, these methods are known to dissolve eumelanin, altering its structure by encouraging reactivity within the DHI and DHICA monomers.[15] Alternatively, enzymatic digestion has been tentatively explored as a milder alternative.[16] While enzymatic treatments do not perturb the architecture of melanin nanoparticles, they are reported to be inefficient at removing proteins and lipids and suffer from low product yield.[17] Small molecule contaminants are typically removed with organic solvent extractions (petroleum ether, ethyl acetate, acetone or ethanol). These can be combined with redox modulation using dithiothreitol (DTT)[18] and proteolysis (proteinase)[19] as a means to reduce protein contaminants even further. Together, this purification method can provide quantitative analysis of melanin content within tissues, which can be extended to other organisms. Thus, we sought to develop an improved approach using novel sets of treatments.

We began by developing a method to first digest the primary structural components of a tissue, followed by steps to digest any remaining non-melanin biopolymers. We designed a four-step protocol (Fig. 2.1) beginning with surgical isolation of a tissue sample (step 1, Fig. 2.1). Once

removed and washed with physiological buffer, the process began by digestion of the connective tissues within the sample using an enzyme such as collagenase (step 2, Fig. 2.1). Next, a cocktail of enzymes (step 3, Fig. 2.1) was applied to degrade the remaining biopolymers (DNA with deoxyribonuclease (DNase) and lipids with lipase) into readily extracted aqueous soluble fragments. The final step removes all small molecules and undigested cellular contaminants by using the high density of melanin as a tool for selective gravimetric purification from an organic solvent mixture (step 4, Fig. 2.1). This four-step process was able to yield pure melanin that could be validated by SEM imaging and ultimately using ss-NMR analysis on the final extract.

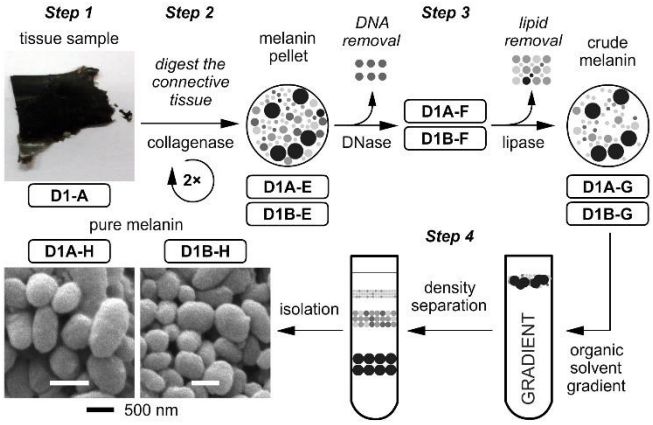


Figure 2.1: A four-step procedure was developed to isolate melanin from a mammalian specimen. (Step 1) This process began by dissecting the melanin-containing tissue sample D1-A from a specimen. Care was taken to surgically remove tissues that did not contain the pigment. (Step 2) This was then followed by digestion of the connective tissue to deliver a melanin pellet after gentle centrifugation. The final two steps served to purify the melanin. (Step 3) This began with degradation of DNA by a DNase and lipids with a lipase. (Step 4) Finally, an organic solvent density gradient was used to remove small molecule metabolites and select for the dense melanin particles. Fractions within this study are given by D1A-B to D1A-H or D1B-B to D1B-H, where D1A or D1B indicates the respective insoluble (D1A) or soluble (D1B) fractions after the first collagenase treatment. The last letter denotes the step in the process, as given by: A, after dissection; B, after first collagenase treatment; C, after sonication; D, pellet after second collagenase treatment; E, after washing the pellet with deionized H_2O; F, after DNA removal with DNase; G, after lipid removal with lipase; and, H, after organic solvent density gradient separation.

Accessing local tissue banks for marine mammals (NOAA West Coast Regional Stranding Network permit file 151401WCR2019PR00028 to D.D.D.), we acquired a sample of tissue from the short beaked dolphin, *Delphinus delphis*, with which to evaluate this approach. This specimen was selected because melanin from dolphin has not been previously studied (providing an excellent test of a methods applicability), and the thick, leathery tissue presented a significant experimental challenge to test our method. As shown in Fig. 2.2a-2.2b, we surgically removed 1 cm × 2 cm sections from pigmented skin tissue proximal to the dorsal fin. Careful surgical procedures were used to isolate the dark stratum corneum (outermost layer, Fig. 2.2c). Surgical preparation of the melanin-containing tissue (step 1, Fig. 2.1) profoundly enhanced the ease in conducting the process and purity of the final melanin isolated.

Following dissection, digestion of the connective tissue (step 2, Fig. 2.1) was performed with a 2.28 g sample **D1-A** incubated at 37 °C for 24 h with 115 mg of collagenase (0.05 g/g of tissue) in 20 mL of deionized H_2O. After incubation, the solution was decanted to provide insoluble, undigested **D1A-B** and soluble, digested **D1B-B** (Fig. 2.2d) fractions. At this point, the insoluble samples were homogenized by sonication. This process repeated until all materials were digested. When complete, the supernatant was decanted from all fractions, and the melanin was thoroughly washed with deionized H_2O to remove any residual enzyme. Following centrifugation and supernatant removal, both samples were subjected to a second collagenase treatment, which delivered samples **D1A-D** and **D1B-D** (Fig. 2.2e).

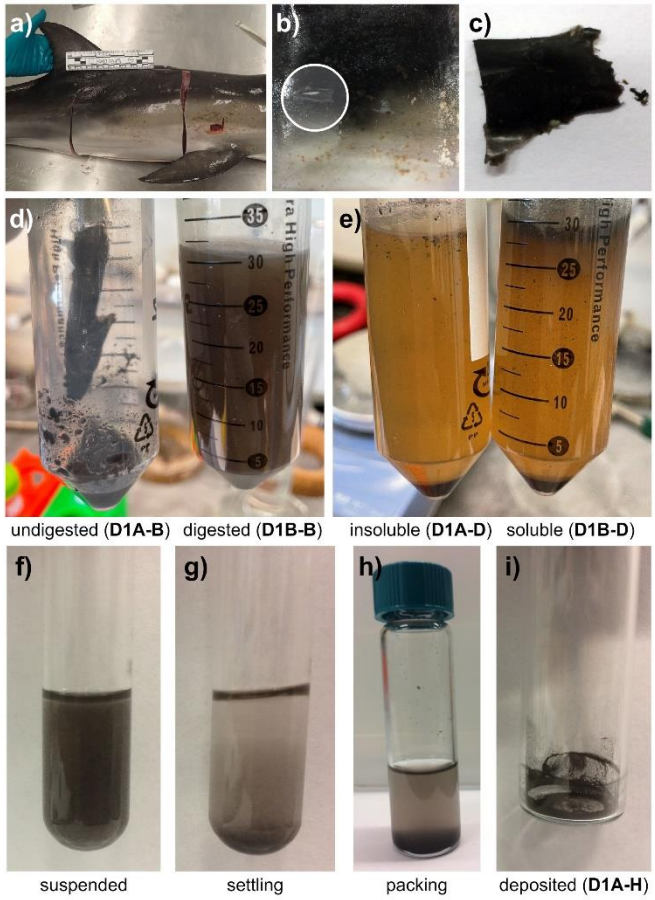


Figure 2.2: Images documenting the key steps of the chemoenzymatic process applied to the skin of dolphin *Delphinus delphis* to extract melanin. **a)** Freshly stranded specimen undergoing necropsy analysis by NOAA employee. **b)** Skin tissue that was surgically removed from this specimen (white circle). **c)** An exemplary sample was prepared for this study after the removal of all non-melanin containing tissues. **d)** Insoluble (**D1A-B**) and soluble (**D1B-B**) fractions from the first digestion with collagenase. **e)** Samples **D1A-D** and **D1B-D** after sonification, additional collagenase treatment and centrifugation at 4200 rpm. Black melanin pigment is seen settled at the bottom. **f)-i)** Steps of the density gradient process. **f)** Homogeneously suspend melanin power after enzymatic purification in a mixture of CH_2Cl_2 and MeOH such that it remains suspended when shaken. **g)** After sitting at 23 °C for 15 min, the melanin particles settle. **h)** This procedure is repeated 3 times allowing final transfer and **i)** depositing in a storage vial. During the settling stage, the heavy melanin particles sink while other biological particles (white to grey color) float, which is readily removed by decanting or removal by syringe. This experiment was repeated three times, with analytical data collected from the third repetition. Samples from the first two repetitions were used for ss-NMR analyses and select samples from the third repetition are shown.

Finding that repeated treatment with collagenase was required for optimal extraction yield, we realized a diagnostic tool could be used to monitor the purification process. After exploring a

panel of stains, we identified three fluorescent dyes able to detect lipids (Nile red), oligonucleotides (4,6-diamidino-2-phenylindole, DAPI) and proteins (Alexa fluor 488 NHS ester, an amine reactive stain) in the sample. Each stain was developed in parallel by placing a 20 µL drop of each sample on a glass slide and adding 5 µL of aqueous dye solutions: 1 mM Nile Red, 1 mM DAPI, or 1 mM Alexa fluor 488 NHS. After incubating for 20 min at 23 °C, a coverslip was placed on the slide and the slide imaged on an inverted fluorescent microscope (EVOS, Invitrogen). As shown in Fig. 2.3a, considerable lipid (red in N) and an oligonucleotide, DNA/RNA (blue in D) contamination was observed in the material collected after collagenase treatment **D1A-E** and **D1B-E** (Fig. 2.3a). While not unexpected due to the proteolytic activity of collagenase, we did not observe significant levels of protein contamination.

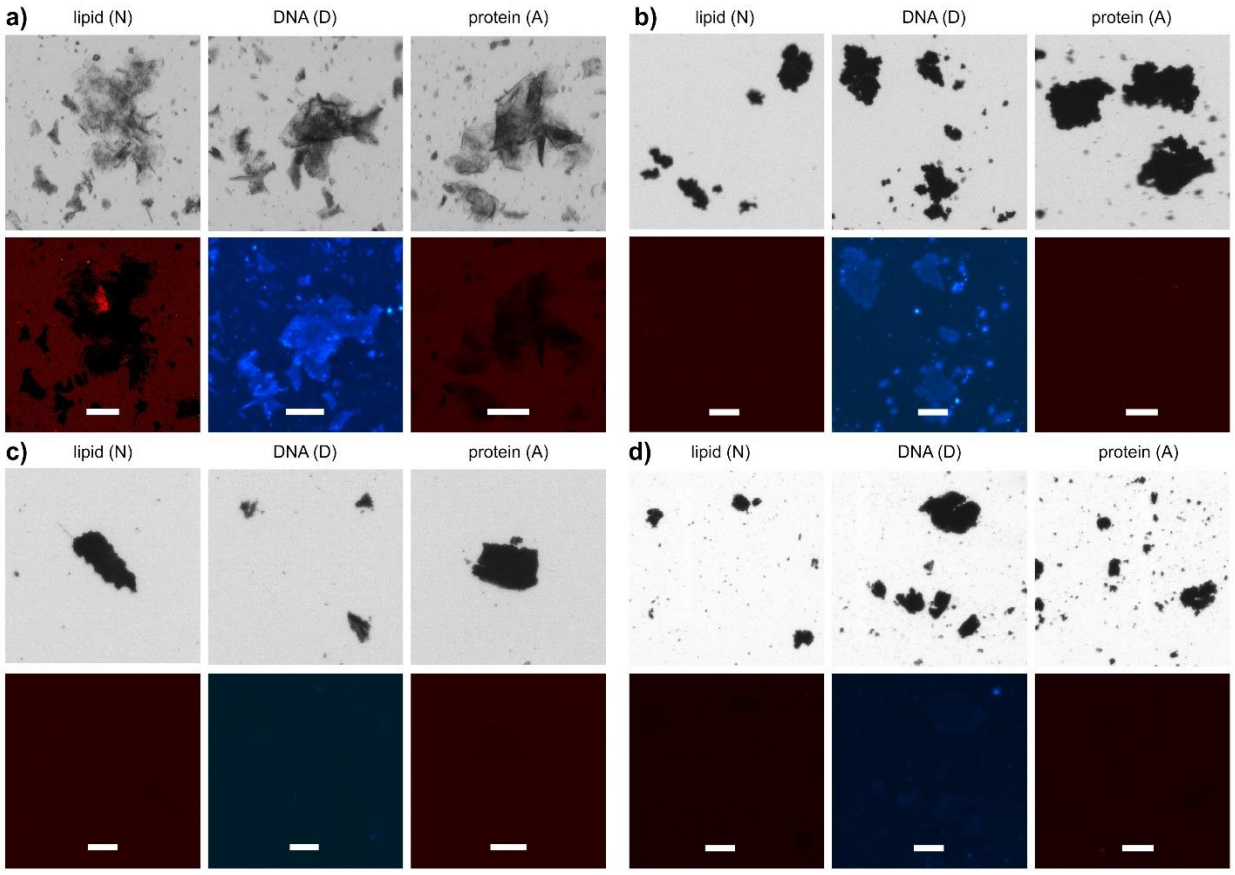


Figure 2.3: Screening for melanin purity using combination of bright field and fluorescence microscopy. A three-stain protocol was developed to label samples during the enzymatic degradation process for biopolymeric impurities. This consisted of using Nile red to stain lipids (red fluorescent, N), DAPI (4′,6-diamidino-2-phenylindole) to stain DNA (blue, D) and Alexa fluor 488 NHS, an amine-reactive stain for protein (green, A). a) Images of stained samples of the pellet after collagenase treatment and washing (D1B-E). b) Images of stained samples of the pellet after lipase treatment (D1A-G). c) Images of stained samples of the pellet after DNase treatment (**D1A-F** is shown). **d)** Images of stained samples of the pellet after the organic solvent density purification (**D1B-H**). After each step, both **D1A** and **D1B** samples were comparable, only one example was shown as given by sample number. For each stain, a bright field (transmission) and fluorescent image are provided. Fluorescence was collected by imaging with excitation (λ_{ex}) and emission (λ_{em}) at 531±40 and 593±40 nm (Nile red), 357±44 and 446±60 nm (DAPI), or 470±22 and 525±50 nm (Alexa fluor 488 NHS). Scale bars denote 100 µm for all panels. Dark images denote the lack of fluorescence (all taken under the same acquisition settings).

With this knowledge at hand, we moved on to digest lipid and oligonucleotide contaminants. The black pellets from **D1A-E** and **D1B-E** were resuspended in 25 mL of 2 µM bovine pancreas DNase and sonicated for 5 min on ice. After precipitation by gravity at 23 °C, the

pellet was collected and washed 3× with deionized H_2O to yield 23.6 mg (dry weight) of DNA free melanin **D1A-F** and **D1B-F** (Fig. 2.3b). Using a comparable procedure, the samples were then suspended in 25 mL of 2 µM Amano lipase from *Pseudomonas fluorescens* and sonicated. After precipitation by gravity at 23 °C, the pellet was collected and washed 3× with deionized H_2O to yield 17.4 mg (dry weight) of lipid free melanin **D1A-G** and **D1B-G** (Fig. 2.3b).

While enzymatic treatment was sufficient to remove the majority of biomolecular contamination within the pigments, samples after **D1A-G** and **D1B-G** (Fig. 2.1) remained contaminated by small residual molecules. After exploring different options for purification, we found that melanin density (1.5-1.8 g/cm^3)[20] could serve as a means of purification. By mixing CH_2Cl_2 (1.33 g/cm^3) and MeOH (0.79 g/cm^3) in varying ratios, we identified conditions allowing only the melanin particles to settle *via* gravity. For the example shown in Fig. 2.2f-2.2i, we suspended the crude melanin pellet **D1A-G** or **D1B-G** in 1 mL of MeOH and added 2.5 mL of CH_2Cl_2 and mixed by shaking. Once suspended (Fig. 2.2f), the melanin particles slowly settled over 20-30 min. Removal of the solvent, transfer to a storage vial and replication of the process (Fig. 2.2h) allowed the reproducible isolation of packed melanin samples (note: when pure melanin is dried (Fig. 2.2i), it is highly electrostatic and difficult to transfer to storage vials).

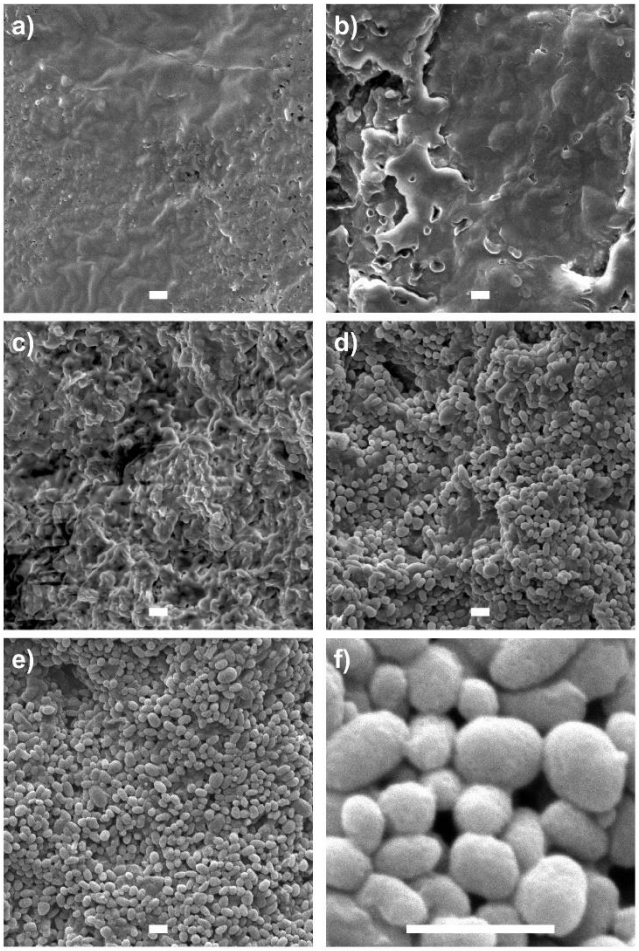


Figure 2.4: SEM imaging to assess structural purity, with images of: **a)** the skin after dissection as shown macroscopically in Fig. 2c; **b)** the tissue observed after the first collagenase treatment (**D1A-B**); **c)** the pellet after DNase treatment (**D1A-F**); **d)** the pellet after lipase treatment (**D1A-G**); **e)** purified melanin after the organic solvent density gradient purification (**D1A-H**), and; **f)** melanin from sample **D1A-H** at higher magnification. After each step, both **D1A** and **D1B** samples were comparable, only one example was shown as given by sample number. All scale bars denote 1 µm.

Next, we applied scanning electron microscopy (SEM) to validate the purification and strategy and characterize the materials (Fig. 2.4). Here, SEM analysis allowed for the evaluation of the structural purity of the melanin in parallel with its overall chemical purity. Not unexpectedly, samples collected at early steps in the purification process returned images of tissue surfaces (Fig. 2.4b), and with additional enzymatic treatment and density purification (Fig. 2.4c-2.4d) returned uniform spheroids 0.42 ± 0.06 nm diameter and 0.60 ± 0.10 nm in length (Fig. 2.4e-2.4f).

Interestingly, while collagenase treatment delivered a black pigment, pure melanin was not obtained until digestion of both lipid and oligonucleotides within these samples.

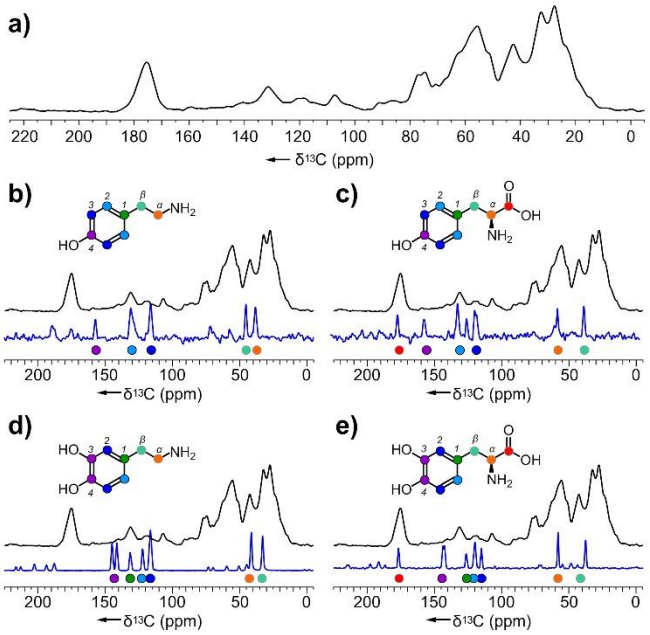


Figure 2.5: Molecular purity of melanin assessed by ss-NMR. a) A 1D ^{13}C CP NMR spectrum of a 30 mg sample of purified melanin (D1A-H and D1B-H from two repetitions of the study). Comparative analyses against ^{13}C NMR spectra of solid monomers b) tyramine, c) L-tyrosine, d) dopamine and e) L-DOPA (all shown in blue) allows one to suggest the monomeric composition of the melanin.[11a] We further validated the sample purity by comparison with samples of melanin from jumbo squid and cuttlefish ink (Supporting Fig. S2.2).[21] Supporting Fig. S2.3 provides a comparison of the NMR spectrum collected from D1A-H and D1B-H to that obtained from jumbo squid and cuttlefish ink. Supporting Fig. S2.4 provides a comparison of the NMR spectral data to melanin prepared by the action of the action of the *Agaricus bisporus* (mushroom) tyrosinase on L-tyrosine or L-DOPA.

Finally, we turned to ss-NMR to evaluate the isolated melanin at the molecular level. Combining samples **D1A-H** and **D1B-H** from two repetitions of this experiment, as noted in Fig. 1, provided a sufficient sample of pure melanin (30 mg) for ss-NMR analyses. As shown in Fig. 2.5a, we collected a 1D ^{13}C CP NMR spectrum from this batched sample. Using an approach recently implemented in our laboratory,[11a] we also collected ss-NMR spectra from the four monomers that can be observed during the Raper-Mason pathway, including tyramine (Fig. 2.5b), L-tyrosine (Fig. 2.5c), dopamine (Fig. 2.5d), and L-DOPA (Fig. 2.5e) (see further discussion on

these substrates and their role in the Raper-Mason pathway in our previous work[11a]). Overlays of these spectra with that obtained from the purified dolphin melanin were used to guide the assignments. While the substrates shown in Fig. 2.5, represent possible substrates or precursors for the dolphin tyrosinase, detailed screening efforts on cloned and expressed dolphin tyrosinase are required before the true substrate scope can be identified.

As shown in Fig. 2.5b-2.5e, the carbonyl peak at 174 ppm of the dolphin melanin indicates the dominance of tyrosine and L-DOPA. The peak at 155 ppm (weak peak observed) representing phenol groups is in agreement with those from tyrosine and tyramine. The aromatic region between 100 and 150 ppm shows distinct chemical shifts in agreement with those arising from tyrosine, tyramine and L-DOPA at 131 and 120 ppm. The aliphatic region between 0 to 100 ppm of the dolphin spectrum agrees more with tyrosine and L-DOPA. Together, we can gather that the dolphin melanin consists mostly of tyrosine, L-DOPA, and minor amounts of tyramine, with no evidence of dopamine intermediates. We hypothesize the precursor makeup results in a specific melanin species that contributes to melanosome morphology and plays a role in pigment organization.

In addition to its identification of the core unit of eumelanin, ss-NMR spectra collected on dolphin melanin (Fig. 2.5a) contained significant peaks at 20-30 ppm, typical of lipids or proteins. While not commonly reported, natural melanin often covalently-attaches peptides, proteins or lipids that are not removed during the isolation process.[22] While use of harsh acid and basic treatments may lead to ester or amide hydrolyses (in part removing covalently attached lipids or proteins/peptides) these resulting melanins no longer reflect their natural state. While a matter of further study, a detailed understanding of the post-biosynthetic modifications (covalent

modification by lipids, peptides or proteins) is an important next step towards understanding of the natural structure of melanin pigments.

2.2 Conclusion

Overall, this study demonstrates an important synergy between chemoenzymatic sample preparation guided by microscopic monitoring for molecular, structural, and atomic purity. Here we defined a method leveraging mild enzymatic conditions to free melanin particles entrained in tissue, which also enables the systematic removal of contaminating biomolecules. We also illustrated how organic solvent density gradients can be used to further enrich dense melanin particles. This method offers an alternative to conventional methods, many of which provide excellent quantitative utility.[23] Applying this methodology to short beaked dolphin tissue, we achieved 16.2 mg of total melanin from 2.28 g of wet tissue, indicating this sample of *Delphinus delphis* skin contained about 0.7% melanin. Using visual (Fig. 2.2) and microscopic (Figs. 2.3-2.4) methods to analyze the discarded liquid fractions at each stage, we were able to ensure that <1% of the melanin was lost through this process. The isolated melanin was evaluated by electron microscopy (Fig. 2.4) and ss-NMR (Fig. 2.5), and we deduced that the melanin polymers are dominated by tyrosine and L-DOPA, resulting in spheroids with highly regular shape. Studies are now underway to apply this approach to survey melanin from diverse bacterial,[24] fungal,[25] and mammalian[26] specimens.

2.3 Supporting Information

A. General methods. Chemical reagents and enzymes were purchased from Worthington, Sigma-Aldrich, or Fisher Scientific and used ‘as is’ unless noted. Milli-Q deionized water was used at all stages of the procedure.

B. Tissue sampling. Samples were cut from the skin of the tissue sample using a fillet knife (Berkeley Fishing). The samples were cut into 1 cm × 2 cm sections and the non-melanin containing tissues (white, colorless) were removed using a #3 scalpel (Fisher Scientific).

C. Collagenase digestion. Skin sample (2.28 g) D1-A was placed in a 50 mL Falcon tube and H2O (10 mL) was added until the sample was submerged. Collagenase Type I (Worthington Biochemical) (0.05 g/g sample) was added as a powder and the tube was shaken until the enzyme dissolved. The sample was then digested at 37 °C for 24 h to provide a sample that was soft enough to tear. Note: while this sample was easily broken and soft at this stage, not all tissues may be sufficiently digested with this procedure. We recommend that the sample be checked after the first treatment. If the sample was not soft enough to easily break apart the aqueous media was removed and the process was repeated. At this stage, the sample was separated into two fractions insoluble, undigested (D1A-B) and soluble, digested (D1B-B), and sonicated for 20 min (3 s on, 4 s off). After sonication, the solution was left to precipitate by gravity. Layers of solution lacking black hue were removed and another 2.5 mg of collagenase was added for every 1 mL of black solution. The sample was shaken and incubated at 37 °C for another 24 h. At this point, the sample was highly uniform and black. The sample was then washed three times by suspending in an equal volume of water and collecting by centrifugation (4200 rpm). Note: this may not be the case for all samples. If the sample was not highly uniform and black, then the process of adding in 2.5 mg of collagenase for every 1 mL of black pellet and incubation was repeated until uniformity was achieved. Once washed, the sample was air-dried and weighed to afford 41.2 mg of melanin pellet (11.2 mg from D1A-E and 30.0 mg of D1B-E) with 98 % of tissue mass lost, resulting in 1.8 % of the original mass.

D. DNAase treatment. Samples D1A-E and D1B-E were processed independently and not combined. Each sample was resuspended in H2O (25 mL) and 1.5 mg of deoxyribonuclease I (2 μM, 31 kDa) from bovine pancreas (EC number: 3.1.21.1, Sigma Aldrich) was added per mL of black solution. The sample was sonicated for 5 min (3 s on, 4 s off). The resulting suspension was allowed to precipitate by gravity for 24 h. The opaque supernatant was removed. The sample was washed three times with water and centrifugation (4200 rpm). The pellet was collected, air dried and weighed to afford 23.6 mg of DNA-free melanin (10.6 mg from D1A-F and 13.0 mg from D1B-F) with 43 % of the mass lost in this process, resulting in 1.0 % of the original mass.

E. Lipase treatment. Again individually, D1A-F and D1B-F were suspended in H2O (25 mL) and 1.5 mg of Amano lipase (2 μM, 33 kDa) from Pseudomonas fluorescens (EC number: 232.619.9, Sigma Aldrich) was added per mL of black solution. The sample was sonicated for 5 min (3 s on, 4 s off) and the suspension was allowed to precipitate by gravity. The opaque supernatant was removed and the sample was washed with water and centrifuged three times. The pellet was collected, air dried and weighed to afford 17.4 mg of melanin (7.8 mg from D1A-G and 9.6 mg from D1B-G with 26.2 % of the previous mass lost, resulting in 0.7 % of the original mass.

F. Gravimetric purification over an organic solvent gradient. The resulting pellets were joined by suspension in anhydrous MeOH (1 mL) as illustrated in Fig. 2.2f. Anhydrous CH2Cl2 (2.5 mL) was added until the black precipitate began to settle (Fig. 2.2g). After 15 min, the supernatant containing colorless particles was removed by pipette. The process was repeated 3×, with the third being transferred into a storage vial (see Fig. 2.2h-j). After air-drying, the pellet was weighed S4 to afford 16.2 mg of melanin (8.4 mg from D1A-H and 7.8 mg from D1B-H), indicating that Delphinus delphis skin contained 0.7 % m/m of melanin (wet weight).

G. Melanin purity monitoring through fluorescent staining. A 20 ± 10 µL sample was placed on a 25 × 75 × 1mm microscope glass slide (22038103, Fisher Scientific) and 5 µL of a 1 mM solution of the fluorescent stain dissolved in H2O containing 5% DMSO was added. The mixture was briefly stirred using a pipette tip and allowed to sit for 20 min. At this point, a rectangular cover glass (12-542B, Fisher Scientific) was placed on the sample. Images were collected on an EVOS fluorescent microscope using 20× magnification with excitation (λex) and emission (λem) at 531 ± 40 and 593 ± 40 nm respectively (Nile red), 357 ± 44 and 446 ± 60 nm (DAPI), or 470 ± 22 and 525 ± 50 nm (Alexa fluor 488 NHS). Images were processed in Photoshop (Adobe) with identical cropping and exposure adjustments (+1) for each image. Images were assembled into a figure using Animate (Adobe) and identical image processing and the display was used for all images. Raw image files can be provided upon request.

H. Melanin purity monitoring through SEM imaging. Dried melanin powder was placed on conductive carbon tape on an aluminum sample holder and coated using an iridium-sputter coating for 8 s. An FEI Quanta FEG 250 microscope was used for acquiring images using a 3-5 kV energy source with a spot size of 3 under high vacuum at a working distance of 9 mm. Images were processed in Photoshop (Adobe) with identical cropping for each image. Images were assembled into a figure using Animate (Adobe) and identical image processing and the display was used for all images. Raw image files can be provided upon request.

I. ssNMR-analyses. Data was acquired on a 750 1 H MHz spectrometer (Bruker). Spinning sidebands from the NMR were labeled in the spectra and do not overlap with peaks of interest. When there was uncertainty, the CP experiment was run at multiple spinning frequencies to verify peaks. All NMR spectra were recorded with complete proton decoupling.

FID files were processed using TopSpin 3.2 (Bruker) and MestreNova 6.0.2 (Mestrelab Research).

Dolphin melanin. Dried natural abundance melanin sample (~30 mg) was packed in 3.2 mm Bruker rotor. 1D 13C CP spectra of the precursor monomers were acquired with 12k scans and a recycle delays of 3 s. The CP contact times were optimized to 2 ms, with a 1 H decoupling of 83 kHz and a spinning frequency of 13.5 kHz. The intrinsic heterogeneous nature of melanin biopolymer results in broad line width in the NMR spectra, which cannot be improved with higher field, spinning frequency, or temperature of the experiments. The cross polarization experiment was chosen due do to its usefulness and applicability in what we wanted to achieve. Spectral editing and direct polarization experiments were not feasible due to the long acquisition time of the experiments.

Precursor monomers (L-Tyr, Tyramine, L-DOPA and Dopamine). Natural abundance monomers (30 mg) were packed in 3.2 mm Bruker rotor. 1D 13C CP spectra of the monomers were acquired with 512 scans and recycle delays of 3 s. The CP contact times were optimized to 2 ms, with a 1 H decoupling of 83 kHz and a spinning frequency of 13.5 kHz. Overlays of these spectra with the dolphin melanin 13C CP spectrum were used for structural assignment as shown in Fig. 2.5 of the manuscript.

J. Supporting Spectra. The following Figures provide supporting spectral analyses and assignments.

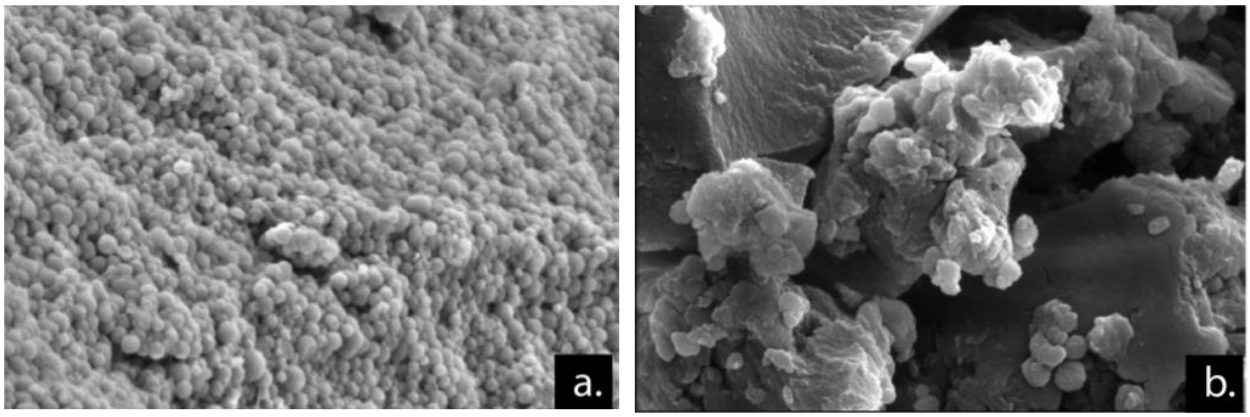


Figure S2.1: Octopus ink extracted and purified enzymatically. b). Octopus ink from the same organism purified with 6M HCl, neutralized with base showing lost nanoparticle features.

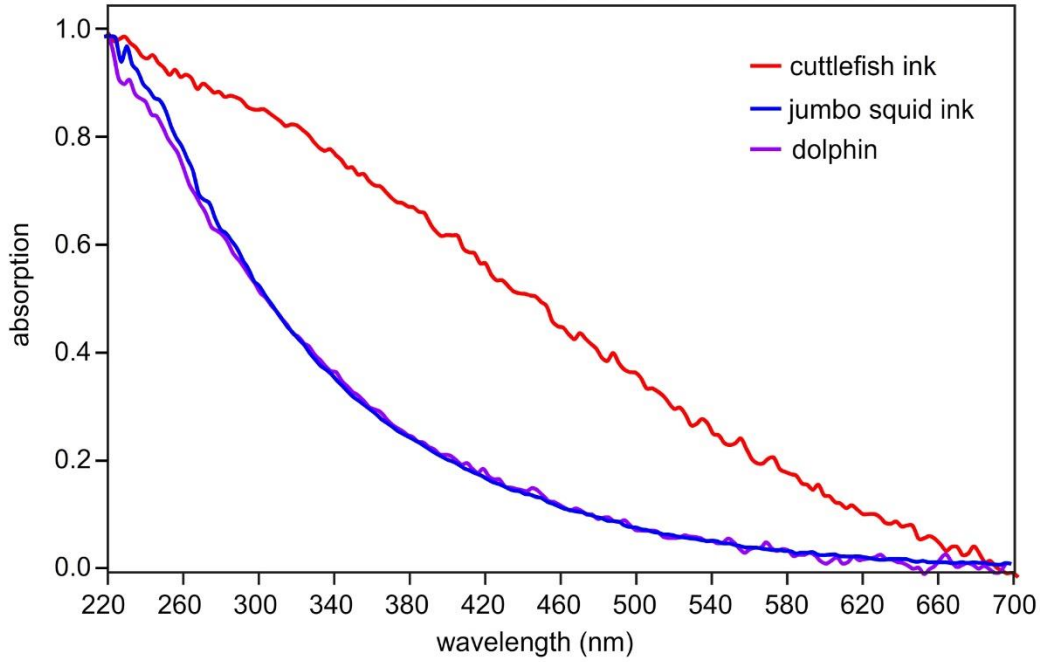


Figure S2.2: UV Visible spectra depicting the comparison between purified samples of jumbo squid ink (blue) and cuttlefish ink (red) with the dolphin sample prepared in this study (purple).

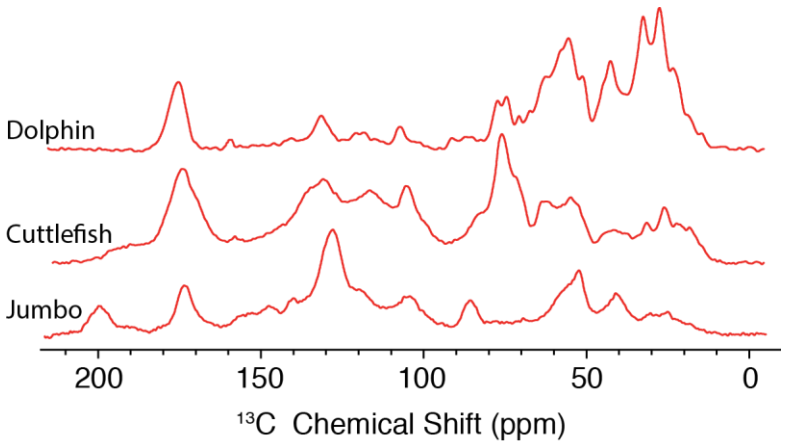


Figure S2.3: ss-13C-NMR spectra depicting the comparison between purified samples of jumbo squid ink (jumbo) and cuttlefish ink (cuttlefish) with the dolphin sample prepared in this study).

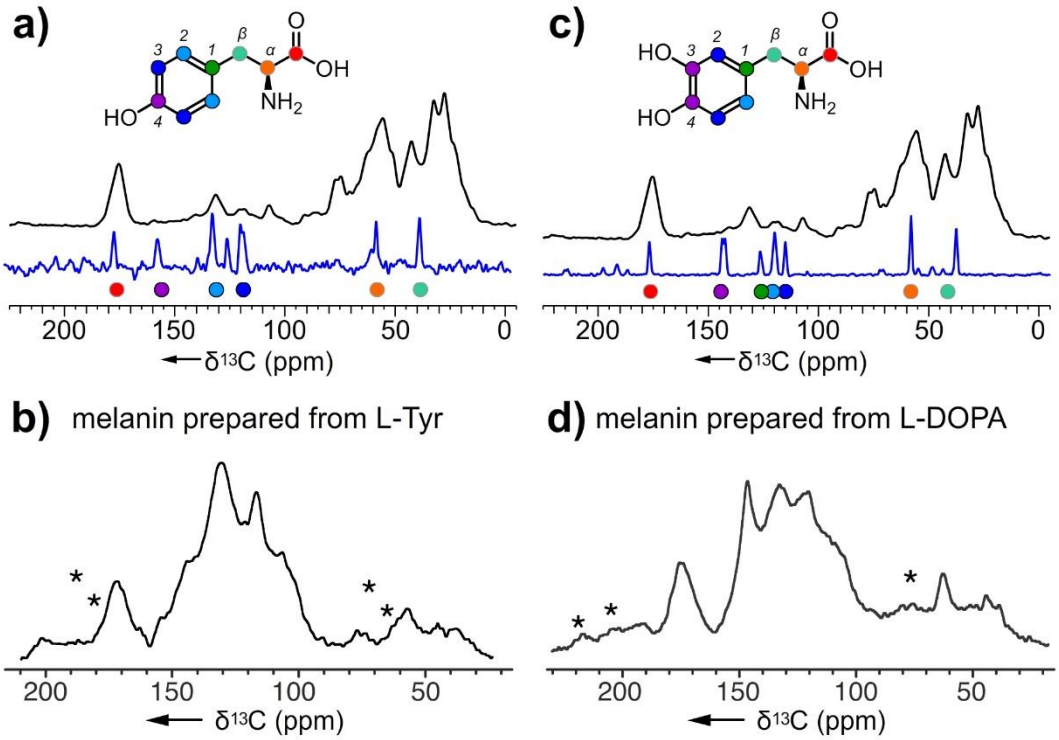


Figure S2.4: ss-13C-NMR spectra depicting the comparison the dolphin melanin and melanin prepared in vitro by Agaricus bisporus (mushroom) tyrosinase. a) copy of panel c from Fig. 2.5, b) ss-13C-NMR obtained from melanin prepared by the action of Agaricus bisporus (mushroom) tyrosinase on L-tyrosine. c) copy of panel e from Fig. 2.5, d) ss-13C-NMR obtained from melanin prepared by the action of Agaricus bisporus (mushroom) tyrosinase on L-DOPA. Methods and further details on the preparation of the melanins in panels b and d can be found in Ni QZ, Sierra BN, La Clair JJ, Burkart MD. Chemoenzymatic elaboration of the Raper-Mason pathway unravels the structural diversity within eumelanin pigments. Chem Sci. 2020, 11, 7836-7841.

2.4 Acknowledgements

Chapter 2, in full, is a reprint of the material as it appears, Kalaj, Brianna N.; Ni, Qing Zhe.; La Clair, James J.; Burkart, Michael D. “Chemoenzymatic isolation and characterization of high purity mammalian melanin.” *ChemBioChem*, **2022**, *23*, e202200021. The dissertation author was the co-primary investigator and author of this paper.

2.5 References

1. a) P. A. Riley, *Int. J. Biochem. Cell Biol.* **1997**, *29*, 1235-1239; b) S. A. D'Mello, G. J. Finlay, B. C. Baguley, M. E. Askarian-Amiri, *Int. J. Mol. Sci.* **2016**, *17*, 1144.

2. a) J. J. Negro, C. Finlayson, I. Galván, *Int. J. Mol. Sci.* **2018**, *19*, 230; b) C. M. Eliason, M. D. Shawkey, J. A. Clarke, *Evolution* **2016**, *70*, 445-455; c) Q. Li, J. A. Clarke, K. Q. Gao, C. F. Zhou, Q. Meng, D. Li, L. D’Alba, M. D. Shawkey, *Nature* **2014**, *507*, 350-353; d) F. Zhang, S. Kearns, P. Orr, M. J. Benton, Z. Zhou, D. Johnson, X. Xu, X. Wang, *Nature* **2010**, *463***,** 1075-1078; e) K. Glass, S. Ito, P. R. Wilby, T. Sota, A. Nakamura, C. R. Bowers, J. Vinther, S. Dutta, R. Summons, D.E.G. Briggs, K. Wakamatsu, J. D. Simon, *Proc. Nat. Acad. Sci U.S.A.* **2012**, *109*, 10218-10223

3. a) W. Cao, X. Zhou, N. C. McCallum, Z. Hu, Q. Z. Ni, U. Kapoor, C. M. Heil, K. S. Cay, T. Zand, A. J. Mantanona, A. Jayaraman, A. Dhinojwala, D. D. Deheyn, M. D. Shawkey, M. D. Burkart, J. D. Rinehart, N. C. Gianneschi, J. *Am. Chem. Soc.* **2021**, *143*, 2622-2637; b) M. E. McNamara, V. Rossi, T. S. Slater, C. S. Rogers, A. L. Ducrest, S. Dubey, A. Roulin, *Trends Ecol. Evol.* **2021**, *36*, 430-443; c) L. D’Alba, M. D. Shawkey, *Physiol. Rev.* **2018,** *99*, 1-19.

4. D. J. Field, L. D’Alba, J. Winter, S. M. Webb, G. Gearty, M. D. Shawkey, *PLoS One* **2013**, *8*, e59451.

5. a) H. S. Raper, *Biochem. J.* **1927,** *21*, 89-96; b) H. S. Mason, *J. Biol. Chem.* **1948**, *172*, 83-99; c) L. Panzella, A. Ebato, A. Napolitano, K. Koike, *Int. J. Mol. Sci.* **2018**, *19*, 1753.

6. a) R. Micillo, L. Panzella, K. Koike, G. Monfrecola, A. Napolitano, M. d'Ischia, *Int. J. Mol. Sci.* **2016**, *17*, 746; b) J. D. Simon, D. N. Peles, *Acc. Chem. Res.* **2010**, *43*, 1452-1460.

7. a) L. Zecca, D. Tampellini, M. Gerlach, P. Riederer, R. G. Fariello, D. Sulzer, *Mol. Pathol.* **2001**, *54*, 414-418; b) F. A. Zucca, E. Basso, F. A. Cupaioli, E. Ferrari, D. Sulzer, L. Casella, L. Zecca, *Neurotox. Res*. **2014**, *25*, 13-23.

8. a) X. Zhou, X. Gong, W. Cao, C. J. Forman, J. Oktawiec, L. D'Alba, H. Sun, M. P. Thompson, Z. Hu, U. Kapoor, N. C. McCallum, C. D. Malliakas, O. K. Farha, A. Jayaraman, M. D. Shawkey, N. C. Gianneschi, *Angew. Chem. Int. Ed. Engl.* **2021**, *60*, 17464-17471; b) H. C. Eisenman, A. Casadevall, *Appl Microbiol Biotechnol*, **2012**, *93*, 931-940.

9. a) M. Al Khatib, J. Costa, D. Spinelli, E. Capecchi, R. Saladino, M. C. Baratto, R. Pogni, *Int. J. Mol. Sci.* **2021**, *22*, 1739; b) J. Schmaler-Ripcke, V. Sugareva, P. Gebhardt, R. Winkler, O. Kniemeyer, T. Heinekamp, A. A. Brakhage, *Appl. Environ. Microbiol.* **2009**, *75*, 493-503.

10. a) F. Solano, *New J. Sci.*, **2014,** *2014*, 1-28; b) J. D. Nosanchuk, R. E. Stark, A. Casadevall, *Front. Microbiol*. **2015**, 6, 1463.

11. a) Q. Z. Ni, B. N. Sierra, J. J. La Clair, M. D. Burkart, *Chem. Sci*, **2020**, *11*, 7836-7841; b) S. Chatterjee, R. Prados-Rosales, B. Itin, A. Casadevall, R. E. Stark, *J. Biol. Chem.* **2015**, *290*, 13779-13779; c) J. Zhong, S. Frases, H. Wang, A. Casadevall, R. E. Stark, *Biochem*. **2008**, *47*, 4701-4710; d) S. Tian, J. Garcia-Rivera, B. Yan, A. Casadevall, R. E. Stark, *Biochem*. **2003**, *42*, 8105-8109.

12. a) C. Xin, J. H. Ma, C. J. Tan, Z. Yang, F. Ye, C. Long, S. Ye, D. B. Hou, *J. Biosci. Bioeng.* **2015**, *119*, 446-454; b) J. B. Nofsinger, S. E. Forest, L. M. Eibest, K. A. Gold, J. D. Simon, *Pigment Cell Res*, **2000**, *13*, 179-184.

13. a) I. E.Pralea, R. C. Moldovan, A. M. Petrache, M. Ilieș, S. C. Hegheș, I. Ielciu, R. Nicoară, M. Moldovan, M. Ene, M. Radu, A. Uifălean, C. A. Iuga, *Int. J. Mol. Sci*. **2019**, 20, 3943; b) Y. Liu, J. D. Simon, *Pigment Cell Res*. **2003**, 16, 606-618; c) Y. J. Kim, M. J. Kim, D. K. Kweon, S. T. Lim, S. J. Lee, *J. Vis. Exp.* **2019**, *145*, e58185.

14. a) I. E. Pralea, R. C. Moldovan, A. M. Petrache, M. Ilieș, S. C. Hegheș, I. Ielciu, R. Nicoară, M. Moldovan, E. Ene, M. Radu, A. Uifălean, C.A. Iuga, *Int. J. Mol. Sci.* **2019**, 20, 3943; b) A. Filson, J. Hope, *Nature* **1957**, *179*, 211.

15. a) S. Ito, *Biochem. Biophys. Acta*. **1986**, *883*, 155-161; b) S. Ghiani, S. Baroni, D. Burgio, G. Digilio, M. Fukuhara, P. Martino, K. Monda, C. Nervi, A. Kiyomine, S. Aime, *Mag. Reson. Chem*. **2008**, *46*, 471-479.

16. Y. Liu, V. R. Kempf, J. B. Nofsinger, E. E. Weinert, M. Rudnicki, K. Wakamatsu, S. Ito, J. D. Simon, *Pigment Cell Res*. **2003**, 16, 355-365.

17. a) Y. Liu, V. R. Kempf, J. B. Nofsinger , E. E. Weinert, M. Rudnicki, K. Wakamatsu, S. Ito, J. D. Simon JD. *Pigment Cell Res*. **2003**, *16*, 355-365. b) M. R. Chedekel, B. L. Murr, L. Zeise, *Pigment Cell Res*. **1992**, *5*, 143-147; c) A. Wilczek, H. Kondoh, Y. Mishima, *Pigment Cell Res*. **1996**, *9*, 63-67; d) K. L. Double, L. Zecca, P. Costi, M. Mauer, C. Griesinger, S. Ito, D. Ben-Shachar, G. Bringmann, R. G. Fariello, P. Riederer, M. Gerlach, J. *Neurochem*. **2000**, *75*, 2583-2589.

18. L. Novellino, A. Napolitano, G. Prota, *Biochim. Biophys. Acta*. **2000**, *1475*, 295-306.

19. a) D. N. Hu, S. A. McCormick, J. A. Seedor, D. C. Ritterband, M. K. Shah, *Exp. Eye Res*. **2007**, *84*, 655-662; b) D. N. Hu, *Photochem. Photobiol*. **2008**, *84*, 645-649.

20. C. T. Chen, V. Ball, J. J. de Almeida Gracio, M. K. Singh, V. Toniazzo, D. Ruch, M. J. Buehler, *ACS Nano*. **2013**, *26*, 1524-1532.

21. M. d'Ischia, K. Wakamatsu, A. Napolitano, S. Briganti, J.-C. Garcia-Borron, D. Kovacs, P. Meredith, A. Pezzella, M. Picardo, T. Sarna, J. D. Simon, S. Ito, *Pigment Cell Melanoma Res*. **2013**, *26*, 616-633

22. a) I. Mani, V. Sharma, I. Tamboli, G. Raman, *Pigment Cell Res*. **2001**, *14*, 170-179; b) S. Ghiani, S. Baroni, D. Burgio, G. Digilio, M. Fukuhara, P. Martino, K. Monda, C. Nervi, A. Kiyomine, S. Aime, Magn. Reson. Chem. **2008**, *46*, 471-479.

23. a) S. Ito, K. Wakamatsu, *Pigment Cell Res*. **2003**, 16, 523-531; b) Wakamatsu, K. Ito, S. *Pigment Cell Res*. **2002**, 15, 174-183.

24. a) M. E. Pavan, N. I. López, M. J. Pettinari, *Appl. Microbiol. Biotechnol.* **2020**, *104*, 1357-1370; b) S. Singh, S. B. Nimse, D. E. Mathew, A. Dhimmar, H. Sahastrabudhe, A. Gajjar, V. A. Ghadge, P. Kumar, P. B.. Shinde, *Biotechnol. Adv*. **2021**, 107773; c) J. Lindgren, A. Moyer, M. H. Schweitzer, P. Sjövall, P. Uvdal, D. E. Nilsson, J. Heimdal, A. Engdahl, J. A. Gren, B. P. Schultz, B. P. Kear, *Proc. Biol. Sci*. **2015**, *282*, 20150614.

25. a) D. F. Q. Smith, A. Casadevall, *Curr. Top. Microbiol. Immunol.* **2019**, *422*, 1-30; b) H. C. Eisenman, A. Casadevall, *Appl. Microbiol. Biotechnol.* **2012**, *93*, 931-940; c) J. D. Nosanchuk, R. E. Stark, A. Casadevall, *Front .Microbiol*. **2015**, *6*, 1463.

26. A. Slominski, D. J. Tobin, S. Shibahara, J. Wortsman, *Physiol. Rev.* **2004**, *84*, 1155-1122.

Chapter 3 Investigating Methylene Diphenyl Diamine Alteration and Uptake by Soil and Compost Organisms

3.1 Introduction

Over the last seven decades, the world has seen a tremendous surge in the generation of plastic waste with an estimated 8 billion metric tons having been generated as of 2015, 79% of which has accumulated in landfills and the environment.[1] Due to its durability, plastics and the microplastics generated from these wastes are persisting in the environment and working their way into our bodies through the food chain.[2] Microplastics ingestion leads to adverse health effects including poor metabolism and reduced fertility, and chemical leaching from plastics can also lead to endocrine disruption.[2] The ubiquitous presence of litter in oceans and landfills has highlighted the need for biodegradable plastics from renewable feedstocks, which has come to fruition recently with bio-based polyester polyurethanes (PUs).[3,4] Polyester polyurethanes are largely biodegraded using esterases, ureases, amidases, and proteases from bacteria and fungi.[5] Through the hydrolysis and breaking of ester and urethane bonds, the microbial community can access the monomeric segments of the polymer for food.[6,7]

Characteristically, PUs contain repeating urethane linkages produced from a polyol and a diisocyanate, though the identity of these molecules can vary greatly depending on the desired properties of the product. A common diisocyanate incorporated into PUs is 4,4'-methylenediphenyl diisocyante (MDI), a petroleum product used to give the polymer rigidity due to its aromatic structure. The hardness provided by MDI's diphenyl structure unfortunately exacerbates the challenges relating to its biodegradability. Recent developments in biodegradable PUs are a unique opportunity to reduce the amount of forever plastics with eco-friendly alternatives. With these promising new PUs in their ability to stop microplastic persistence in the environment, it does in turn raise the question of component toxicity, where these monomers were

not previously believed to be released. One hypothesized degradation product of MDI-based polyurethanes is 4,4'-methylenediphenyl diamine (MDA), predicted to be released during the hydrolysis of the urethane bonds of a PU forming after water cleaves the ester group, followed by decarboxylation.[8,9] MDA is considered a potential human carcinogen and an environmental risk, thus examining how this intermediate is processed is vital to evaluating the environmental impact and safety of biodegradable PU formulations that use MDI as the diisocyanate component.[10] Although various bacteria including *Acinetobacter*, *Corynebacterium*, and members of the genus *Pseudomonas* have been extensively studied for PU degradation, these studies have provided no conclusive evidence on what happens to MDA during biodegradation.[11,12] There has also been a recent fungal study published, however the assertion that an organism consumes MDA was not fully supported by the data and the approach outlined in this work takes a different aspect while not fully convinced there is an organism out there which consumes this monomer.[13] Though there has been no conclusive evidence yet on MDA consumption, there is observation of yellowing in PUs which occurs over time, which can be attributed to UV breakdown of aromatics within PU backbones through a photo-fries rearrangement followed by quinone formation. [14,15] Given this, if there is potential for MDI-based polyurethanes to generate MDA that could persist in the environment, it is a barrier to generating 100% biodegradable PU materials and there is still a need to determine if this small molecule gets metabolized or if it is toxic to microbial organisms.

To better understand how MDA is utilized by microorganisms responsible for PU degradation, this study set out to investigate if the organisms that can grow using these PUs as their sole carbon source could also breakdown MDA itself. Understanding if organisms in a soil or compost environment can biologically or abiotically remove MDA from the metabolic pool can help guide future biodegradation studies of PUs and innovation in modification of bacteria toward

MDA consumption if needed. In this work a multi-faceted approach was employed to analyze the biodegradation and modification of MDA, by PU consuming bacteria including activity assays and passaging studies. Mass spectrometry, NMR, and metagenomic sequencing, were able to provide an in depth model of the biological and physiochemical modifications of MDA that can occur in a terrestrial PU biodegradation environment. Herein, it is demonstrated that the presence of at least one of three bacteria are needed to survive in the presence of MDA and have potential to breakdown the monomer.

3.2 Methods and Materials

3.2.1. Passaging experiment for analysis of environmental consortia growth utilizing PU monomers in the presence of MDA

Organisms previously identified to grow on PU formulations[3] were streaked from frozen glycerol stocks and grown up on Luria Broth-Agar plates. Colonies were then picked and grown up in 100 mL of Luria Broth (LB) media for 5 days shaken on a shake table at a low speed to produce each starting culture. All starting cultures were combined at an OD_{600} value of 0.05 into 50 mL total culture volume along with M9 salt media with 0% dextrose for the inoculant culture to acclimate to M9 salt media while being shaken on a shake table at a low speed. After 3 days this culture was then used as an inoculant at an OD_{600} value of 0.01 into a total culture volume of 25 mL of M9 salt media, M9 salt media with 25 mg/L MDA, or M9 salt media with PU monomers without MDA, M9 salt media with PU monomers with 25 mg/L MDA. M9 salt media with 25 mg/L MDA were made 24 hours before needed in order to solubilize the small molecule. These experiments were done in triplicate and shaken on a shake table at a low speed. OD_{600} values were taken on day 0, 7, and 14 to measure how well each culture was growing. For each time point 1 mL of each culture was transferred to 1.5 mL Eppendorf tube and frozen for Orbitrap MS studies. After an OD_{600} value was taken on day 14, this solution was used as an inoculant in a fresh solution

of appropriate M9 salt media for a total volume of 25 mL and a total culture OD_{600} of 0.01 for the next passage of the experiment. This was repeated for 4 consecutive passages over 8 weeks of this experiment.

3.2.2. Metagenomic sequencing of organisms after passaging experiment

Cell pellet samples from the final day of passaging, Day 56, were collected by Swinnex filtration of the 25mL sample. DNA was then extracted using a QIAGEN DNeasy PowerSoil Pro DNA Kit (cat# 47014) on the cellulose filter membranes from the filter collection. DNA was analyzed for purity using $OD_{260/280}$ measurements reading a 1.87-1.89 for all samples and concentrations were taken with an Invitrogen Qubit 2.0 Fluorometer (cat# Q32866). Initial readings were saturated at >600 ng/mL and were diluted until final concentrations of 40-70 ng/mL were obtained. 16S sequencing was conducted and analyzed by Novogene.

3.2.3. Activity assay to determine uptake of monomers released from PU by organisms

To identify what organisms from the passaging study can degrade each monomer within a PU formulation, a plate assay was created in which organisms previously identified were grown individually in 10 mL of LB media for 5 days shaken on a shake table at a low speed to produce each starting culture. These cultures were then spun down and the supernatant was removed from the pellet. Pellets were resuspended in M9 minimal media to be washed then pelleted and the supernatant was discarded again, this was repeated to thoroughly wash the rich media from the organisms. Each pellet was then resuspended in 1 mL of M9 minimal media to be used as an inoculant for the activity assay. Each small molecule was dissolved in 10 mL M9 minimal media at 1 g/L concentration. To set up the assay, 100 uL of small molecule solution was added to 42

wells of a 48 well plate, followed by dilution up to 1 mL by M9 media creating a 100 mg/L solution. Each of the 13 organisms was used to inoculate individual wells in triplicate, as well as a blank sample. OD_{600} values were taken of each plate for a day 0 time point, followed by shaking on a shake table at room temperature and low speed for 1 week to allow for growth. OD_{600} values were taken of each plate on day 7 and day 14 to monitor growth.

3.2.4. Individual microorganism uptake study with MDA

Pseudomonas aeruginosa was identified as a candidate for being able to survive in the presence of MDA and was therefore studied individually for MDA alteration. *Pseudomonas aeruginosa* was inoculated in M9 media with 0.4% dextrose in 125 mL flasks. The flasks were shaken at room temperature for 3 days to promote growth, indicated by increased turbidity. Cultures were then spun down, and previous media was replaced with either M9 salt media with 0% dextrose for same day- spike in controls or M9 salt media with 0% dextrose and 3 mg/L MDA for experimental. Five mL fractions of each culture were allotted into 16 mL culture tubes in triplicate and shaken on a shake table at a low speed for the duration of the study. Time points were taken on day 0, 2, and 3. For each time point 1 mL of each culture was transferred to 5 mL Eppendorf tubes and spun down.

Same day-spike in controls were determined necessary to account for matrix effect issues between MDA and the media in the cells. For these controls, to the spun down cultures, 10x solution of MDA (100 mg/L) was spiked in at a 1:10 ratio, in order to account for matrix effects of the media to cells. Samples were then centrifuged in order to pellet cells and the samples were diluted 1:10 in MilliQ water to be analyzed by Orbitrap LCMS.

3.2.5. Identification of individual microorganism MDA conversion products using Liquid Chromatography-Mass Spectrometry

4,4-methylenediamine (97% purity), 4',4''-methylenebis(acetanilide) and N-(4-(4-Aminobenzyl)phenyl)acetamide were purchased from Sigma Aldrich. $^{13}C,^{15}N$ labelled 4,4-methylenediamine was purchased from Iso Sciences. Unlabeled standards in M9 media were prepared from 0.5-10 mg/L and shaken overnight to promote full dissolution. Each standard was diluted 1:10 in MilliQ water and run on a Thermo Fisher Alliance HPLC/Orbitrap Elite with an ACE PFP C18 column 150 x 2.1 mm, 3µm particle size. A 7 minute isocratic method was run using 60% methanol and 40% 0.1% formic acid buffer pH 3.3 and a flow rate of 0.21 mL/min. The column temperature was set at 45 ̊C.The Orbitrap HESI was run in positive mode at 3.5 eV. The capillary temperature was set at 350 ̊C. Detection parameters for the FTMS was at a resolution of 15000 full scan mode between 90-1500 m/z. Five data dependent MS2 scans were run between 100-1500 m/z. MDA, MMDA, and DMDA calibration curves were created using peak areas of the extracted ion chromatograms (M+H) for m/z values 199.12, 241.13, and 283.14. Identities were also verified by MS/MS. Retention times were determined to be 2.50, 3.09, and 3.85 mins, limit of quantification for each was 0.5 mg/L, and coefficients of determination were 0.997, 0.979, and 0.996, respectively.

3.2.6. NMR analysis of individual organism MDA conversion products

Whole cell samples were taken of the organisms inoculated with 3 mg/L $^{13}C,^{15}N$-MDA, purchased from Iso Sciences, over time. Time points of the experiment were taken on day 0, 2, and 3. For each time point, 1.3 mL of each culture were transferred to Eppendorf tubes and sonicated on ice for 20 s (2 s on, 3 s off). Samples were then lyophilized and stored at -20 ̊C until

used for NMR experiments. In order to prepare samples to go on the magnet, they were defrosted and brought to room temperature as 0.5 mL of deuterated chloroform was added. Organic matter that was solubilized was then transferred to a 5mm NMR tube for experiments. 1D ^{15}N INEPT spectra of the sample contents were acquired with 3k scans and a recycle delay of 5 s at 298 K. Data was acquired on an 800 MHz Bruker Avance.

3.2.7. Soil and compost consortia uptake study with DMDA

$^{13}C,^{15}N$-DMDA purchased from Iso Sciences, was used for these studies. Six organisms constituting each consortia, with soil consortia being made up of *Heraspiriliums* SP2 and SP3, *Rhodococcus'* SP1 and SLB2, *Ochrobactrum* SLB1, and *Pseudomonas* F5 and compost consortia being made up of *Pseudomonas* CP1, *Ochrobactrums* CP2, CLB2, and CLB5, *Achromobacter* CLB3, and *Stenotrophomonas* CLB4 were streaked and grown up on LB-Agar plates. Colonies were then picked and grown up in LB to produce each inoculant starting culture. Consortia were inoculated in M9 media with 0.4% dextrose in 125 mL flasks. The flasks were shaken at room temperature for 3 days to promote growth, indicated by increased turbidity. Cultures were then spun down, and previous media was replaced with either M9 salt media with 0% dextrose for same day- spike in controls or M9 salt media with 0% dextrose and either 3 or 10 mg/L diacetylated methylene diphenyl diamine (DMDA) for experimental studies. Five mL fractions of each culture were aliquoted into 16 mL culture tubes in triplicate and shaken on a shake table at a low speed for the duration of the study. Time points were taken on day 0, 2, 5, and 14. For each time point 1 mL of each culture was transferred to 5 mL Eppendorf tubes and spun down.

3.2.8. Liquid Chromatography-Mass Spectrometry analysis of consortia DMDA conversion

Using the same LCMS/Orbitrap method as previously described, the ^{15}N-labelled MDA, MMDA, and DMDA peaks were identified by M+H m/z values of 202.12, 243.13, and 286.14, respectively and confirmed by MS/MS.

3.2.9. NMR analysis of consortia DMDA conversion

Whole cell samples were taken of the organisms inoculated with 3 or 10 mg/L ^{13}C,^{15}N-DMDA over time. Time points of the experiment were taken on day 0, 2, and 5 for the first compost and soil studies with 3 mg/L DMDA or day 0 and 14 for 14 day soil studies with 10 mg/L DMDA. For each time point, 1.3 mL of each culture were transferred to Eppendorf tubes and sonicated on ice for 20 s (2 s on, 3 s off). Samples were then lyophilized and stored at -20 C until used for NMR experiments. In order to prepare samples to go on the magnet, they were defrosted and brought to room temperature as 0.5 mL of deuterated chloroform was added. Organic matter that was solubilized was then transferred to a 5mm NMR tube for experiments. 1D ^{15}N INEPT spectra of the sample contents were acquired with 3k scans and a recycle delay of 5 s at 298 K. Data was acquired on an 800 MHz Bruker Avance.

3.3 Results and Discussions

3.3.1 Co-metabolism passaging experiment for analysis of consortia growth utilizing PU monomers in the presence of MDA

As a community of organisms is typically present in any compost or soil environment bio-based polyurethanes would be subjected to for biodegradation, this study set out to determine the viability of a consortia of organisms in the presence of MDA over an elongated time period. In order to ensure the organisms had a representative pool of small molecules similar to what would constitute a polyurethane foam's breakdown products this was conducted as a co-metabolism study

with the monomers used in the relevant foam formulation. During the first two passages of this study it is seen in Figure 1a that the growth of the consortia levels off after the first week and it is expected that the organisms were acclimating to using the monomers available as a metabolite source and began competing with one another for survival. As passage 3 began, continued positive growth was seen over the two week period. As it was observed in both the 3rd and 4th passages (Figure 3.1), it is assumed the consortia has self-regulated and the culture is productively balanced as it begin the uptake of each of the monomers. The growth of the organisms over time in the experimental samples can be seen to be surviving at a comparable rate to the control experiment without MDA present. This led to the conclusion that MDA is being processed in one of three ways, it is either being consumed if able, or if toxic, being removed from the organism's metabolite pool or being physically removed through insolubilizing the small molecule, as some precipitates were seen during the passages.

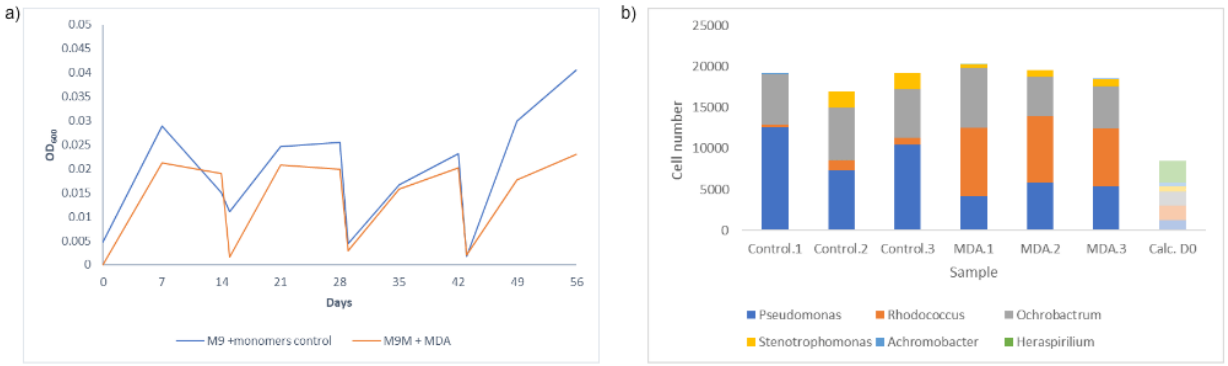


Figure 3.1: Passaging and sequencing data. A) 8 week passaging study of relevant PU monomers and MDA in M9 minimal media. B) 16S sequencing data for control of solely organisms with monomers in triplicate, organisms with monomers and MDA in triplicate, and a calculated Day 0 time point for comparison.

3.3.2. Metagenomic sequencing of organisms after passaging experiment

To achieve the goal of identifying MDA's ecotoxicology on the consortia of organisms previously identified to feed on PU foams, 16S metagenomic sequencing was utilized to analyze the change in PU monomers over time were used. After 8 weeks and 4 separate passages, it was hypothesized that the bacteria remaining was able to survive in MDA's presence and either utilize MDA or physiochemically or biologically remove MDA. Figure 1b displays which bacteria were identified their relative amounts as enriched through the passaging studies. After calculating a day 0 time point based off of the original composition of the starting culture, through taking the average original OD_{600} for the samples and converting that to cell number, followed by relative amount adjustment based on OD_{600} value of both Day 0 and Day 56, the change in absolute cell number in relation to the blank samples and experimental samples can be observed. Cell numbers for the data were calculated by taking absolute abundances and accounting for 16S copy number of each organism. In comparing cell number of day 0 starting point vs day 56 end point for the passaging studies, it can be observed that the cultures did significantly grow and almost doubled the cell counts raising from around 8,500 calculated day 0 to 19,000 on average for control and experimental day 56. Another positive within the sequencing data is the cell numbers staying relatively consistent across experimental and control samples, leading to reliable comparisons. Notably, the *Rhodococcus* was greatly enriched in the MDA containing samples compared to the control samples and the *Pseudomonas* and *Ochrobactrum* species seemed to still be present in large quantities unlike *Stenotrophomonas*, *Achromobacter* and *Heraspirillium* which either disappeared or died off. This led to the consideration of three organisms, *Rhodococcus*, *Pseudomonas,* and *Ochrobactrum* as main candidates for allowing a consortia to adapt and grow in the presence of MDA.

3.3.3. Activity assay to determine uptake of monomers released from PU by organisms

In order to determine what microorganisms were responsible for degradation of each monomer within the foam formulation and identify organisms that possibly grew solely on MDA, a plate assay was created. All organisms from compost or soil that had originally been shown to consume foam and were used in the passaging study were fed the individual diol and diacid monomers of a polyurethane foam formulation along with MDA. The activity assay showed that each organism had one if not two major monomers it consumed as illustrated in Figure 2. The *Pseudomonads* favoring diacid 1 and diols 1 and 3, and *Heraspirillums* also favoring diacid 1 and diol 1. *Rhodococcus* while also favoring diol 1 also tended to grow better on diacid 2 rather diacid 1, opposite to *Achromobater* which favored diol 2 and diacid 1. While the variety of *Ochrobactrum* tended to each favor a diol, with organism CLB2 having the best growth out of all organisms on diol 3, SLB1 and CP2 also grew well on all diacids. *Stenotrophomonas* had the best growth on all diacids and diols as a whole, while having the second best growth overall in the assay on diol 3.

		diacid 1	diacid 2	diacid 3	diol 1	diol 2	diol 3	MDA
Pseudomonas	F5	0.003	0.204	-0.002	0.070	-0.156	0.143	0.023
	CP1	0.032	0.039	-0.020	0.177	-0.037	0.055	0.008
Heraspirilium	SP2	0.299	-0.002	-0.022	0.154	-0.123	0.009	0.012
	SP3	0.001	0.009	-0.048	0.087	-0.036	0.016	0.008
Rhodococcus	SP1	0.005	0.177	0.028	0.155	-0.075	-0.035	0.033
	SLB2	0.015	0.131	0.001	0.044	0.034	-0.015	0.008
Ochrobactrum	SLB1	0.145	0.201	0.041	-0.013	0.170	-0.042	-0.028
	CP2	-0.003	0.025	0.286	0.317	0.011	0.083	-0.034
	CLB2	0.033	0.046	-0.051	0.200	-0.002	0.571	-0.042
	CLB5	0.042	0.113	-0.033	-0.019	0.268	-0.018	0.049
Achromobacter	CLB3	0.143	0.026	0.012	0.010	0.199	0.001	0.003
Stenotrophomonas	CLB4	0.045	0.140	0.023	0.050	0.165	0.371	-0.003
	blank	0.016	0.011	0.097	0.014	-0.029	0.064	0.003

Figure 3.2: Activity Assay Study. Results from monomer feeding study, heat map with red to green relating to negative to positive growth respectively.

With this data in hand, it was determined that each monomer had an organism or two which could reliably consume it, with MDA being the only monomer in the renewable PU foam formulation that does not have a microorganism observed to have significant enrichment on the monomer alone over time. However, *Rhodococcus* SP1, *Pseudomonas* F5, and *Ochrobactrum* CLB5 showed the most potential for growing on MDA alone, matching up well with the sequencing results for which organisms seemed to survive the best over a time period of 8 weeks.

3.3.4. Individual microorganism uptake study with MDA

The initial studies involved using *Pseudomonas aeruginosa* for a first look at determining how the small molecule was getting metabolized or altered within the biodegradation of renewable polyester polyurethanes. *Pseudomonads* have been known to have acetylcholinesterases and C-acyltransferases which catalyze the acetylation of amines and alcohols, leading to us using this as a starting point.[16,17] With *Pseudomonads* already having infrastructure to modify amines within their metabolic pools, as well as having seen positive growth of this organism on MDA within the activity assay with little negative growth on other monomers as well as positive data from the 16S sequencing data on the passaging studies done before, this was a preferable starting point.

First performing an LCMS study, over the course of 4 days the media containing *Pseudomonas aeruginosa* and unlabeled MDA experienced a decrease in MDA concentration at 2.5 mins and new peaks were seen as peaks on the LCMS chromatogram at 3.1 and 3.8 mins, with the corresponding m/z values 241.13 and 283.14. Chemical formula calculations predicted these two m/z values to have the formulas $C_{15}H_{17}N_2O$ and $C_{17}H_{19}N_2O_2$ to within 5 ppm, which corresponded to the M+H adducts of the mono- and di- acetylated MDA products.

In follow up to the unlabeled MDA study with Pseudomonas, we performed a study with 10 mg/L ^{13}C, ^{15}N-labeled MDA and F5, to observe if any significant peak shifts using ^{15}N NMR (Figure 3.1) were able to be visualized. At the 0-hour time point we see the peak for MDA with its free amines corresponding to the peak within the proper range at 55ppm. By 24 hours, two peaks were observed, the previously seen amine peak for ^{15}N-MDA at 55 ppm, and a new peak at 130 ppm, in the range of amide bonds, which was a shift from the standard signal (Figure 3.3). After 72 hours, the 55 ppm peak corresponding to the starting amine was no longer present, instead all detectable nitrogen in the sample was present in a single peak at 130 ppm. This peak shift signifies the movement of the amine into the range of amide peaks for nitrogen by ^{15}N NMR chemical shifts.

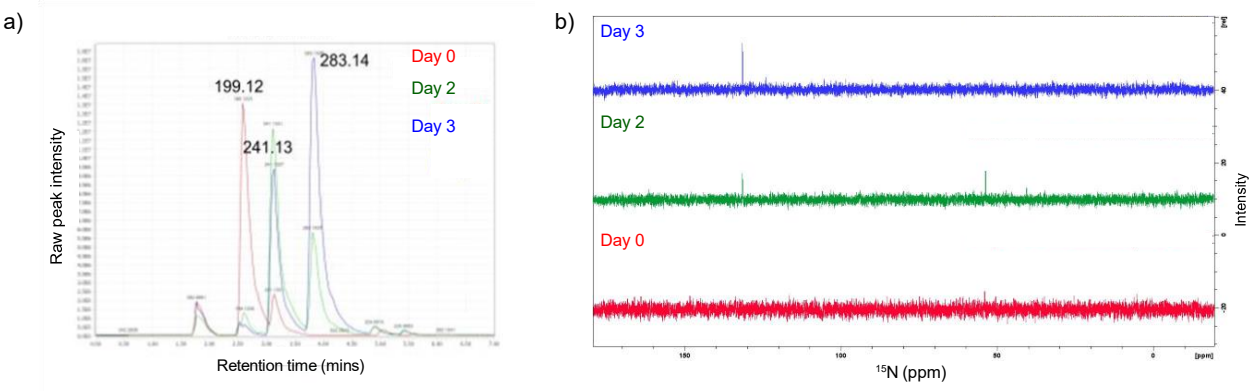


Figure 3.3: Pseudomonas aeruginosa studies for MDA uptake a) LCMS data over 0 days (red), 2 days (green), and 3 days (blue) b) ^{15}N labeled MDA NMR study shown over 0 days (red), 2 days (green), and 3 days (blue)

In a corresponding test using the ^{13}C, ^{15}N-labeled MDA as a substrate for MS, the labeled MDA peak at 202.12 m/z value decreased and two new peaks having m/z values 244.13 and 286.14 appeared. These two peaks corresponded to the chemical formulas $^{13}CC_{14}H_{17}{}^{15}N_2O$ and

$^{13}CC_{16}H_{19}{}^{15}N_2O_2$ to within 5 ppm. This corresponded with the results for the unlabeled MDA test, with the ^{15}N-MDA byproducts having masses of 244.13 and 286.14, indicating an additional 3 mass value due to the two ^{15}N and single ^{13}C in the labeled-MDA compound. Thus, it can be concluded that *Pseudomonas aeruginosa* converted MDA into alternative molecules as the amines terminating the molecule are converted into an amide form through acetylation with correlating changes observed in both LCMS and NMR.

3.3.5. Environmental consortia uptake of DMDA

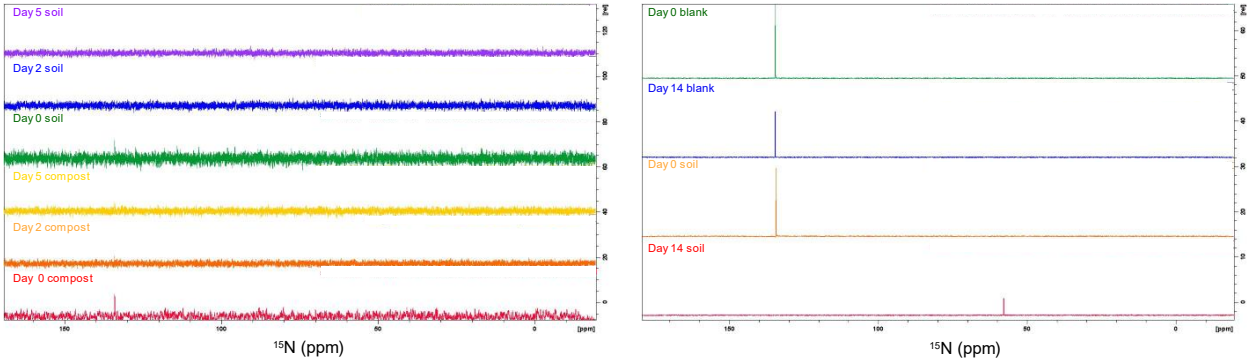


Figure 3.4: ^{15}N INEPT NMR studies on DMDA a) 5 day compost and soil consortia study b) 14 day soil consortia study

After determining the modifications of MDA to MMDA and DMDA by Pseudomonas aeruginosa, next studies investigated how DMDA was molecularly modified by the previously identified consortia of organisms. In an experiment with 3 mg/L ^{15}N-DMDA and a consortia of soil and compost organisms, a slight peak disappearance was able to be seen using ^{15}N NMR (Figure 3.4). At the 0-hour time point, one peak was observed, the expected acetylated amine peak for ^{15}N-DMDA at 130 ppm as confirmed from the standard signal (Figure 3.4). Two days later in both soil and compost experiments the small DMDA peak was no longer able to be visualized at

130 ppm, however no new peaks appeared. This lead to the hypothesis that the small molecule must be insolubilized and is not getting altered by the organism communities. In order to better visualize what was taking place, a 10 mg/L ^{15}N-DMDA experiment was developed to monitor how a soil consortia of organisms altered the small molecule over an extended two week period. After 2 weeks, we visualized the disappearance of the 130 ppm peak corresponding to the starting amide and a shift of all detectable nitrogen in the sample to a peak at 55 ppm, however this peak is significantly lower in signal than the original peak. This decrease in peak intensity correlates with the five day consortia studies in which the peaks disappeared over time. The visualized peak shift signifies that the nitrogen within the sample has been converted back into an amine form, showing the consortia's use of acetylation and deacetylation over long term. Combined with the decrease in signal over time we determined the organisms in the consortia must be able to convert between both forms readily depending on the needs for uptake and use by each organism.

3.4 Conclusion

MDA is the primary hypothesized toxin that can be released from biodegradable polyurethanes currently made, leading to questions on its toxicity within the environment. In order to begin addressing questions regarding whether it is released from these biodegradable PUs, this work set out to determine if it gets metabolized or if it is toxic to a microbial community. Co-metabolism studies on MDA along with PU monomers were the first studies conducted to investigate the effects of MDA on microbiota over time, determining its toxicology. These passaging studies determined that the consortia continued to grow over time with cell numbers doubling by the end of an 8 week study. Using metagenomic sequencing technology it was determined that three main organisms were able to survive and grow best in the presence of this small molecule, *Rhodococcus*, *Pseudomonas*, and *Ochrobactrum*. Follow up studies using an

activity assay on a community of bacterial organisms determined what organisms contributed to the intake of each monomer used within the formulation, with *Pseudomonas*, *Ochrobactrum*, or *Rhodococcus* coming in as privileged strains which could survive best in the presence of MDA. In order to analyze the modification of MDA, a ^{13}C,^{15}N-labeled MDA's alteration was monitored by MS and NMR after incubation with organisms of relevance. This led to the determination that there were microorganisms within the consortia which were able to either acetylate MDA or deacetylate DMDA with *Pseudomonas* being a confirmed bacteria to acetylate the small molecule. This study has led to the conclusion that there must be at least one of the three bacterial genus' *Pseudomonas*, *Ochrobactrum*, or *Rhodococcus* present in order to survive and grow in the presence of MDA. These three bacteria hold the most potential for physiochemically removing MDA from the available metabolite pool by acetylating it for insolubility and giving time for other breakdown methods, such as aromatic breakdown by UV light to take place.

3.5 Acknowledgements

Chapter 3, in full, is currently being prepared for submission for publication of the material as Kalaj, Brianna N.; Tessman, Marissa; Simkovsky, Ryan; Mayfield, Stephen P.; Burkart, Michael D. "Investigating methylene diphenyl diamine alteration and uptake by soil and compost organisms." The dissertation author was the primary researcher and author of this material.

3.7 References

1. R. Geyer, J. R. Jambeck, K. L. Law, Production, use, and fate of all plastics ever made. *Sci. Adv.*, **2017**, *3*, e1700782

2. S. Anbumani, P. Kakkar, Ecotoxicologial effects of microplastics on biota: a review. *Environ. Sci. Pollut. R.*, **2018**, *25*, 14373-14396.

3. N. R. Gunawan, M. Tessman, A. C. Schreiman, R. Simkovsky, A. A. Samoylov, N. K. Neelakantan, T. A. Bemis, M. D. Burkart, R. S. Pomeroy, S. P. Mayfield, Rapid Biodegradation of renewable polyurethane foams with identification of associated microorganisms and decomposition products. *Bioresour. Technol. Rep.* **2020**, *11*, 100513.

4. T. A. Phung Hai, N. K. Neelakantan, M. Tessman, S. D. Sherman, G. Griffin, R. S. Pomeroy, S. P. Mayfield, M. D. Burkart, Flexible polyurethanes, renewables fuels, and flavorings from a microalgae oil waste stream, *Green Chem*, **2020**, *10*, 3088-3094.

5. A. Magnin, E. Pollet, V. Phalip, L. Averous, Evaluation of biological degradation of polyurethanes. *Biotechnol. Adv.* **2019**, *39*, e107457.

6. G. T. Howard, Biodegradation of polyurethane: a review. *Int. Biodeterior. Biodegradation.* **2002**, *49*, 245-252.

7. M. Osman, S. M. Satti, A. Luqman, F. Hasan, Z. Shah, A. A. Shah, Degradation of Polyester Polyurethane by *Aspergillus* sp. Strain S45 Isolated from Soil. *J. Polym. Environ.* **2017**, *26*, 301-310.

8. Y. Matsumiya, N. Murata, E. Tanabe, K. Kubota, M. Kubo. Isolation and characterization of an ether-type polyurethane-degrading micro-organism and analysis of degradation mechanism by *Alternaria* sp. *J. Appl. Microbiol.* **2010**, *108*, 1946-1953.

9. P. Bhavsar, M. Bhave, H. K. Webb, Effective multi-stage biodegradation of commercial bulk polyurethane by Clonostachys and Purpureocillium spp. *Sci. Total Environ.* **2024**, *905*, 168329.

10. D. C. Allport, D. S. Gilbert, S. M. Outterside, *MDI and TDI: Safety, Health and the Environment: A Source Book and Practical Guide*. John Wiley & Sons, 2003

11. G. T. Howard, Polyurethane Biodegradation, In: Singh, S. (eds) *Microbial Degradation of Xenobiotics. Environmental Science and Engineering*, Springer, Berlin, Heidelberg, 2012

12. M. Cregut, M. Bedas, M. J. Durand, G. Thouand, New insights into polyurethane biodegradation and realistic prospects for the development of a sustainable waste recycling process. *Biotechnol. Adv.* **2013**, *31*, 1634-1647.

13. J. Liu, Q. Zeng, H. Lei, K. Xin, A. Xu, R. Wei, D. Li, J. Zhou, W. Dong, M. Jiang, Biodegradation of polyester polyurethane by *Cladosporium* sp. P7: Evaluating its degradation capacity and metabolic pathways. *J. Hazard. Mater.* **2023**, *448*, 130776.

14. R. P. Singh, N. S. Tomer, S. V. Bhadraiah, Photo-oxidation studies on polyurethane coating:effect of additives on yellowing of polyurethane. *Polym. Degrad. Stab.*, **2001**, *73*, 443-446.

15. D. Rosu, L. Rosu, C. N. Cascaval, IR-change and yellowing of polyurethane as a result of UV irradiation. *Polym. Degrad. Stab.*, **2009**, *94*, 591-596.

16. H. Inoue, T. Tachibana, T. Bito, J. Arima, Acetylation of amines and alcohols catalyzed by acetylcholinesterase from *Pseudomonas aeruginosa* PAO1. *Enzyme Microb. Technol.* **2023,** *165*, e110208.

17. A. Zadlo-Dobrowolska, N. G. Schmidt, W. Kroutil, Promiscuous activity of C-acyltransferase from *Pseudomonas protegens*: synthesis of acetanilides in aqueous buffer, *Chem. Commun.* **2018**, *54*, 3387.

Chapter 4 Quantitative characterization of chain-flipping of acyl carrier protein of *Escherichia coli* using chemical exchange NMR

4.1 Introduction

Acyl carrier proteins (ACPs) are highly conserved and ubiquitously distributed proteins that facilitate the biosynthesis of fatty acids by sequestering and shuttling acyl intermediates to their various enzymatic binding partners (1-5). In eukaryotes, ACPs are expressed as subdomains of large multifunctional fatty acid megasynthases (type I FAS). In contrast, in most bacteria and plants, ACPs exist as small monomeric proteins that interact with discrete, monofunctional enzymes (type II FAS). Despite these differences, the chemical aspects of fatty acid synthesis are remarkably preserved across all kingdoms of life and include an iterative series of enzymatic reactions involving an initial priming step, followed by elongation of the acyl chain, and its eventual termination (Fig. 4.1A). First, the conserved serine residues of functionally inactive forms of ACPs (*apo*-ACPs) are linked to coenzyme A derived 4′-phosphopantetheine (Ppant) prosthetic moieties to create *holo*-ACPs. In the priming step, acetyl moieties are conjugated to the terminal thiol groups of the Ppant arms of *holo*-ACPs *via* thioester bonds. These ACP-tethered acetyl groups are then elongated and modified *via* a series of decarboxylative Claisen condensation and reduction reaction steps (an addition of two-carbon units per cycle) to generate acyl intermediates that are hydrolyzed upon reaching the desired chain lengths to produce fatty acids.

ACPs have thus evolved to perform two seemingly paradoxical functions: a) to compartmentalize growing fatty acid chains to prevent their premature termination and unwanted side reactions, and b) to deliver these acyl cargos to the active sites of appropriate enzymes for the corresponding native chemical reactions. The molecular basis of ACP-mediated sequestration of acyl chains has been ascertained from numerous structural studies, including X-ray crystallography and solution nuclear magnetic resonance (NMR) spectroscopy [reviewed in Ref.

(1, 5)]. These studies revealed that many ACPs adopt a helical bundle architecture with an expandable hydrophobic pocket that can accommodate the growing acyl chain. Similarly, the translocation of ACP-sequestered acyl chains to the active sites of enzymes, a process termed "chain-flipping", was elucidated using X-ray structures of complexes between an ACP of *Escherichia coli* (*E. coli*) called AcpP and its enzymatic partners [summarized in Ref. (6)]. In these studies, the thioester of AcpP was replaced with stable reactive groups to stabilize transient AcpP-enzyme complexes. The resulting cross-linked structures revealed that AcpP-tethered acyl chains and Ppant arms move into the active sites of enzymes, thus allowing enzymes to gain access to the acyl chains for the subsequent condensation/reduction reactions. Despite these advances, the biophysical principles that govern the interconversions between ACP-sequestered and enzyme-sequestered conformations of the acyl chains are unclear. One likely possibility is that ACP-tethered acyl chains can spontaneously sample ACP-sequestered and solvent-exposed conformations, with the latter being sparsely populated states, and that the ACP-interacting enzymes preferentially bind to these "excited" states, thereby altering the underlying equilibrium. The direct experimental evidence of this proposed equilibrium between ACP-sequestered and lowly-populated solvent-exposed conformations of ACP-tethered acyl chains is, however, lacking. This is because although biological processes, including macromolecular recognition and enzyme catalysis, often rely on lowly-populated states, their very nature makes these states invisible to conventional structural methods that probe homogenous, highly populated states of systems of interest (7, 8).

Here, we make use of *E. coli* AcpP and a synergistic combination of chemical biology and NMR spectroscopy to demonstrate that acyl chains with six, eight, and ten carbons spontaneously flip in and out of the hydrophobic pocket of AcpP. Using chemical exchange saturation transfer

(CEST) and Carr−Purcell−Meiboom−Gill (CPMG) relaxation dispersion experiments that are exquisitely-suited to characterize lowly-populated states (9, 10), we determine the kinetic rate constants of these chain-flips and show that each successive two-carbon unit increase in the acyl chains is accompanied by a progressive decrease in the populations of their solvent-exposed conformations.

4.2 Results and Discussion

Synthesis of selectively ^{15}N-labeled Ppant mimics. The replacement of labile sulfur atoms in Ppant moieties with nitrogen and the subsequent incorporation of these non-hydrolysable pantetheinamides into recombinant ACPs using chemoenzymatic methods have been used to perform numerous structure-function studies of ACPs [reviewed in Ref. (11)]. However, to facilitate chemical-exchange-based NMR studies of *E. coli* AcpP, we first set out to synthesize Ppant analogs with varying acyl chain lengths in which the thioester linkages were replaced by ^{15}N-labeled amide bonds. This strategy was devised because, in addition to generating stable non-hydrolysable acyl-AcpPs conducive for the above-mentioned NMR methods that often require long acquisition times (ranging from hours to days), the selectively ^{15}N-labeled pantetheinamides provided direct experimental evidence for the spontaneous interconversion between AcpP-sequestered and solvent-exposed conformations of AcpP-tethered acyl chains (described below). Note that, using molecular dynamics simulations, we have previously demonstrated that the replacement of sulfur for nitrogen produces minimal perturbations in AcpP's hydrophobic pocket, and thus, an amide linkage is a suitable substitute for Ppant's thioester (12).

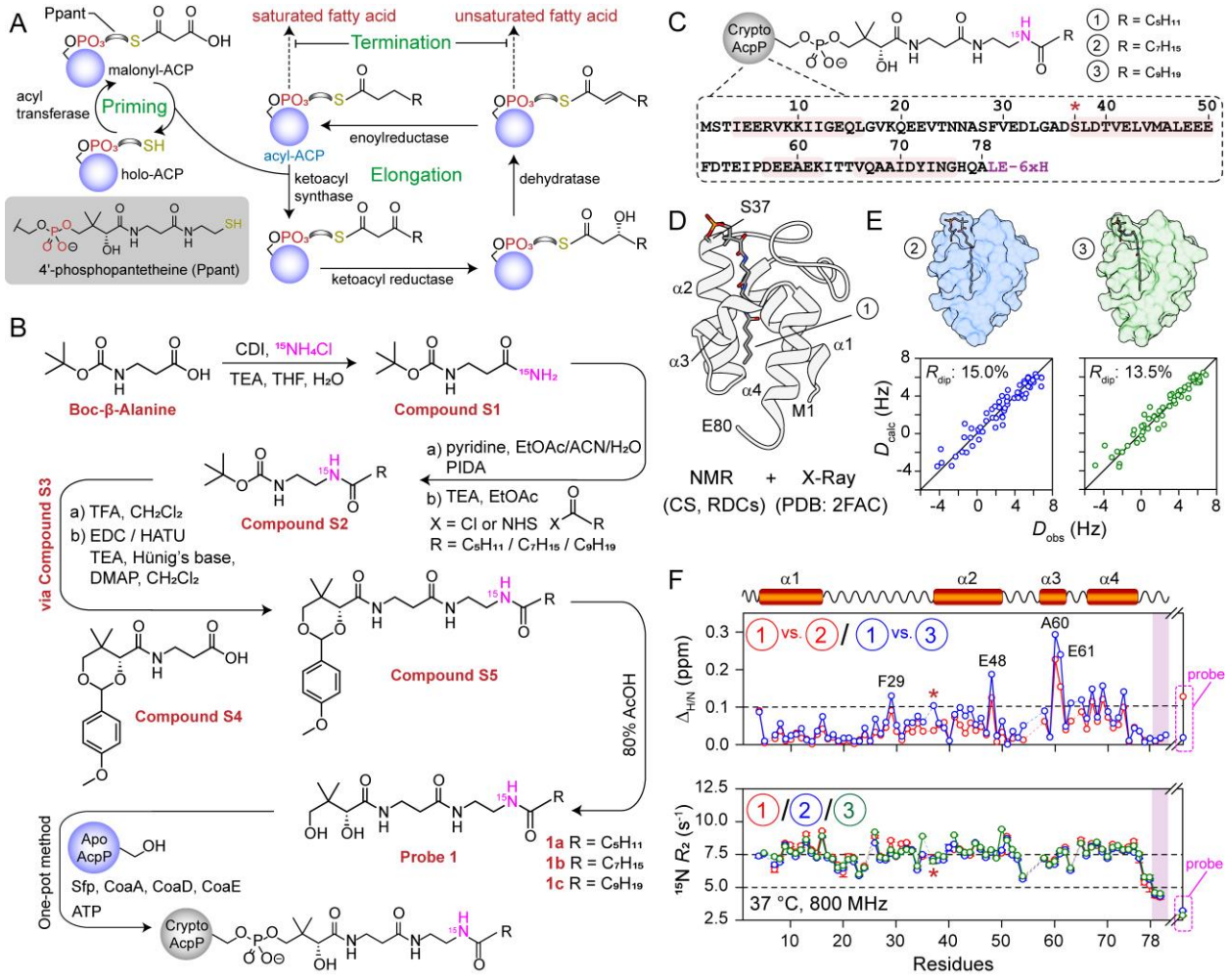


Figure 4.1: (**A**) Schematic overview of general fatty acid biosynthesis; major steps are labeled in green. R = $(CH_2\text{-}CH_2)_n\text{-}CH_3$. (**B**) Synthesis scheme of selectively ^{15}N-labeled pantetheinamides and their conjugation to *apo*-AcpP using the one-pot method (16). (**C**) Schematic and primary sequence of *crypto*-AcpP used in current study. Non-native residues and polyhistidine tag are in magenta, and the ^{15}N atom of tethered pantetheinamide is in pink. The locations of helices are marked by semi-transparent red rectangles; asterisk marks residue S37. (**D**) Ribbon diagram of the model of C_6-*crypto*-AcpP calculated from backbone chemical shifts and $^1D_{NH}$ RDCs using CS Rosetta (17) and Xplor-NIH (18). The C_6-pantetheinamide and residue S37 are in stick representation; throughout the figure, the circled numbers signify acyl chain lengths shown in (**C**). (**E**) Upper: semi-transparent surface representation of models of C_8/C_{10}-*crypto*-AcpPs; tethered pantetheinamides and residue S37 are in stick representation. Lower: best-fit agreement obtained by SVD between $^1D_{NH}$ RDCs measured in pf1 for C_8/C_{10}-*crypto*-AcpPs and those calculated from their corresponding models (blue and green, respectively). The RDC *R*-factor, R_{dip}, is given by $\{<(D_{obs}\text{-}D_{calc})^2>/(2<D_{obs}^2>)\}^{1/2}$, where D_{obs} and D_{calc} are the observed and calculated $^1D_{NH}$ RDC values, respectively (41). (**F**) Perturbations in $^1H_N/^{15}N$ chemical shifts (upper) and ^{15}N-relaxation analysis (lower) of crypto-AcpP samples; tethered pantetheinamides are highlighted in pink dashed rectangles. Secondary structure elements are indicated above the panel. Non-native residues are highlighted in semi-transparent magenta rectangles. Dashed lines indicate proline residues or residues that could not be assigned unambiguously.

To synthesize selectively ^{15}N-labeled pantetheinamides, we devised a general scheme shown in Fig. 4.1B; also see SI appendix, Schemes S4.1-S4.3. Briefly, ^{15}N-labeled ammonium chloride was coupled to Boc-β-alanine to generate an ^{15}N-labeled carbamate derivative (referred to as Compound **S1**). Compound **S1** was subjected to a Hofmann rearrangement (13) in the presence of phenyliodine (III) diacetate (PIDA) to produce the corresponding amine. The latter was capped with appropriate acyl chains (six/eight/ten carbon atoms) *in situ* by reaction with either an acid chloride or an N-hydroxysuccinimide (NHS) ester to yield Compound **S2**. Compound **S2** was subjected to a three-step process involving the removal of the Boc-protecting group [resulting in **S3**, followed by coupling to p-methoxybenzyl (PMB) protected pantetheine **S4** (14), and the removal of the PMB group to generate ^{15}N-labeled pantetheinamides **1a-1c** comprising acyl chains of six/eight/ten carbons linked to an ^{15}N atom (experimental details found in the Supporting Information). The chain lengths were selected because the hydrophobic cavity of AcpP can adequately accommodate acyl chains of six to ten carbons, resulting in progressive stabilization of its structure as a function of increasing acyl chain lengths (4, 15). The ^{15}N-labeled pantetheinamides **1a–1c** were individually tethered to recombinant *apo*-AcpP using our one-pot chemoenzymatic method (16); see SI appendix, Fig. S4.1. The resulting acyl-AcpPs carrying these non-native cargos are hereafter referred to as C_6/C_8/C_{10}-*crypto*-AcpPs (Fig. 4.1C).

NMR analysis of *crypto*-AcpPs. To explore the impact of tethered pantetheinamides on the structure and conformational dynamics of *crypto*-AcpPs, we conducted a detailed NMR investigation. Uniformly ^{15}N-labeled C_6/C_8/C_{10}-*crypto*-AcpPs carrying selectively ^{15}N-labeled pantetheinamides **1a**, **1b**, and **1c**, respectively, yielded excellent ^{1}H-^{15}N HSQC spectra (SI appendix, Fig. S4.2), indicating well-folded stable conformations. The structure of C_6-*crypto*-AcpP was determined from backbone amide residual dipolar couplings ($^{1}D_{NH}$ RDCs) and

backbone chemical shifts using CS-Rosetta (17) and revealed a canonical AcpP fold comprising a four-helix bundle (Fig. 4.1D). The AcpP-sequestered conformation of the tethered C_6-pantetheinamide was modeled using Xplor-NIH (18). Briefly, the C_6-pantetheinamide was first conjugated to residue S37 of the above CS-Rosetta structure, and the resultant complex was refined against an X-ray structure of AcpP carrying a sequestered C_6-pantetheine (15). Note that the numbering of *crypto*-AcpP is based on the UniProt entry: P0A6A8, with the initiator methionine as the first residue (cf. Fig. 4.1C). The C_6-*crypto*-AcpP structure was used as a template to model the AcpP-sequestered conformations of the tethered C_8- and C_{10}-pantetheinamide probes. The resultant structures showed that longer acyl chains adopt linear orientations and were buried deep in the hydrophobic cavity of *crypto*-AcpP (SI appendix, Fig. S4.3). These structures were cross validated against $^1D_{NH}$ RDCs measured for samples of C_8/C_{10}-*crypto*-AcpPs (Fig. 4.1E). The corresponding singular value decomposition (SVD) fits yielded excellent *R*-factors (ranging between ~14–15%), confirming that the pantetheinamides were predominantly sequestered inside the hydrophobic cavity of crypto-AcpP. If these AcpP-tethered pantetheinamides sampled solvent-exposed conformations, based on *R*-factors, the upper limit for their populations is <10% (19-21).

Further NMR analysis of C_6/C_8/C_{10}-*crypto*-AcpP samples revealed large $^1H_N/^{15}N$ chemical shift perturbations in a few AcpP residues upon increasing the acyl chain lengths ($\Delta_{H/N} \geq 0.1$ ppm; Fig. 4.1F). The ^1H-^{15}N cross-peaks of the ^{15}N atoms of the pantetheinamides could be readily identified in the ^1H-^{15}N HSQC spectra of C_6/C_8/C_{10}-*crypto*-AcpP samples (cf. SI appendix, Fig. S4.2). C_6- and C_{10}-pantetheinamides exhibited similar $^1H_N/^{15}N$ chemical shifts ($\Delta_{H/N}$ ~0.02 ppm).

In contrast, $^1H_N/^{15}N$ chemical shift perturbations between C6- and C8-pantetheinamides were ~0.13 ppm. The large $^1H_N/^{15}N$ perturbations in a few AcpP residues and the probe ^{15}N atoms

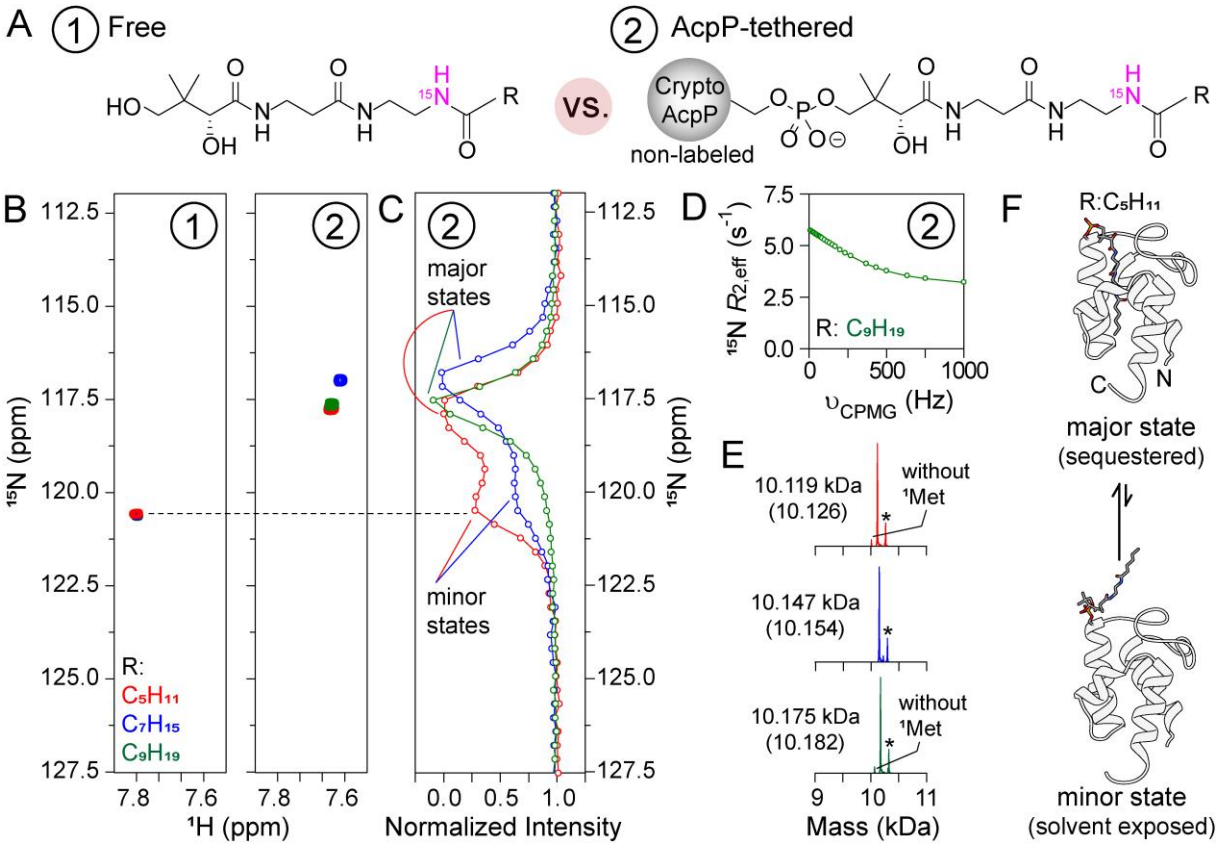


Figure 4.2: Overlay of ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled *crypto*-AcpPs (A) C6 (B) C8 and (C) C10 and ^{15}N labeled free pantetheinamides.

likely arise from side-chain repacking, as the overall conformational changes in $C_6/C_8/C_{10}$-*crypto*-AcpP structures are negligible, evidenced by the above RDC data. Analysis of ^{15}N-relaxation rates, measured at 800 MHz at 37 °C, revealed average R_2 values of ~7.5 s^{-1} for majority of AcpP residues, further confirming the globular nature of $C_6/C_8/C_{10}$-*crypto*-AcpPs (Fig. 4.1F). The corresponding R_2 values for ^{15}N atoms of AcpP-tethered pantetheinamides were ~3 s^{-1}, indicating that, unlike AcpP, the probes exhibit rapid backbone dynamics. The latter is consistent with prior NMR structural studies of AcpP that showed dynamic disorder in AcpP-tethered Ppant moieties, resulting in weak or no interactions between protein backbone and Ppant atoms (22, 23). Taken

together, the above results show that *crypto*-AcpPs studied here adopt an all-helical architecture with tethered pantetheinamides predominantly sampling AcpP-sequestered conformations.

Direct evidence of solvent-exposed states of AcpP-tethered pantetheinamides. To determine whether AcpP-tethered pantetheinamides transiently sample solvent-exposed conformations, we compared **1a–1c** in their free forms to non-labeled AcpPs (Fig. 4.2A). First, probes **1a–1c** were dissolved in a buffer in the absence of AcpP. The corresponding ^{1}H-^{15}N HSQC spectra revealed that the ^{15}N-atoms of solvent-exposed pantetheinamides were superimposed and all ^{15}N resonances were downfield-shifted (^{15}N shift of ~120.6 ppm; Fig. 4.2B). In contrast, the ^{15}N resonances of AcpP-tethered pantetheinamides exhibited notable upfield shifts (ranging from 117–118 ppm). Based on these results and the above-described structural and RDC analyses, we conclude that the latter represent AcpP-sequestered states of the pantetheinamides. To uncover the existence of solvent-exposed (i.e., NMR-invisible) states of AcpP-tethered pantetheinamides, ^{15}N-CEST methodology was employed, which "visualizes" sparsely populated states (≥0.5%) that are in exchange with NMR-visible major species on a time scale of 2–50 ms (9, 24, 25). The corresponding ^{15}N-CEST profiles of AcpP-tethered C_6- and C_8-pantetheinamides showed well-resolved minor dips in intensities at resonance frequencies of ~120.5 ppm, which coincided with ^{15}N-frequencies of solvent-exposed probes (Fig. 4.2C). The magnitude of these minor dips, which are proportional to the populations of the corresponding minor states, was larger in C_6-*crypto*-AcpP as compared to C_8-*crypto*-AcpP, whereas no discernible minor CEST dips were observed in C_{10}-*crypto*-AcpP sample. The latter may result from the timescale of exchange not being accessible to ^{15}N-CEST (τ_{ex} <2 or >50 ms) and/or because the population of the solvent-exposed conformation of AcpP-tethered C_{10}-pantetheinamides is below the limit of detection of ^{15}N-CEST (≤0.5%). To determine the evidence or lack thereof of chain flips exhibited by C_{10}-*crypto*-AcpP,

we measured ^{15}N-CPMG relaxation dispersions, which probe exchange between NMR-visible major and NMR-invisible minor states on a time scale of 0.1–10 ms (10, 24, 26). The ^{15}N-atom of AcpP-tethered C_{10}-pantetheinamide yielded small but noticeable dispersions (R_{ex} ~2.5 s^{-1}; Fig. 4.2D), indicating that like C_6- and C_8-*crypto*-AcpP samples, C_{10}-*crypto*-AcpP undergoes conformational exchange, albeit on a faster timescale. Further analysis of these samples by liquid chromatography–electrospray ionization–time-of-flight mass spectrometry (LC-ESI-TOFMS) revealed complete conjugation of selectively ^{15}N-labeled pantetheinamides to non-labeled AcpPs (Fig. 4.2E). Collectively, the above results establish that the AcpP-tethered pantetheinamides transiently flip out of the hydrophobic cavity of AcpP (Fig. 4.2F). Additionally, we note that performing these analyses on non-labeled proteins tethered to selectively ^{15}N-labeled pantetheinamides (as opposed to uniformly ^{15}N-labeled proteins) permitted us to rapidly ascertain the existence of underlying conformational exchange due to the simplicity of the corresponding spectra (1 peak per spectrum, ^{15}N-CEST/^{15}N-CPMG acquisition time: ≤3 h/sample).

Quantitative characterization of chain flipping exhibited by AcpP-tethered pantetheinamides. To characterize the equilibrium between AcpP-sequestered and solvent-exposed conformations of AcpP-tethered pantetheinamides, we recorded ^{15}N-CEST experiments on samples comprising uniformly ^{15}N-labeled AcpPs tethered to selectively ^{15}N-labeled pantetheinamides (Fig. 4.3A). Here, uniformly ^{15}N-labeled (as opposed to non-labeled) AcpP was used. The rational for this approach was to identify protein residues that were affected by the conformational fluctuations of the tethered pantetheinamides and to reliably determine the underlying rate constants (described below). For C_6- and C_8-*crypto*-AcpP samples, many protein residues (26 and 18, respectively) along with ^{15}N-atoms of tethered pantetheinamides exhibited well-resolved second CEST dips at resonance frequencies of minor states (Fig. 4.3B; also see SI

appendix, Fig. S4.4–S4.5 and Table S4.1). Like our above-described results, the magnitude of CEST dips representing minor states became progressively smaller as a function of increasing acyl chain lengths, and no characteristic second CEST dips were observed for the C_{10}-*crypto*-AcpP sample (cf. Fig. 4.2C and Fig. 4.3B). The latter was subjected to ^{15}N-CPMG relaxation dispersion experiments, which showed noticeable dispersions for 11 protein residues along with the ^{15}N-atom of the tethered C_{10}-pantetheinamide (Fig. 4.3C; also see SI appendix, Fig. S4.6). These results confirm our earlier observations (see Fig. 4.2D) and show that the C_{10}-*crypto*-AcpP sample also undergoes conformational exchange.

Fig. 4.3D highlights the residues of C_6-*crypto*-AcpP that showed large CEST effects. These residues primarily lie around the hydrophobic cavity of AcpP, encompassing the sequestered C_6-pantetheinamide. Similar residues exhibited ^{15}N-CEST or ^{15}N-CPMG effects for C_8 and C_{10}-*crypto*-AcpP samples, respectively (SI appendix, Fig. S4.5–S4.6). These observations further support transient sampling of the solvent-exposed conformations by the tethered pantetheinamides, which will alter the local environment of AcpP's hydrophobic pocket. Additionally, a vast majority of residues of C_6/C_8-*crypto*-AcpPs exhibited notable downfield ^{15}N shifts for their corresponding minor CEST dips (SI appendix, Fig. S4.4–S4.5). The downfield change in the isotropic chemical shifts of backbone ^{15}N atoms has long been known to correlate well with increased hydrogen-bonding strengths (27-29). We, therefore, hypothesize that solvent-exposed configurations of tethered pantetheinamides will rearrange AcpP's hydrophobic cavity, building an intramolecular hydrogen bond network between its residues and consequently compacting its hydrophobic lumen. These inferences agree with a prior study that established a gradual decrease in AcpP's hydrophobic cavity as a function of decreasing the acyl chain lengths of tethered pantetheinamides (15).

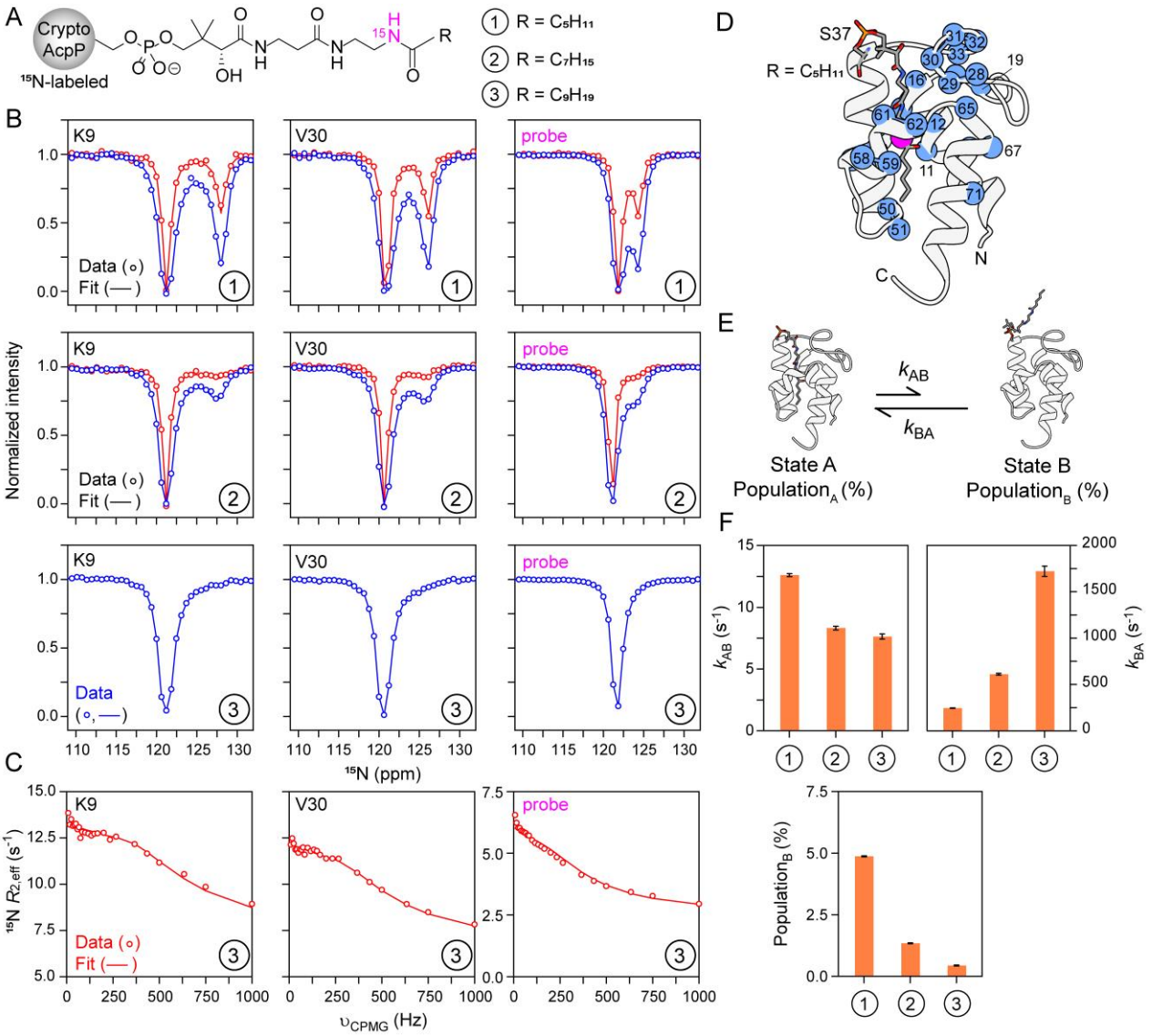


Figure 4.3: (**A**) Schematic representation of samples used for ^{15}N-CEST or ^{15}N-CPMG experiments (protein concentration: 1 mM); acyl chain lengths of AcpP-tethered pantetheinamides are designated by circled numbers. (**B**) Examples of ^{15}N-CEST profiles observed for samples shown in panel (**A**) obtained with the constant wave (CW) saturation field applied for T_{CEST} = 400 ms at two RF field strengths: 14.4 (red) and 28.6 (blue) Hz; calculated values: 15 and 30 Hz, respectively. For C_6- and C_8-*crypto*-AcpP samples (upper and middle, respectively), the experimental CEST data are displayed as circles, and the global best-fits to a two-state exchange model (see panel **E**) as solid lines. For C_{10}-*crypto*-AcpP (lower), no characteristic CEST effect was observed, and solid lines are, therefore, used to guide the eye. (**C**) Examples of ^{15}N-CPMG relaxation dispersion profiles obtained for C_{10}-*crypto*-AcpP. (**D**) Residues of C_6-*crypto*-AcpP (blue spheres) and the ^{15}N atom (pink sphere) of the tethered pantetheinamides that were used for global fitting of the ^{15}N-CEST data. The kinetic model (**E**) and rate constants (**F**) for exchange between sequestered (State A) and solvent-exposed configuration (State B) of AcpP-tethered pantetheinamides. Populations of State **B** derived from ^{15}N-CEST/-CPMG data are shown at the bottom of panel (**F**).

Quantitative analyses of CEST or CPMG profiles via propagation of the Bloch-McConnell equations (30) yield populations, chemical shifts of the minor state(s), and the rate constants of the underlying equilibrium (9, 24-26). We, therefore, analyzed the conformational exchange exhibited by AcpP-tethered pantetheinamides using the kinetic scheme shown in Fig. 4.3E. This scheme involves exchange between AcpP-sequestered (State **A**) and lowly-populated solvent-exposed (State **B**) configurations of the tethered pantetheinamides. Global fitting of the ^{15}N-CEST profiles of C_6-*crypto*-AcpP (26 AcpP residues and ^{15}N atom of the C_6-pantetheinamide, acquired at two RF fields; SI appendix, Fig. S4.4) to this two-state model yielded excellent fits, establishing that the observed exchange process indeed involves two-states and the transient sampling of the solvent-exposed conformation by the tethered pantetheinamide. The latter conformation was populated at ~4.9%, and the forward and backward exchange rate constants, k_{AB} and k_{BA}, were ~13 and ~246 s^{-1}, respectively (Fig. 4.3F and Table 4.1). Global fitting of the ^{15}N-CEST profiles of C_8-*crypto*-AcpP (18 AcpP residues and ^{15}N atom of the C_8-pantetheinamide; SI appendix, Fig. S4.5) yielded the exchange rates of k_{AB}: ~8 s^{-1} and k_{BA}: ~611 s^{-1} and the solvent-exposed state population of ~1.3%. In the case of C_{10}-*crypto*-AcpP, a model with exchange rates of k_{AB}: ~8 s^{-1} and k_{BA}: ~1722 s^{-1} and a solvent-exposed state population of ~0.4% was able to adequately reproduce experimental ^{15}N-CPMG dispersions (11 AcpP residues and ^{15}N atom of the C_{10}-pantetheinamide; SI appendix, Fig. S4.6). The ^{15}N chemical shift differences (^{15}N-$\Delta\omega$) between major and minor states of C_{10}-*crypto*-AcpP derived from fits to ^{15}N-CPMG data matched closely with those obtained from the experimental ^{15}N-CEST profiles of C_6/C_8-*crypto*-AcpP (SI appendix, Fig. S4.7). In addition to cross-validating the goodness of the CPMG fit to the above-described two-state model, the close correlation between ^{15}N-$\Delta\omega$ values indicates that the local environment of AcpP's hydrophobic cavity is essentially the same for the lowly-populated states of $C_6/C_8/C_{10}$-*crypto*-

AcpP samples, further confirming that the tethered pantetheinamides flip out of AcpP's hydrophobic pocket in these transient states. The trends observed in exchange rate constants and minor state populations derived from CEST/CPMG profiles (Fig. 4.3F and Table 4.1) are in excellent agreement with our above-described experimental observations (see Figs. 4.2C-D and 4.3B-C). An increase in acyl chain length is accompanied by a progressive decrease in the populations of solvent-exposed configurations of AcpP-tethered pantetheinamides (~4.9, ~1.3, and ~0.4% for $C_6/C_8/C_{10}$, respectively), largely because of the substantial increase in corresponding backward rate constants (k_{BA}: ~246, ~611, ~1722 s^{-1} for $C_6/C_8/C_{10}$, respectively). The forward rate constants for the formation of solvent-exposed states were not significantly different (k_{AB}: ~8–13 s^{-1}), implying similar underlying processes. The lifetimes of solvent-exposed states were ~3.9, ~1.6, and ~0.6 ms for C_6, C_8, and C_{10} acyl chains, respectively [lifetime = 1 / (k_{AB} + k_{BA})]. Collectively, these results explain why no CEST effect was observed for the C_{10}-*crypto*-AcpP sample, as both the population (~0.4%) and the lifetime (~0.6 ms) of the solvent-exposed state are beyond the limit of detection of the CEST experiment.

Table 4.1: Populations (P_A and P_B), exchange (k_{AB} and k_{BA}), and equilibrium (K_{eq}) rate constants for the exchange between AcpP-sequestered (State **A**) and solvent-exposed (State **B**) conformations of AcpP-tethered pantetheinamides.[a]

	P_A (%)	P_B (%)	k_{AB} (s^{-1})	k_{BA} (s^{-1})	K_{eq}[b]
C_6-*crypto*-AcpP[c]					

95.13 ± 0.02	4.87 ± 0.02	12.60 ± 0.12	246.12 ± 2.43	19.53 ± 0.08
C_8-*crypto*-AcpP[(c)]				
98.66 ± 0.02	1.34 ± 0.02	8.32 ± 0.15	611.07 ± 10.69	73.63 ± 1.10
C_{10}-*crypto*-AcpP[(d)]				
99.56 ± 0.02	0.44 ± 0.02	7.63 ± 0.20	1722.13 ± 54.79	226.27 ± 10.29

a. NMR experiments were performed at spectrometer ^{1}H frequency of 800 MHz at 37 °C. The buffer conditions were as follows: 50 mM potassium phosphate, pH 7.4, and 0.01% sodium azide. Data were collected on uniformly ^{15}N-labeled AcpP tethered to selectively ^{15}N-labeled pantetheinamides (protein concentration = 1 mM).
b. The K_{eq} values are given by $K_{eq} = P_A / P_B$ and have been rounded to two decimal places.
c. For ^{15}N-CEST datasets, the set of global fitted parameters were $\{k_{AB}; P_B\}$. The corresponding population of State **A** (P_A) and backward rate constat (k_{BA}) were calculated using the relationships: $P_A = 100 - P_B$; and $k_{BA} = P_A \times k_{AB} / P_B$.
d. For ^{15}N-CPMGs, the set of global fitted parameters were $\{k_{AB}; k_{BA}\}$. The corresponding populations of State **B** (P_B) and State **A** (P_A) were calculated using the relationships: $P_B = (100 \times k_{AB}) / (k_{AB} + k_{BA})$; and $P_A = 100 - P_B$.

4.3 Conclusion

In summary, we report a chemical synthesis of selectively ^{15}N-labeled $C_6/C_8/C_{10}$-pantetheinamides, which were conjugated to *E. coli* AcpP to generate *crypto*-AcpP samples conducive for heteronuclear NMR studies. NMR structural analysis showed that these pantetheinamides predominantly adopt an AcpP-sequestered conformation. Chemical-exchange NMR methods, ^{15}N-CEST and ^{15}N-CPMG, provided direct experimental evidence of lowly-populated solvent-exposed states of AcpP-tethered pantetheinamides and facilitated a thorough kinetic analysis of their interconversion to AcpP-sequestered states. Populations of solvent-exposed conformation of acyl chains were inversely proportional to the chain lengths and varied from ~5–0.4% for chains with six, eight, and ten carbons. The rate constant (k_{BA}) for the AcpP-sequestered conformation of C_{10}-pantetheinamide was about three and seven times higher than

those for the C_8- and C_6-pantetheinamides, respectively, indicating that AcpP-sequestered state of C_{10}-pantetheinamide is energetically more favorable than that of its C_8- and C_6-counterparts. The unitless equilibrium constant, K_{eq}, for the interconversion of solvent-exposed to AcpP-sequestered states of the tethered pantetheinamide is defined as $K_{eq} = [P_A]/[P_B]$, where P_A and P_B are the populations of AcpP-sequestered and solvent-exposed states, respectively (Table 4.1). Under the conditions employed here (pH 7.4 and 37 °C), K_{eq} values were ~20, ~74, and ~226 for $C_6/C_8/C_{10}$-pantetheinamides, respectively, further confirming that there is an energetic penalty for exposing the aliphatic acyl chains to the solvent and that this energy barrier is directly proportional to the chain lengths investigated in this study.

The simplest explanation for the above-observed phenomenon is that the growing acyl chain builds a dense network of complementary interactions within AcpP's tunnel-like hydrophobic core, thus making it difficult to expose the chain to the solvent. Examination of NMR-derived models of $C_6/C_8/C_{10}$-*crypto*-AcpP confirms this hypothesis (Fig. 4.4A). Among these, each C_6- and C_{10}-*crypto*-AcpP exhibited 11 hydrophobic interactions between the tethered pantetheinamide and AcpP residues. Surprisingly, for C_8-*crypto*-AcpP, the number of hydrophobic interactions was larger (15 interactions) due to additional contacts between the acyl chain and AcpP residues, providing a likely explanation for the $^1H_N/^{15}N$ chemical shift perturbations observed between the ^{15}N-atoms of AcpP-tethered C_6/C_{10} vs. C_8-pantetheinamides. For all *crypto*-AcpPs, the side-chain hydroxyl group of residue T40 formed hydrogen bonds with native or non-native nitrogen atoms of the tethered pantetheinamides. Furthermore, for the C_{10}-pantetheinamide, additional hydrogen bonds are formed between the polar atoms of the pantetheinamide and residues at the entrance to AcpP's hydrophobic cavity (namely, E61 and I63). This combination of hydrophobic and polar interactions is likely responsible for the increased stabilization of its

AcpP-sequestered state. It is important to note that ^{15}N-R_2 values reported in this study imply a high degree of mobility for AcpP-tethered Ppant moieties, and as such, the corresponding acyl chains may adopt multiple conformations within AcpP's hydrophobic core, leading to variable interactions with the core residues. Nevertheless, the cumulative effect of these interactions will result in a progressive stabilization of AcpP-sequestered states of acyl chains with six, eight, and ten carbons, as evidenced by chemical-exchange NMR. The above-described kinetic equilibrium will be notably different for longer (12–18 carbons) and shorter (2–4 carbons) acyl chains. This is because longer chains may protrude from AcpP's pocket, as seen in an NMR structure of spinach ACP tethered to an acyl chain comprising 18 carbons (31). Such a protrusion would destabilize the protein fold, culminating in a substantial increase in populations of the solvent-exposed states of tethered acyl chains and subsequently, increased rates of their solvent-mediated hydrolysis (32, 33). Furthermore, the shorter chains may predominantly sample solvent-exposed configurations due to the lack of optimal interactions with AcpP's core, as was observed in prior X-ray and NMR studies of AcpP comprising an acyl chain of four carbons (34, 35). Studies to characterize the kinetic equilibrium exhibited by these longer and shorter AcpP-tethered acyl chains are currently ongoing in our laboratories.

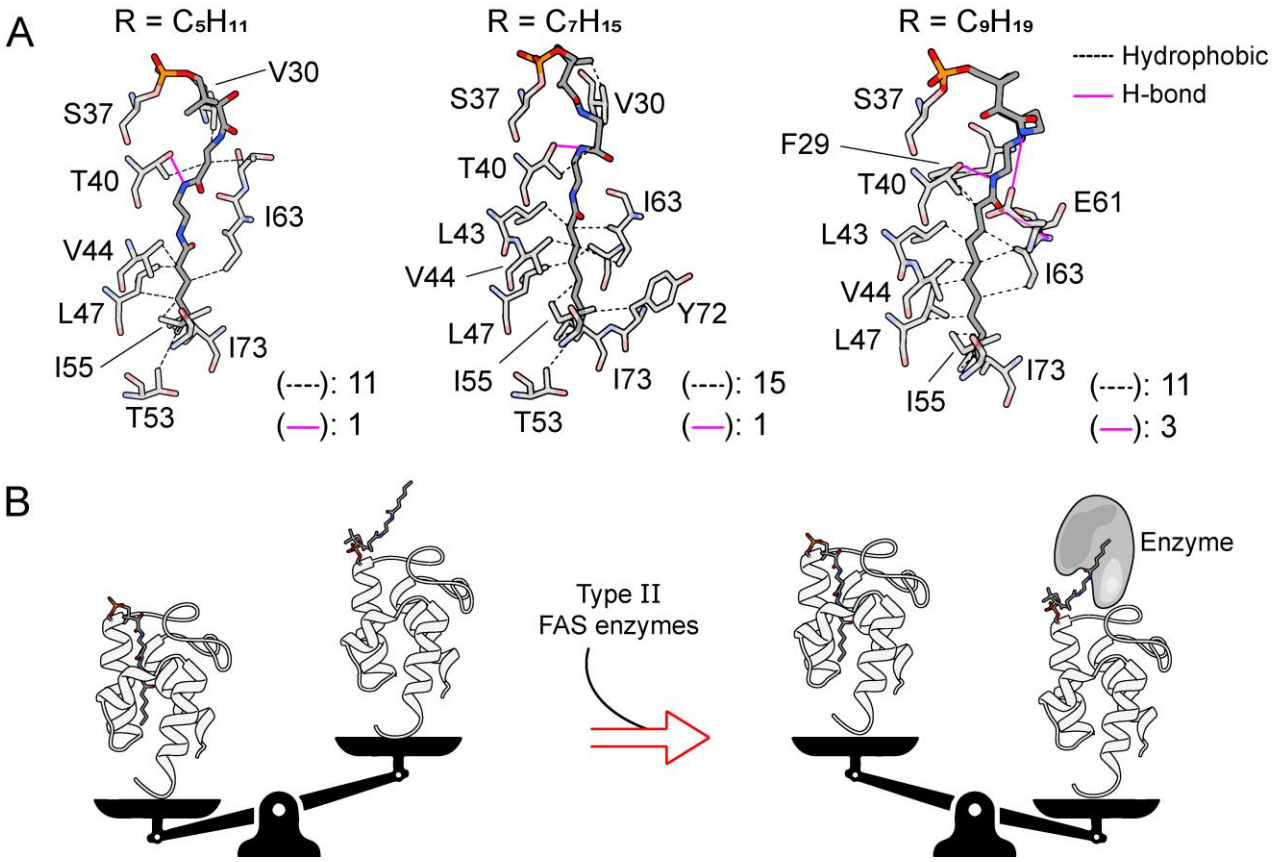


Figure 4.4: (**A**) Hydrophobic interactions (dashed black lines) and hydrogen bonds (pink lines) formed by C_6/C_8/C_{10}-pantetheinamides within AcpP's hydrophobic core. The individual interactions are tallied at the bottom of each panel. Residues of AcpP that interact with tethered pantetheinamides are shown in a semi-transparent stick representation. Structural models derived from NMR data (see Fig. 1D-E) are used to make this comparison using the PLIP (Protein Ligand Interaction Profiler) webserver (42). (**B**) Conformation interplay between AcpP-sequestered and solvent-exposed configurations of AcpP-tethered pantetheinamide (left) and the proposed effect of AcpP-interacting FAS enzymes on this equilibrium (right).

The spontaneous sampling of solvent-exposed conformations by the AcpP-tethered acyl chains is likely to be functionally important. This is because such states will allow appropriate enzymes to gain access to the acyl chains for the corresponding chemical reactions. Additionally, the transient nature of these states will shield AcpP's active thioester from the solvent, preventing premature chain termination. Many enzymes within type II FAS are promiscuous for acyl chain lengths (36). We envision that solvent-exposed conformations of AcpP-tethered acyl chains may be recognized by these enzymes, consequently culminating in the translocation of exposed acyl

chains into their hydrophobic catalytic sites, leading to fatty acid production (Fig. 4.4B). Although it is tempting to correlate the kinetic rates and populations of solvent-exposed states of AcpP-tethered acyl chains observed here with the previously reported catalytic rates of such promiscuous FAS enzymes (37, 38), we refrain from doing so. This is because whether such conformational changes can influence the catalytic rates of enzymes is unclear and still being debated (39, 40). Finally, we note that the symbiotic combination of chemical biology and modern solution NMR methodology described in this study would be applicable to many other biological pathways and systems (e.g., biosynthesis of polyketides and non-ribosomal peptides, protein ubiquitination, etc.) that employ thioester linkages.

Ongoing work on partner proteins

Building from the aforementioned studies, our next step was to test this new methodology on its ability to analyze the chain-flipping interactions of FabF, an elongating ketosynthase responsible for catalyzing the extension of fatty acyl chain by two-carbon units from C4 up to C18. Utilizing ^{15}N-CEST experiments, FabF was introduced to the C_6 and C_8 ^{15}N-labeled pantetheinamides attached to ^{15}N-labeled AcpP (Figure 4.5). These dynamic experiments were able to preliminarily monitor the movement of backbone amides within AcpP with **1a** and **1b** moving between AcpP-sequestered (State **A**) into lowly-populated chain-flipped (State **C**) on the millisecond timescale. Analysis of the pantetheinamides, in the presence of the partner protein, reveals a higher sequestered population than when attached to the carrier protein alone (Table 1). The C_6-*crypto*-AcpP sees an increase from 95.13% to 96.18% sequestered, while C_8-*crypto*-AcpP sees a smaller increase from 98.66 to 99.05%. The increase in sequestered population was analyzed alongside a surge in the backward exchange rate constant, $k_{CA,}$ for the probes. The experiment quantified the rate for C_6-*crypto*-AcpP being sequestered rose from 246.12 ± 2.43 s^{-1} to 420.5 ± 9.6 s^{-1} and C_8-

crypto-AcpP was seen to increase from 611.07 ± 10.69 s^{-1} to 901.8 ± 28.1 s^{-1}. However, the forward exchange rate from State **A** to State **C** were not substantially different in the presence of the partner protein, going from k_{AC} : 12.6 ± 0.1 s^{-1} to 16.7 ± 0.4 s^{-1} for C_6-crypto-AcpP and k_{AC} : 8.3 ± 0.2 s^{-1} to 9.2 ± 0.3 s^{-1} for C_8-crypto-AcpP.

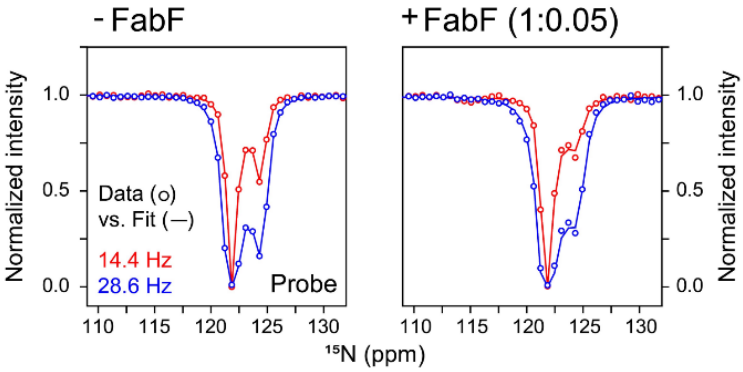


Figure 4.5: ^{15}N CEST profiles for C_6-*crypto*-AcpP without FabF (left) and with the addition of 0.05 equivalents of FabF (right).

The increase in the backward rate constants, k_{CA}, suggests the partner protein is testing the tethered pantetheinamide, specifically the pantetheine portion at the start of the probe, at a faster rate than we can visualize. Thus it is hypothesized the pantetheinamide is quickly returning it to AcpP's hydrophobic pocket once it cannot be transferred onto FabF's active site. This leads to the increase in State **A** that is visualized, demonstrating that the longer the acyl chain, the faster it is getting returned to AcpP, most likely due to the hydrophobic interactions within AcpP's cavity reforming. With the hydrophobic and polar interactions of AcpP's pocket being responsible for the stabilization of the sequestered state instead of the chain-flipped state. This agrees with data seen from the *crypto*-AcpP studies alone.

Table 4.2: Populations (P_A and P_C), exchange (k_{AC} and k_{CA}), and equilibrium (K_{eq}) rate constants for the exchange between AcpP-sequestered (State **A**) and chain-flipped (State **C**) conformations of AcpP-tethered pantetheinamides with FabF.[(a)]

	P_A (%)	P_C (%)	k_{AC} (s^{-1})	k_{CA} (s^{-1})	K_{eq}[(b)]
1 C_6-*crypto*-AcpP : 0.05 FabF[(c)]					
	96.18 ± 0.02	3.82 ± 0.02	16.7 ± 0.4	420.5 ± 9.6	25.18 ± 0.08
1 C_8-*crypto*-AcpP : 0.05 FabF[(c)]					
	99.05 ± 0.02	0.95 ± 0.02	9.2 ± 0.3	901.8 ± 28.1	104.26 ± 1.10

a. NMR experiments were performed at spectrometer ^1H frequency of 800 MHz at 37 °C. The buffer conditions were as follows: 50 mM potassium phosphate, pH 7.4, and 0.01% sodium azide. Data were collected on uniformly ^{15}N-labeled AcpP tethered to selectively ^{15}N-labeled pantetheinamides (protein concentration = 1 mM).

b. The K_{eq} values are given by $K_{eq} = P_A / P_C$ and have been rounded to two decimal places.

c. For ^{15}N-CEST datasets, the set of global fitted parameters were $\{k_{AC}; P_C\}$. The corresponding population of State **A** (P_A) and backward rate constat (k_{BA}) were calculated using the relationships: $P_A = 100 - P_C$; and $k_{CA} = P_A \times k_{AC} / P_C$.

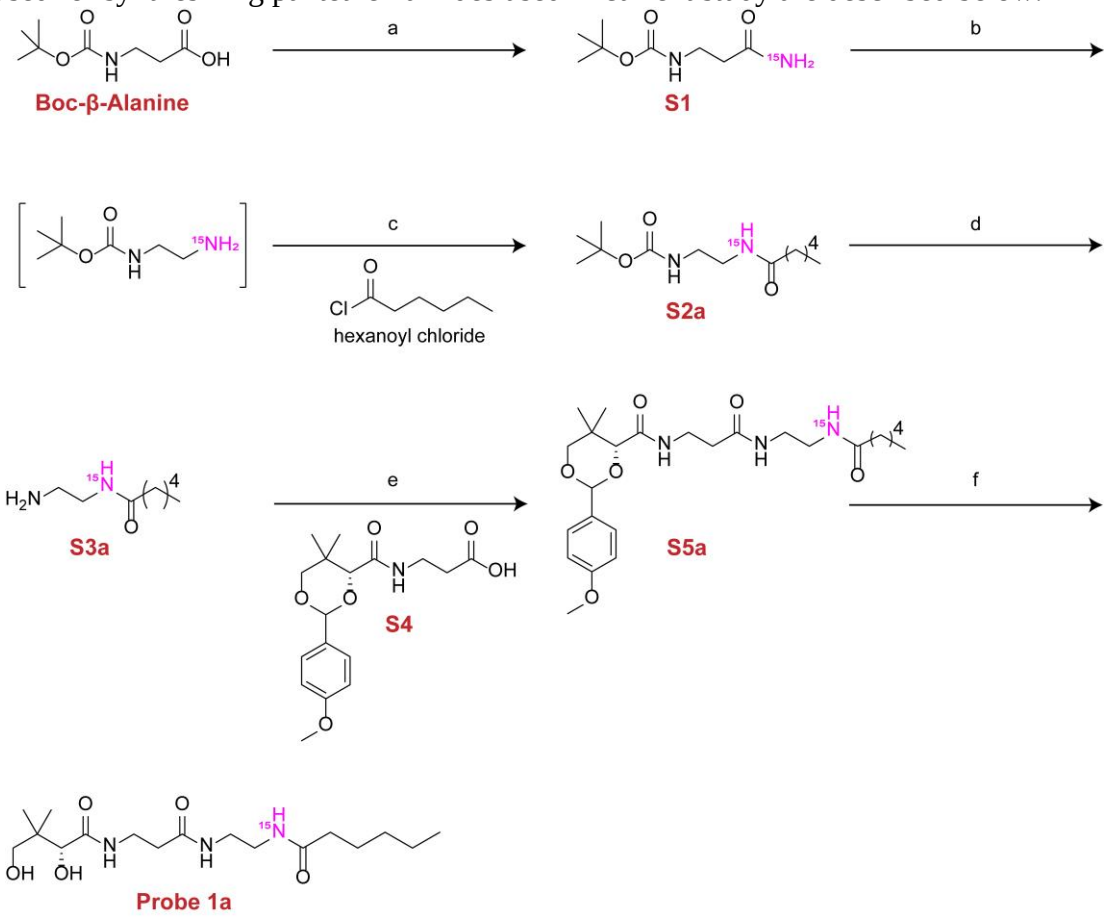
The experiments conducted in this study are the first reported quantification of the rate of chain-flipping interactions between AcpP and a partner protein within the fatty acid biosynthetic pathway. Follow up studies are currently being conducted to ensure the robustness of the method with the ^{15}N-labeled pantetheinamides tethered to unlabeled AcpP in the presence of FabF. Though a state of sequestration inside of the partner protein pocket cannot be directly visualized, the valuable kinetic data of the chain flipping interactions from ^{15}N-CEST NMR provides insight into the rapid evolution of enzyme modification within the fatty acid biosynthetic pathway.

4.4 Supporting Information

Materials

Reagents for chemical synthesis, protein expression, and purification were purchased from Acros Organics, Honeywell Research Chemicals (Fluka), Sigma-Aldrich, Thermo Fisher Scientific, and/or Tokyo Chemical Industry (TCI) Co., Ltd. Deuterated solvents, and regents for nuclear magnetic resonance (NMR) isotopic enrichment were purchased from Sigma-Aldrich and/or Cambridge Isotope Laboratories, Inc. All chemical reactions were performed in scintillation vials and stirred with Teflon-coated stir bars using mechanical stirrers (IKA laboratory technology). Analytical thin layer chromatography (TLC) and preparative TLC were performed on silica gel glass 60 F254 precoated glass plates (Supelco Analytical, catalog no. 1.00390). TLC plates were visualized with ultraviolet light and/or an appropriate stain (potassium permanganate, bromocresol green, 2,4-dinitrophenylhydrazine, ninhydrin, and ceric ammonium molybdate). Flash chromatography was carried out using silica gel 60 (Fisher Scientific, catalog no. S825).

Chemical synthesis of selectively ^{15}N-labeled pantetheinamide probes. Schemes and individual steps used for synthesizing pantetheinamides used in current study are described below.



Scheme S4.1. Synthesis of a selectively ^{15}N-labeled C_6-pantetheinamide. The letters on the arrows denote the following reagents and reaction conditions: (**a**) 1,1'-carbonyldiimidazole (CDI), dry tetrahydrofuran (THF), triethylamine (TEA), ^{15}N-labeled ammonium chloride ($^{15}NH_4Cl$), water, room temperature, reaction yield: ~70% of compound **S1**; (**b**) phenyliodine(III) diacetate (PIDA), pyridine, hexanoyl chloride, water/ethyl acetate/acetonitrile, 0 °C to room temperature; (**c**) TEA, ethyl acetate, room temperature, reaction yield: ~34% of compound **S2a**; (**d**) Trifluoroacetic acid (TFA), TEA, dichloromethane, room temperature; (**e**) Hünig's base (i.e., N,N-diisopropylethylamine), 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC), 4-dimethylaminopyridine (DMAP), dichloromethane, room temperature, reaction yield: ~74% of compound **S5a**; (**f**) 80% glacial acetic acid, room temperature, reaction yield: ~25% of **1a**. Notes: Compound **S4** was prepared according to our previous protocol (S1). The reported reaction yields are measured using chromatographically isolated and spectroscopically homogeneous materials.

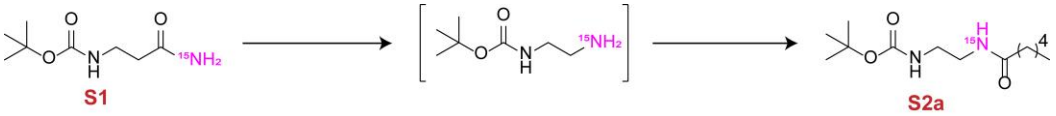
Individual steps in the synthesis of a selectively ^{15}N-labeled C_6-pantetheinamide.

Synthesis of compound S1 [tert-butyl (3-(amino-^{15}N)-3-oxopropyl)carbamate].

Reagents: CDI (97%; Acros Organics), Boc-β-alanine, (99%; Acros Organics), TEA (>99%; Thermo Fisher Scientific), and $^{15}NH_4Cl$ (98% of ^{15}N; Sigma-Aldrich). Notes: All reagents were used without further purification. The numbers in parentheses represent the percent purity of the reagents.

Procedure: CDI (486 mg, 3 mmol) was added in small portions to a Boc-β-alanine solution (568 mg, 3 mmol) in dry THF (1.5 mL). TEA (836 µL, 6 mmol) was mixed in an aqueous solution of $^{15}NH_4Cl$ (54.1 µL) and was then diluted with dry THF (1.5 mL). The resultant two solutions, namely $^{15}NH_4Cl$ and Boc-β-alanine, were mixed and stirred at room temperature. After ~16 h, the solution was concentrated and purified using flash chromatography (mobile phase composition: 2:1 ethyl acetate/acetone), which yielded compound **S1** as a white solid (396 mg, yield: ~70%). The latter (molecular formula: $C_8H_{16}N^{15}NO_3Na$) was analyzed using high-resolution mass spectrometry (HRMS): calculated mass = 212.1023 Da; experimental mass = 212.1026 Da.

Synthesis of compound S2a [tert-butyl (2-hexanamidoethyl)carbamate].

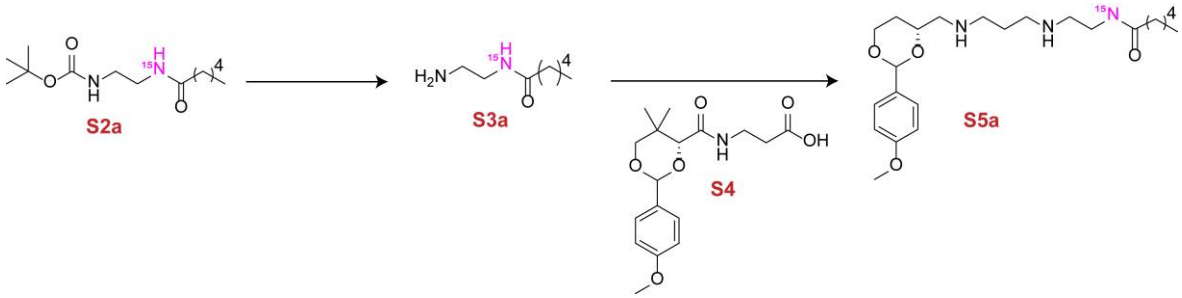


Reagents: PIDA (>97%; TCI Co., Ltd.), pyridine (99%, Thermo Fisher Scientific), TEA (>99%, Thermo Fisher Scientific), hexanoyl chloride (97%, Acros Organics). Notes: All reagents other than pyridine were used without further purification; pyridine was distilled before use.

Procedure: PIDA (103 mg, 0.32 mmol) was added to a solution of compound **S1** (50 mg, 0.27 mmol) dissolved in water/ethyl acetate/acetonitrile (1:2:2 vol/vol; total volume = 1.3 mL) and stirred at 0 °C. After ~15 min, pyridine (21 µL, 0.27 mmol) was added, and the resultant solution was warmed to room temperature. After ~16 h, the reaction mixture was concentrated into an oil,

followed by its dissolution in ethyl acetate (2.7 mL). TEA (74 µL, 0.53 mmol) and hexanoyl chloride (0.04 mL, 0.27 mmol) were added to this solution in a stepwise manner. After ~3 h, the reaction mixture was washed with water, saturated sodium bicarbonate, saturated NH_4Cl, and brine. The organic layer was dried over anhydrous sodium sulfate and evaporated to dryness. The residue was purified using flash chromatography (mobile phase gradient: 1:1 ethyl acetate/hexane to 2:1 ethyl acetate/hexane) to yield compound **S2a** as a white solid (23 mg, yield: ~34%). The latter (molecular formula: $C_{13}H_{26}N^{15}NO_3Na$) was analyzed using HRMS: calculated mass = 282.1806 Da; experimental mass = 282.1806 Da.

Synthesis of compound S5a [(4R)-2-(4-methoxyphenyl)-5,5-dimethyl-N-(3-oxo-3-((2-hexanamidoethyl)amino) propyl)-1,3-dioxane-4-carboxamide]

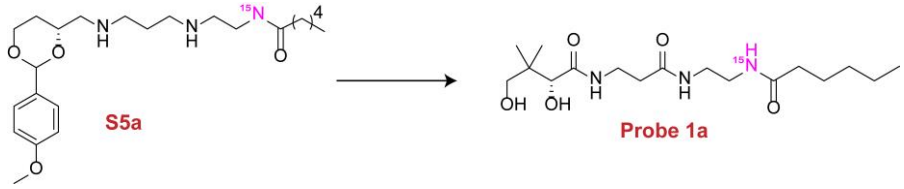


Reagents: TFA (99%, Sigma-Aldrich), TEA (>99%, Thermo Fisher Scientific), EDC (98%; Combi-Blocks, Inc.), DMAP (98%; CreoSalus, Inc.), Hünig's base (99.5+%; Acros Organics). Notes: All reagents other than Hünig's base were used without further purification; Hünig's base was distilled before use.

Procedure: TFA (26 µL, 0.4 mmol) was added to a solution of compound **S2a** (23 mg, 0.08 mmol) in dichloromethane. After ~3 h, TEA (56 µL, 0.4 mmol) was added to the reaction, followed by compound **S4** (27 mg, 0.08 mmol), EDC (12 mg, 0.08 mmol), DMAP (20 mg, 0.16 mmol), and Hünig's base (28 µL, 0.16 mmol). After ~16 h, the reaction mixture was diluted with dichloromethane and washed with water (three times), saturated sodium bicarbonate, saturated NH_4Cl, and brine. The organic layer was dried over anhydrous sodium sulfate and evaporated to

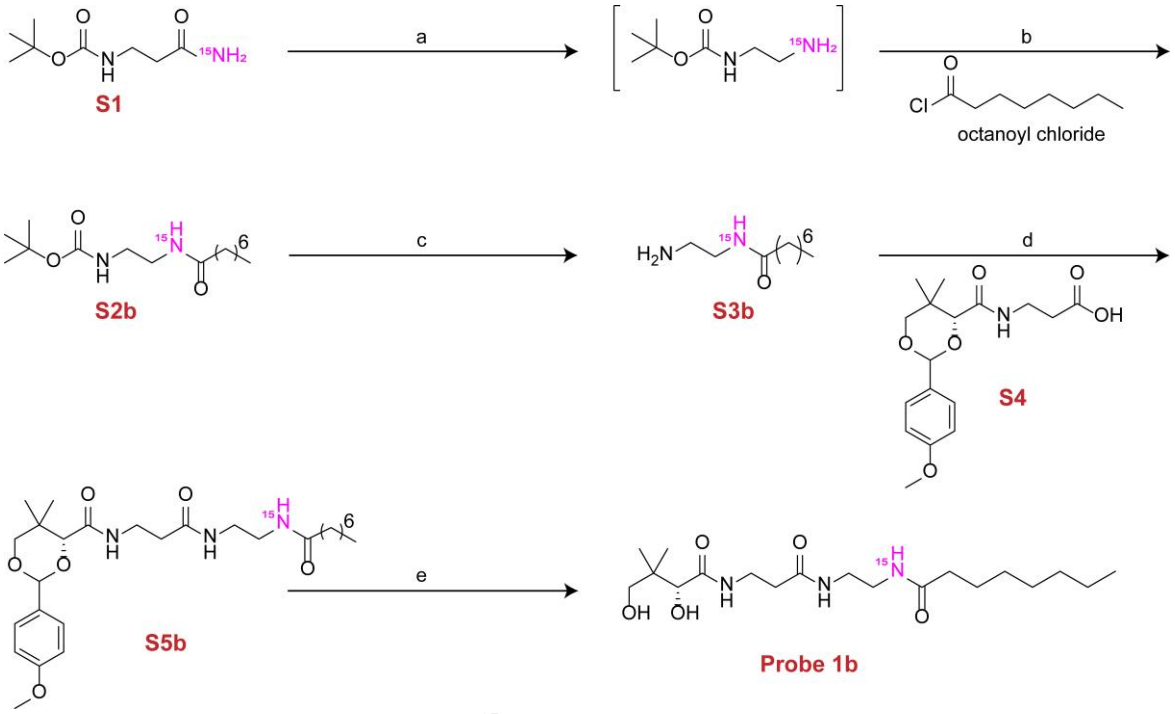
dryness. The resultant residue was purified by flash chromatography (mobile phase gradient: ethyl acetate to 2:1 ethyl acetate/acetone) to yield compound **S5a** as a white solid (30 mg, yield: ~74%). The latter (molecular formula: $C_{25}H_{40}N_2{}^{15}NO_6$) was analyzed using HRMS: calculated mass = 479.2882 Da; experimental mass = 479.2880 Da.

Synthesis of selectively ^{15}N-labeled C$_6$-pantetheinamide probe 1a [(R)-N-(2-(3(2,4dihydroxy3,3dimethylbutanamido)propanamido)ethyl)hexanamide]



Reagents: Acetic acid, glacial (≥99.7% wt/wt; Thermo Fisher Scientific), used without further purification.

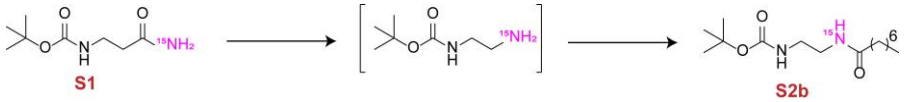
Procedure: Compound **S5a** (12 mg, 24 µmol) was dissolved in 80% acetic acid (0.24 mL) and stirred. After ~20 h, the crude reaction mixture was concentrated into an oil. The residue was purified using flash chromatography (mobile phase composition: 2:1 ethyl acetate/acetone) to yield probe **1a** as a white solid (2.3 mg, yield: ~25%). The latter (molecular formula: $C_{17}H_{34}N_2{}^{15}NO_5$) was analyzed using HRMS: calculated mass = 361.2463 Da; experimental mass = 361.2462 Da.



Scheme S4.2. Synthesis of a selectively ^{15}N-labeled C_8-pantetheinamide. The letters on the arrows denote the following reagents and reaction conditions: (**a**) PIDA, pyridine, water/ethyl acetate/acetonitrile, 0 °C to room temperature; (**b**) TEA, ethyl acetate, room temperature, reaction yield: ~28% of compound **S2b**; (**c**) TFA, dichloromethane, room temperature; (**d**) Hünig's base, EDC, DMAP, dichloromethane, room temperature, reaction yield: ~52% of compound **S5b**; (**e**) 80% glacial acetic acid, room temperature, reaction yield: ~89% of Probe **1b**.

Individual steps in the synthesis of a selectively ^{15}N-labeled C_8-pantetheinamide probe.

Synthesis of compound S2b [tert-butyl (2-octanamidoethyl)carbamate].

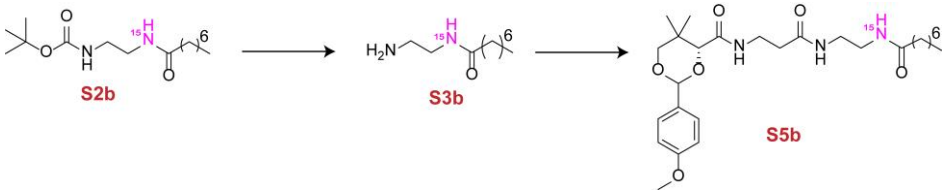


Reagents. PIDA (>97%; TCI Co., Ltd.), pyridine (99%, Thermo Fisher Scientific), TEA (>99%, Thermo Fisher Scientific), octanoyl chloride (97%, Acros Organics). All reagents other than pyridine were used without further purification; pyridine was distilled before use.

Procedure: PIDA (318 mg, 0.98 mmol) was added to a solution of compound **S1** (155 mg, 0.82 mmol) dissolved in water/ethyl acetate/acetonitrile (1:2:2 vol/vol, total volume = 4.12 mL) and stirred at 0 °C. After ~15 min, pyridine (66 µL, 0.82 mmol) was added, and the resultant solution

was warmed to room temperature. After ~16 h, the reaction mixture was concentrated into an oil, which was then dissolved in ethyl acetate (8.2 mL). TEA (0.23 mL, 1.65 mmol) and octanoyl chloride (0.1 mL, 0.82 mmol) were then added to this solution in a stepwise fashion. After ~3 h, the resultant reaction mixture was washed with water, saturated sodium bicarbonate, saturated NH_4Cl, and brine. The organic layer was dried over anhydrous sodium sulfate and evaporated to dryness. The residue was purified using flash chromatography (mobile phase gradient: 1:1 ethyl acetate/hexane to 2:1 ethyl acetate/hexane) to yield compound **S2b** as a white solid (67 mg, yield: ~28%). The latter (molecular formula: $C_{15}H_{30}N^{15}NO_3Na$) was analyzed using HRMS: calculated mass = 310.2119 Da; experimental mass = 310.2122 Da.

Synthesis of compound S5b [(4R)-2-(4-methoxyphenyl)-5,5-dimethyl-N-(3-oxo-3-((2-octanamidoethyl)amino) propyl)-1,3-dioxane-4-carboxamide].

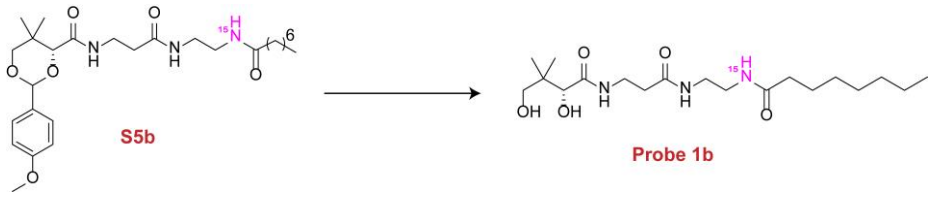


Reagents: TFA (99%, Sigma-Aldrich), TEA (>99%, Thermo Fisher Scientific), EDC (98%; Combi-Blocks, Inc.), DMAP (98%; CreoSalus, Inc.), Hünig's base (99.5+%; Acros Organics). All reagents other than Hünig's base were used without further purification; Hünig's base was distilled before use.

Procedure: TFA (41 µL, 0.63 mmol) was added to a solution of compound **S2b** (36 mg, 0.13 mmol) in dichloromethane. After ~3 h, TEA (88 µL, 0.63 mmol) was added to the reaction, followed by compound **S4** (42 mg, 0.13 mmol), EDC (20 mg, 0.13 mmol), DMAP (44 mg, 0.25 mmol), and Hünig's base (44 µL, 0.16 mmol). After ~16 h, the reaction mixture was diluted with dichloromethane and washed with water (three times), saturated sodium bicarbonate, saturated NH_4Cl, and brine. The organic layer was dried over anhydrous sodium sulfate and evaporated to

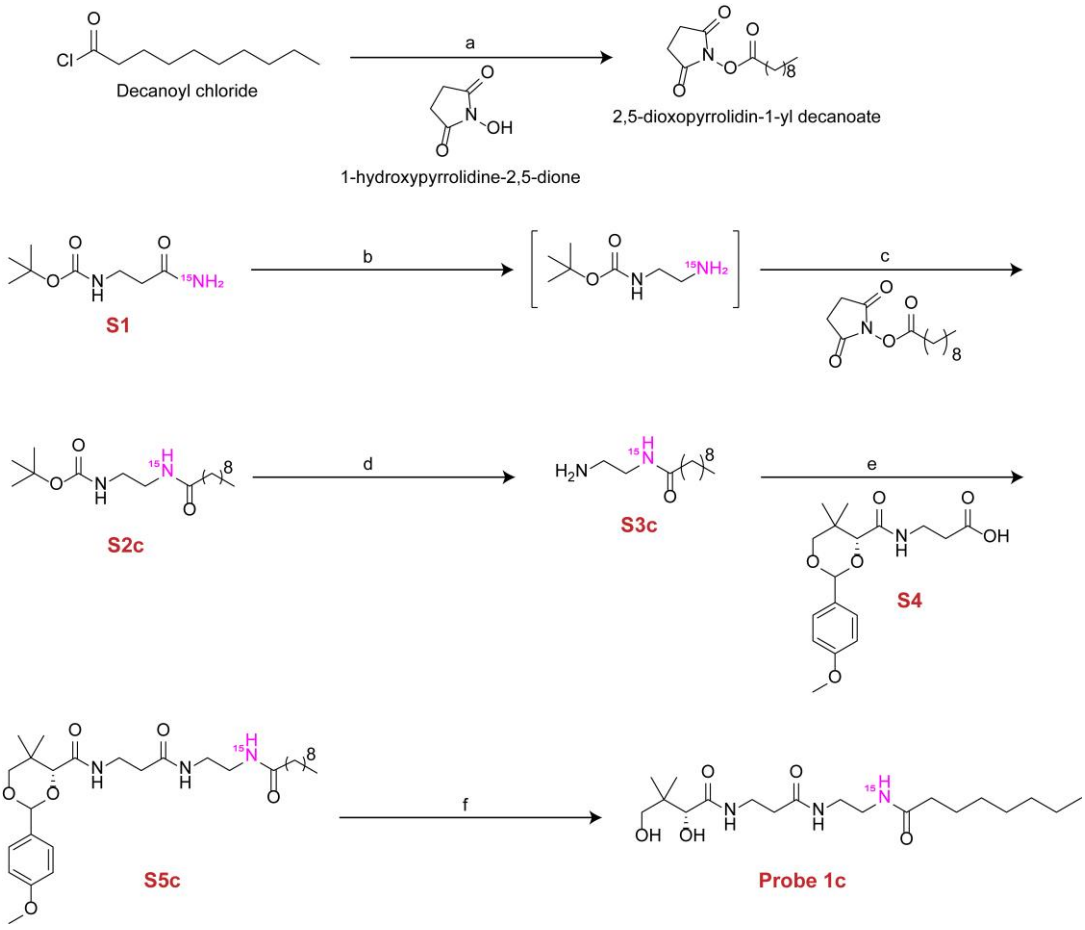
dryness. The resultant residue was purified by flash chromatography (mobile phase gradient: ethyl acetate to 2:1 ethyl acetate/acetone) to yield compound **S5b** as a white solid (30 mg, yield: ~74%). The latter (molecular formula: $C_{27}H_{44}N_2{}^{15}NO_6$) was analyzed using HRMS: calculated mass = 507.3195 Da; experimental mass = 507.3196 Da.

Synthesis of selectively ^{15}N-labeled C_8-pantetheinamide probe 1b [(R)-N-(2-(3-(2,4-dihydroxy-3,3-dimethylbutanamido)propanamido)ethyl) octanamide]



Reagents: Acetic acid, glacial (≥99.7% wt/wt; Thermo Fisher Scientific), used without further purification.

Procedure: Compound **S5b** (28 mg, 0.055 mmol) was dissolved in 80% acetic acid (0.55 mL) and stirred. After ~20 h, the crude reaction mixture was concentrated into an oil. The residue was purified using flash chromatography (mobile phase composition: 2:1 ethyl acetate/acetone) to yield probe **1b** as a white solid (2.3 mg, yield: ~25%). The latter (molecular formula: $C_{19}H_{38}N_2{}^{15}NO_5$) was analyzed using HRMS: calculated mass = 389.2776 Da; experimental mass = 389.2777 Da.



Scheme S4.3. Synthesis of a selectively ^{15}N-labeled C_{10}-pantetheinamide. The letters on the arrows denote the following reagents and reaction conditions: (**a**) EDC, dichloromethane, room temperature, reaction yield: ~62% of 2,5-dioxopyrrolidin-1-yl decanoate; (**b**) PIDA, pyridine, water/ethyl acetate/acetonitrile, 0 °C to room temperature; (**c**) dry THF, room temperature, reaction yield: ~85% of compound **S2c**; (**d**) TFA, dichloromethane, room temperature; (**e**) Hünig's base, hexafluorophosphate azabenzotriazole tetramethyl uronium (HATU), DMAP, dichloromethane, room temperature, reaction yield: ~72% of compound **S5c**; (**f**) 80% glacial acetic acid, room temperature, reaction yield: ~70% of **1c**.

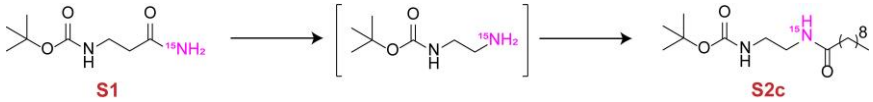
Individual steps in the synthesis of a selectively ^{15}N-labeled C_{10}-pantetheinamide.

Synthesis of 2,5-dioxopyrrolidin-1-yl decanoate.

Reagents: Decanoic chloride (98+%; Thermo Scientific Chemicals), 1-hydroxypyrrolidine-2,5-dione (≥99%; Sigma-Aldrich), and EDC (98%; Combi-Blocks, Inc.). All reagents were used without further purification.

Procedure: 1-hydroxypyrrolidine-2,5-dione (144 mg, 1.25 mmol) and EDC (388 mg, 2.5 mmol) were added in a stepwise fashion to the solution of decanoic acid (250 mg, 1.25 mmol) in dichloromethane. The resultant solution was stirred at room temperature for ~16 h. The reaction mixture was filtered (how?) and concentrated under reduced pressure (how?). The crude solid was recrystallized in isopropyl alcohol to yield 2,5-dioxopyrrolidin-1-yl decanoate as a white solid (230 mg, yield: 62%). The latter (molecular formula: $C_{14}H_{23}NO_4Na$) was analyzed using HRMS: calculated mass = 292.1519 Da; experimental mass = 292.1518 Da.

Synthesis of compound S2c [tert-butyl (2-decanamidoethyl)carbamate].

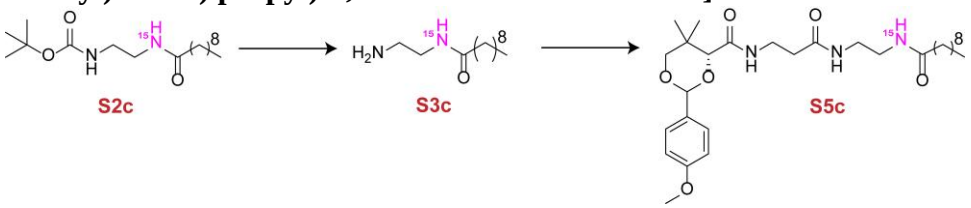


Reagents: PIDA (>97%; TCI Co., Ltd.), pyridine (99%, Thermo Fisher Scientific), 2,5-dioxopyrrolidin-1-yl decanoate (synthesis is described above). All reagents other than pyridine were used without further purification; pyridine was distilled before use.

Procedure: PIDA (97 mg, 0.3 mmol) was added to a solution of compound **S1** (47 mg, 0.25 mmol) dissolved in water/ethyl acetate/acetonitrile (1:2:2 vol/vol; total volume = 1.3 mL) and stirred at 0 °C. After ~15 min, pyridine (20 µL, 0.25 mmol) was added, and the resultant solution was warmed to room temperature. After ~16 h, the reaction mixture was concentrated into an oil, which was then dissolved in dry THF (0.13 mL) and mixed with a solution of 2,5-dioxopyrrolidin-1-yl decanoate (51 mg, 0.19 mmol) in dry THF (XYZ mL). After ~16 h, the crude reaction mixture was evaporated to dryness, and the resultant residue was purified using flash chromatography (mobile phase composition: 2:1 ethyl acetate/hexane) to yield compound **S2c** as a white solid (67 mg, yield:

~85%). The latter (molecular formula: $C_{17}H_{34}N^{15}NO_3Na$) was analyzed using HRMS: calculated mass = 338.2432 Da; experimental mass = 338.2433 Da.

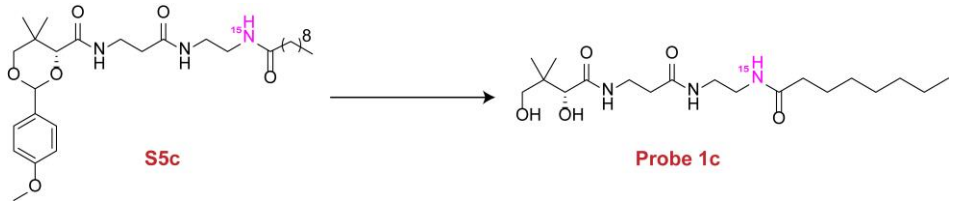
Synthesis of compound S5c [(4R)-2-(4-methoxyphenyl)-5,5-dimethyl-N-(3-oxo-3-((2decanamidoethyl)amino) propyl)-1,3-dioxane-4-carboxamide].



Reagents: TFA (99%, Sigma-Aldrich), TEA (>99%, Thermo Fisher Scientific), EDC (98%; Combi-Blocks, Inc.), DMAP (98%; CreoSalus, Inc.), Hünig's base (99.5+%; Acros Organics). All reagents other than Hünig's base were used without further purification; Hünig's base was distilled before use.

Procedure: TFA (29 µL, 0.45 mmol) was added to a solution of compound **S2c** (28 mg, 0.089 mmol) in dichloromethane. After ~3 h, TEA (62 µL, 0.45 mmol) was added to the reaction, followed by compound **S4** (28 mg, 0.089 mmol), EDC (34 mg, 0.089 mmol), DMAP (31 mg, 0.18 mmol), and Hünig's base (44 µL, 0.16 mmol). After ~16 h, the reaction mixture was diluted with dichloromethane and washed with water (three times), saturated sodium bicarbonate, saturated NH_4Cl, and brine. The organic layer was dried over anhydrous sodium sulfate and evaporated to dryness. The resultant residue was purified by flash chromatography (mobile phase gradient: ethyl acetate to 2:1 ethyl acetate/acetone) to yield compound **S5c** as a white solid (34 mg, yield: ~72%). The latter (molecular formula: $C_{29}H_{48}N_2{}^{15}NO_6$) was analyzed using HRMS: calculated mass = 535.3508 Da; experimental mass = 535.3506 Da.

Synthesis of selectively ^{15}N-labeled C_{10}-pantetheinamide 1c [(R)-N-(2-(3-(2,4-dihydroxy-3,3-dimethylbutanamido)propanamido)ethyl) decanamide]



Reagents: Acetic acid, glacial (≥99.7% wt/wt; Thermo Fisher Scientific), used without further purification.

Procedure: Compound **S5c** (23 mg, 0.043 mmol) was dissolved in 80% acetic acid (0.43 mL) and stirred. After ~20 h, the crude reaction mixture was concentrated into an oil. The residue was purified using flash chromatography (mobile phase composition: 2:1 ethyl acetate/acetone) to yield probe **1c** as a white solid (12.5 mg, yield: ~70%). The latter (molecular formula: $C_{21}H_{41}N_2{}^{15}NO_5Na$) was analyzed using HRMS: calculated mass = 439.2909 Da; experimental mass = 439.2910 Da.

All above-mentioned compounds were characterized using one-dimensional (1D) NMR spectroscopy and mass spectrometry (all NMR spectra are included in the Appendix section). 1D $^{1}H/^{13}C$ NMR spectra were recorded on a JEOL ECA500 spectrometer equipped with a 2-channel inverse-detect broad-band probe or a Varian VX500 spectrometer equipped with an Xsens cold probe. ^{13}C NMR spectra were recorded with complete proton decoupling. Spectra were processed and analyzed using the MestreNova software suite (MestreLab Research S.L.). Low resolution electrospray mass spectrometry (ESI-MS) analyses were performed using an LCQ-Deca spectrometer (Thermo Finnigan) and high-resolution MS (HR-MS) analyses and liquid chromatography HR-MS were conducted using an Agilent 6230 time of flight mass spectrometer with Jetstream ESI source.

Expression and purification of acyl carrier protein (AcpP) of *Escherichia coli* (*E. coli*). Briefly, AcpP (residues 1–78, UniProt entry no. P0A6A8) was expressed in BL21(DE3) competent cells (Agilent Technologies). Note that AcpP was expressed with a C-terminal non-cleavable polyhistidine tag. The primary sequence of the tag is as follows: LEHHHHHH. AcpP was expressed at 16 °C. Briefly, cells were grown at 37 °C in 1 L Luria-Bertani (LB; Thermo Fisher Scientific, catalog no. BP97235) medium at natural isotopic abundance. Cells were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at an optical density of ~0.7 at 600 nm, and concomitantly, the temperature of cell culture was reduced to 16 °C. Cells were harvested after ~16 h (2000 g for 30 min). For isotopic labeling, cells were initially grown at 37 °C in 1 L LB medium to an optical density of ~0.7 at 600 nm. Cells were collected by centrifugation (2000 g for 30 min) and resuspended in 250 mL minimal M9 medium (sans the carbon/nitrogen source). Cells were recollected by centrifugation (2000 g for 30 min) and suspended in 1 L minimal M9 medium containing 1 g/L $^{15}NH_4Cl$ for ^{15}N labeling or 1 g/L $^{15}NH_4Cl$ and 4 g/L $^{13}C_6$-d-glucose for $^{15}N/^{13}C$ labeling. The resultant cell culture was incubated at 37 °C for an hour to acclimatize the cells to M9 medium. Cells were then induced with 1 mM IPTG, followed by the above-described protocol for protein expression and cell harvest.

Briefly for the purification scheme of *E. coli* AcpP, cells were resuspended in a lysis buffer comprising 50 mM Tris(Hydroxymethyl)aminomethane (i.e., Tris), pH 7.4, 150 mM NaCl, and 10% glycerol. Cells were lysed by sonication and cleared by centrifugation (17,400 g, 1 h). The resultant supernatant was loaded onto a HisPur Ni-NTA Resin (Thermo Fischer Scientific) pre-equilibrated with above-mentioned lysis buffer (sans glycerol) and eluted in the same buffer containing 250 mM imidazole. The eluted protein was dialyzed in a buffer comprising 50 mM Tris, pH 7.4, 150 mM sodium chloride, 0.5 mM (tris(2-carboxyethyl)phosphine) or TCEP, 12.5

mM magnesium chloride, and 5 mM manganese chloride at 4 °C. The resultant solution was mixed with 0.1 molar equivalent of acyl carrier protein phosphodiesterase of *E. coli* (also called ACP hydrolase or AcpH) and stirred at 4 °C overnight; note that expression and purification of AcpH have been described in our previous works (S2). The above-described procedure is used to generate homogenous apo-AcpP samples by converting any remnant holo-AcpP created during bacterial expression to apo-AcpP. The resultant product, apo-AcpP, was concentrated (Amicon ultra-15 centrifugal filter units, 3 kDa cutoff; EMD Millipore) and loaded onto a HiLoad 26/600 Superdex 75 prep-grade column (Cytiva) pre-equilibrated with buffer containing 50 mM Tris (pH 7.4) and 150 mM sodium chloride. Relevant fractions were pooled, concentrated (~10 mg/mL), and stored at -20 °C.

Production of *crypto*-AcpP tethered to selectively ^{15}N-labeled pantetheinamides of varying acyl chain lengths. All selectively ^{15}N-labeled pantetheinamides were individually tethered to *apo*-AcpP using our one-pot chemoenzymatic method (S3). This method utilizes three CoA biosynthetic enzymes, namely CoaA, CoaD, and CoaE, to form CoA analogs and a phosphopantetheinyl transferase called Sfp to tether these analogs onto *apo*-AcpP, producing *crypto*-AcpP. Note that the expression and purification of Sfp and CoA biosynthetic enzymes have been described previously (S4, S5). The composition of the final reaction mixture (total volume = 2 mL) was as follows: 50 mM Tris, pH 7.0, 150 mM sodium chloride, 8 mM adenosine triphosphate, 12.5 mM magnesium chloride, CoA biosynthetic enzymes: CoaA, CoaD, and CoaE (each at 0.01 µg/µL), 0.04 µg/µL Sfp, 1 mM TCEP, and 100 µM of *apo*-AcpP. Selectively ^{15}N-labeled pantetheinamides of varying acyl chain lengths (see Schemes S4.1, S4.2, and S4.3) were individually dissolved in a small volume of dimethyl sulfoxide (DMSO) to prepare stock solutions (100 mM). These probes were individually mixed with the above-mentioned reaction mixtures,

and the resultant solutions were incubated at 22 °C for ~12 h. The completion of the reactions and the subsequent formation of *crypto*-AcpP products, hereafter referred to as $C_6/C_8/C_{10}$-*crypto*-AcpP, were monitored using reversed-phase high-performance liquid chromatography (RP-HPLC; Agilent 1100 HPLC system) using an Acclaim reversed-phase C18 column (4.6 x 250 mm, 5 µM particle size, 120 Å; Thermo Fisher Scientific, catalog no. 059149); Fig. S1. The reaction mixtures were loaded onto a HiLoad 26/600 Superdex 75 prep-grade column (Cytiva) pre-equilibrated with 50 mM potassium phosphate (pH 7.4) and 0.01% sodium azide. Relevant fractions of *crypto*-AcpP were pooled, concentrated (~10 mg/mL), and stored at -20 °C. All *crypto*-AcpP samples were further verified by liquid chromatography–electrospray ionization-time-of-flight/mass spectrometry (LC–ESI-TOF/MS) using our previously described protocols (S6-S8).

Protein NMR spectroscopy. Samples of non-labeled/uniformly ^{15}N-labeled/uniformly ^{15}N,^{13}C-labeled $C_6/C_8/C_{10}$-*crypto*-AcpP variants were prepared in a buffer comprising 50 mM potassium phosphate, pH 7.4, and 0.01% sodium azide. Aligned samples of ^{15}N-labeled $C_6/C_8/C_{10}$-*crypto*-AcpP variants were prepared using bacteriophage pf1 (13 mg/mL; Asla Biotech AB). Additionally, ^{15}N-labeled pantetheinamides of varying acyl chain lengths were dissolved in a small volume of DMSO, and the resultant stock solutions (100 mM) were diluted in the above-mentioned buffer to reach a final concentration of ~1 mM (final DMSO concentration: 1%). All NMR samples contained 10% (vol/vol) deuterium oxide.

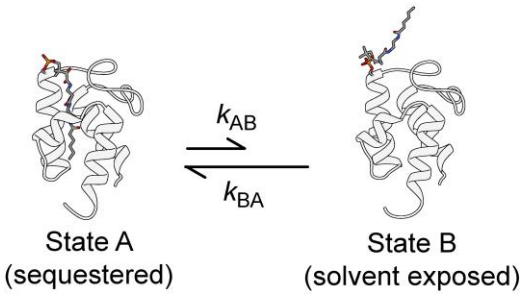
NMR experiments were carried out at 37 °C on Bruker 600 and 800 MHz spectrometers equipped with z-gradient triple resonance cryoprobes. Spectra were processed using NMRPipe (S9) and analyzed using the CCPN software suite (S10). Sequential ^{1}H, ^{15}N, and ^{13}C backbone resonance assignments of $^{15}N/^{13}C$-labeled C_6/C_8-*crypto*-AcpP variants were carried out using three-dimensional (3D) triple resonance experiments (S11). The ^{1}H-^{15}N cross peaks of C_{10}-*crypto*-AcpP

were assigned by reference to the spectrum of C_8-*crypto*-AcpP. NMR chemical shift perturbations between different *crypto*-AcpP variants were determined by analyzing the two-dimensional (2D) HSQC spectra of the corresponding proteins. Perturbations were calculated as follows: $\Delta_{H/N} = \{(\Delta\delta_{HN})^2 + (0.154 \times \Delta\delta_N)^2\}^{1/2}$, where $\Delta\delta_{HN}$ and $\Delta\delta_N$ are the ^1H_N and ^{15}N chemical shift differences in ppm, respectively, between different crypto-AcpP variants. $^1D_{NH}$ residual dipolar couplings (RDCs), given by the difference in $^1J_{NH}$ coupling constants in aligned and isotropic media, were measured using the transverse relaxation optimized spectroscopy (TROSY)-based ARTSY technique (S12) and analyzed with Xplor-NIH (S13, S14). ^{15}N-R_1 and $R_{1\rho}$ measurements (S15) were carried out on uniformly ^{15}N-labeled $C_6/C_8/C_{10}$-*crypto*-AcpP variants.

^{15}N-CEST experiments (S16) were recorded at an 800 MHz spectrometer in an interleaved manner on uniformly ^{15}N-labeled or non-labeled $C_6/C_8/C_{10}$-*crypto*-AcpP variants. The exchange time (T_{CEST}) was set to 400 ms; B1 continuous wave (CW) saturation field strengths of 15 and 30 Hz were used. In each case, the reference spectra were acquired with the T_{CEST} period included but with the B1 field strength set to 0 Hz. For uniformly ^{15}N-labeled AcpP with selectively ^{15}N-labeled pantetheinamides, ^{15}N CW saturation was applied at 50 Hz increments from -900 to +1000 Hz (with the ^{15}N carrier set to 120 ppm). For non-labeled AcpP with selectively ^{15}N-labeled pantetheinamides, ^{15}N CW saturation was applied at 30 Hz increments from -650 to +610 Hz (with the ^{15}N carrier set to 120 ppm). B1 field strengths were calibrated using a 2D version of the nutation experiment (S17). The exact values of the ^{15}N-field strengths were found to be 14.4 and 28.6 Hz for the corresponding calculated values of 15 and 30 Hz, respectively. For C_{10}-*crypto*-AcpP, no characteristic second dips in intensity at the resonance frequencies of the minor species were observed (cf. Fig. 2C and 3B, main text), indicating that the timescale of exchange is not accessible to ^{15}N-CEST measurements and/or because the population of the solvent-exposed conformation

of AcpP-tethered C_{10}-pantetheinamide is below the limit of detection of ^{15}N-CEST (≤0.5%). Therefore, we recorded ^{15}N-Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersions on these C_{10}-*crypto*-AcpP samples. The dispersions were measured at a spectrometer field of 800 MHz with a constant-time period of 120 ms using a pulse scheme with amide proton decoupling that monitors the rates of in-phase ^{15}N coherences (S18). A CW ^{1}H saturation with an 11 kHz radiofrequency (RF) field was applied during the entire CMPG period.

Analysis of ^{15}N CEST and ^{15}N CPMG relaxation dispersions. The ^{15}N CEST data measured on C_6/C_8-*crypto*-AcpP variants were fit globally to a two-state exchange model (**A** ↔ **B**, where **A** is major state; Scheme S4.4) using an in-house MATLAB script (MathWorks Inc.) that is based on the original work of Kay and co-workers (S16) and our previously published report (S19), by numerically solving the McConnell equations (S20). Similarly, the ^{15}N-CPMG dispersions collected on C_{10}-*crypto*-AcpP sample were fit globally to the model shown in Scheme 4. The mathematical and computational framework of this two-state model is described in our earlier work (S6). For ^{15}N-CEST datasets, the set of global parameters included: $\{k_{AB}; P_B\}$, where k_{AB} is the forward rate constat and P_B is the population of State **B**. For ^{15}N-CPMG dataset, the set of global parameters included: $\{k_{AB}; k_{BA}\}$, where k_{AB} and k_{BA} are the forward and backward rate constats, respectively. The uncertainties in the values of the optimized parameters, corresponding to confidence intervals of ± 1 standard deviation were determined from the Jacobian matrix of the non-linear fit. Convergence of the solution was confirmed by varying initial values for all parameters and obtaining the same solution within reported uncertainties.



Scheme S4.4 Two state exchange model **A** ↔ **B**, where **A** is major state and **B** is minor state

Structure determination of C_6/C_8/C_{10}-*crypto*-AcpP. The structure of C_6-*crypto*-AcpP was determined from backbone chemical shifts and backbone amide RDCs using CS-Rosetta (S21-S23). Briefly, a standard CS-Rosetta structure calculation was performed using backbone chemical shifts (^1H_N, ^{15}N, $^{13}C\alpha$, $^{13}C\beta$, and $^{13}C'$) as input to search for the matched three- and nine-residue fragment candidates from the protein coordinate database derived from protein data bank (PDB). Note that the coordinate database did not include previously published NMR and/or X-ray structures of *E. coli* AcpP. As the subsequent step, a de novo Rosetta fragment assembly and structure generation procedure was carried out to use the best matched fragment candidates and backbone amide RDCs as inputs. The calculations were performed to generate 6,000 structures. Among these, the 10 lowest energy structures were used for further analysis and modeling the acyl chains.

The AcpP-sequestered conformation of the tethered C_6 pantetheinamide was modeled using Xplor-NIH (S13, S14) via the following protocol: First, topology and parameter files for the AcpP-tethered C_6 pantetheinamide moiety were generated from the PDB entry 6OKC (S24) using the Xplor-NIH helper program genLigandCIF. Note that this PDB entry represents a cross-linked complex between a ketosynthase called FabB and *crypto*-AcpP tethered to a C_{12} pantetheinamide

determined using X-ray crystallography. First, the C_6 pantetheinamide tag was attached to residue S37 of the lowest-energy CS-ROSETTA model of AcpP, with the entire tag external to the protein. To model the AcpP-sequestered conformation of the C_6-pantetheinamide, initially only S37 and the probe coordinates were allowed to move to fit the heavy atom coordinates of PDB entry 2FAC (S25); the latter comprises a C_6-pantetheine sequestered within the hydrophobic cavity of AcpP. The same PDB entry was used as a template to position C_8-pantetheinamide, whereas PDB entry 2FAE (S25) comprising a C_{10}-pantetheine was used in the case of C_{10}-*crypto*-AcpP.

The modeling calculations employed multiple minimization steps. First, only S37 and the coordinates of tethered pantetheinamides were allowed to move to fit the heavy atom coordinates of the PDB template. This initial minimization employed a PosDiffPot energy term restraining atomic coordinates to appropriate analogs in the PDB template and covalent and TorsionDB energy terms. This initial fit was followed by three rounds of gradient-minimization refinement, using Xplor-NIH's simulationTools.minimizeRefine function, fixing in space the protein backbone and sidechains of the protein not close to the acyl pocket, and adding the RepelPot term to bias against atomic overlap, with the RepelPot scale value sequentially set to values 0.01, 0.1 and 1 during the three minimization rounds. A final gradient minimization was performed allowing all torsion angle degrees of freedom of the protein and the pantetheinamide. The energy minimized in this step was a sum of those terms used in prior steps and terms corresponding backbone amide RDCs measured on the C_6-*crypto*-AcpP, and distance restraints generated from the coordinates of all backbone heavy atoms of the structure which were within 6 Å of each other at this point in the calculation. This procedure was performed separately for C_8 and C_{10} pantetheinamides. The resultant structures were cross-validated against the backbone amide RDCs obtained for the corresponding *crypto*-AcpP samples.

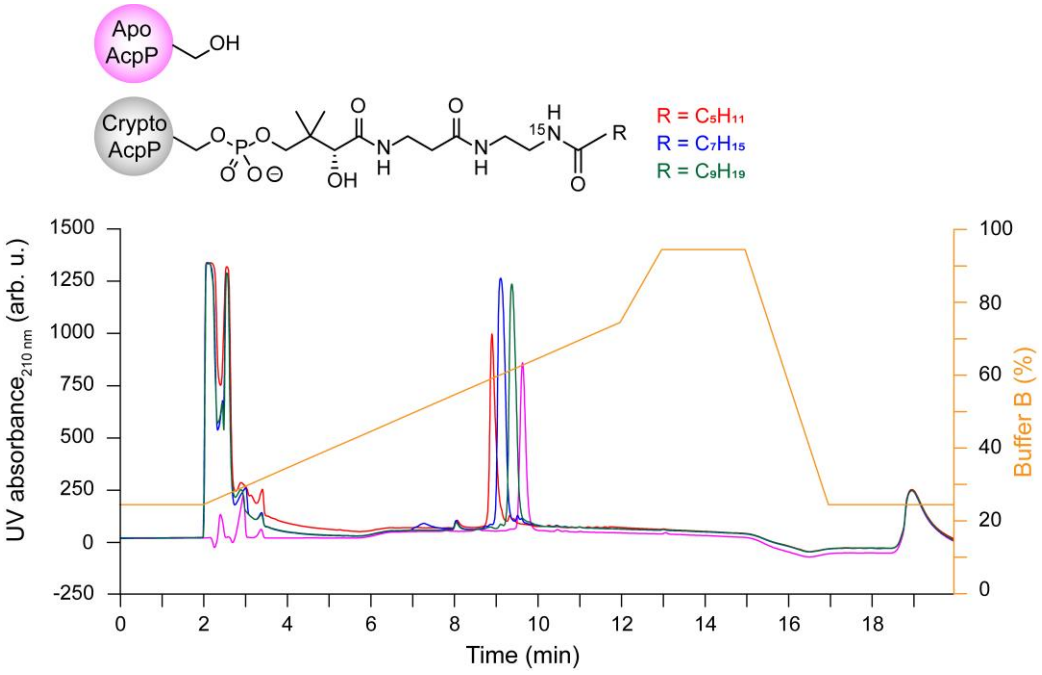


Figure S4.1. Comparison of RP-HPLC elution profiles of *apo*-AcpP (pink) and $C_6/C_8/C_{10}$-*crypto*-AcpP (red, blue, and green, respectively). Crypto-AcpP samples were generated from *apo*-AcpP and selectively ^{15}N-labeled pantetheinamides of varying acyl chain lengths using our one-pot chemoenzymatic method (S3). A schematic representation of the corresponding proteins is shown above the chromatogram. Proteins were detected at an ultraviolet (UV) wavelength of 210 nm (arb. u. = arbitrary units). Chromatography conditions were as follows: Buffer A = 0.05% TFA in water, Buffer B = 0.05% TFA in acetonitrile, and flowrate = 1 mL/min.

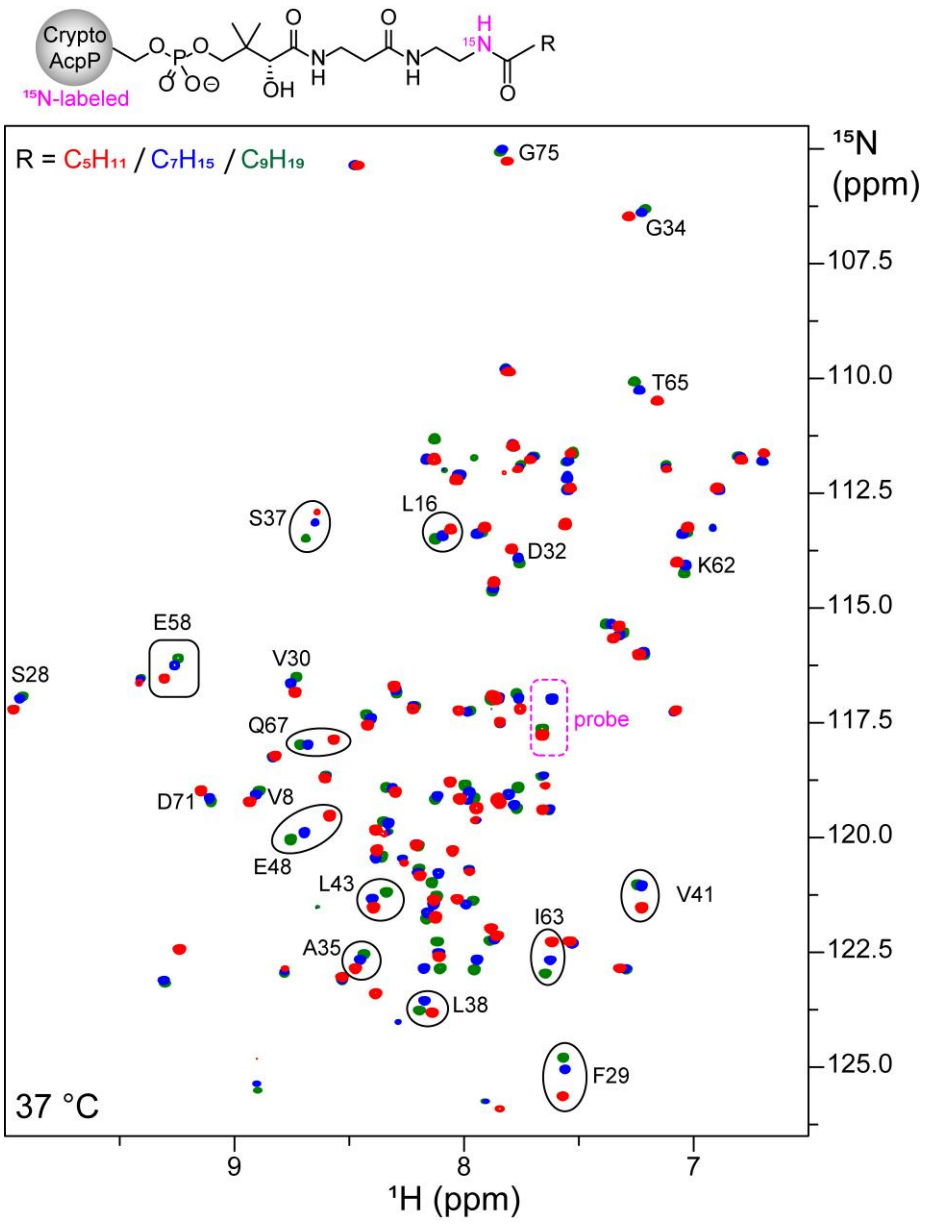


Figure S4.2. NMR spectral analysis of ^{15}N-labeled $C_6/C_8/C_{10}$-*crypto*-AcpP samples. Overlay of expanded regions of ^{1}H-^{15}N HSQC spectra of $C_6/C_8/C_{10}$-*crypto*-AcpP samples (red, blue, and green, respectively). A few residues with isolated ^{1}H-^{15}N cross-peaks that undergo large chemical shift changes upon changing the acyl lengths are labeled. The ^{1}H-^{15}N cross-peaks of the ^{15}N atoms of the pantetheinamides are highlighted using a dashed square (pink). Spectra were recorded at 37 °C and a spectrometer ^{1}H frequency of 800 MHz.

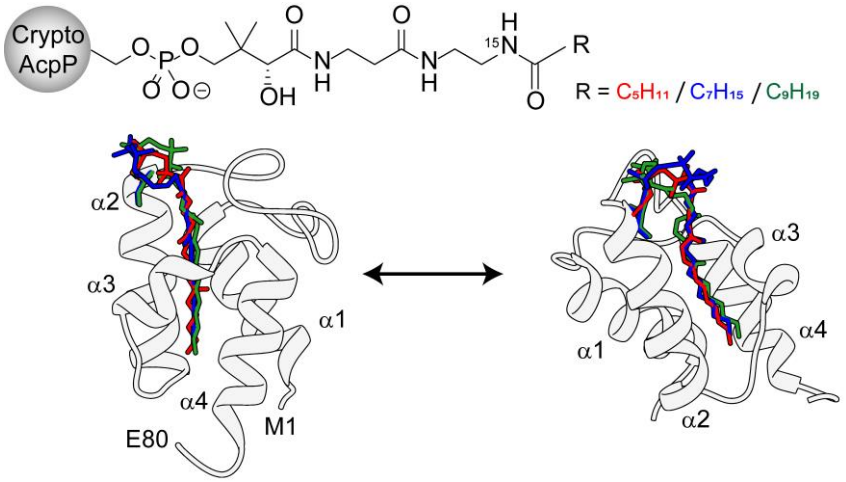


Figure S4.3. Superimposition of C_6/C_8/C_{10}-*crypto*-AcpP models solved using CS-Rosetta and Xplor-NIH. Only the ribbon diagram of C_6-*crypto*-AcpP is shown (white ribbons). Residue S37 and the tethered pantetheinamides of varying acyl chain lengths, C6, C8, and C10, are shown in stick representation (red, blue, and green, respectively).

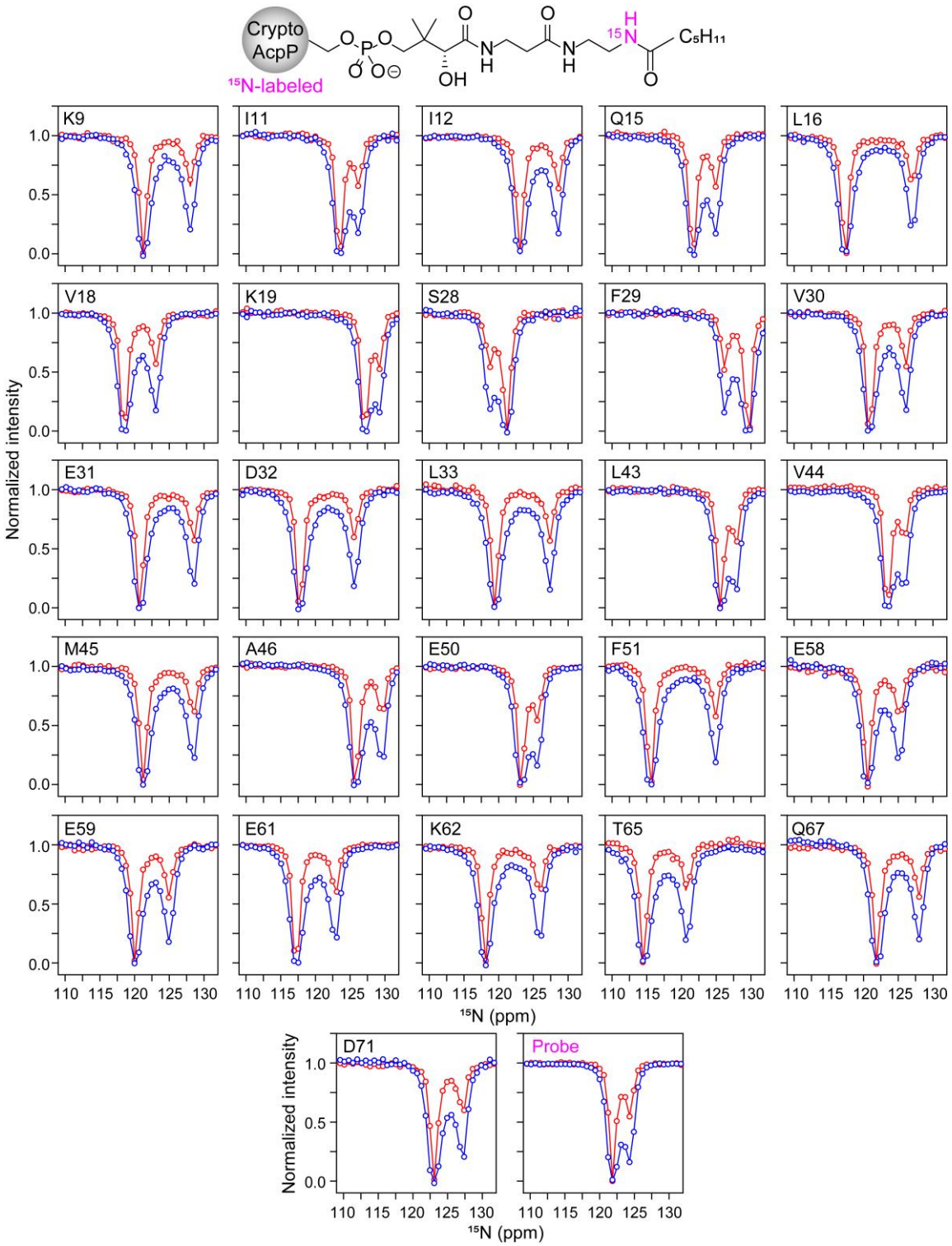


Figure S4.4. NMR analysis of ^{15}N-labeled C_6-*crypto*-AcpP. ^{15}N-CEST profiles at two CW RF fields (14.4 and 28.6 Hz) of 1 mM ^{15}N-labeled C_6-*crypto*-AcpP recorded at a static spectrometer field of 800 MHz. The experimental data (blue, 28.6 Hz; red, 14.4 Hz) are shown as circles and the solid lines represent the global best-fits to a two-state model (see Fig. 4.3E, main text). The data was collected at 37 °C in 50 mM potassium phosphate, pH 7.4, and 0.01% sodium azide.

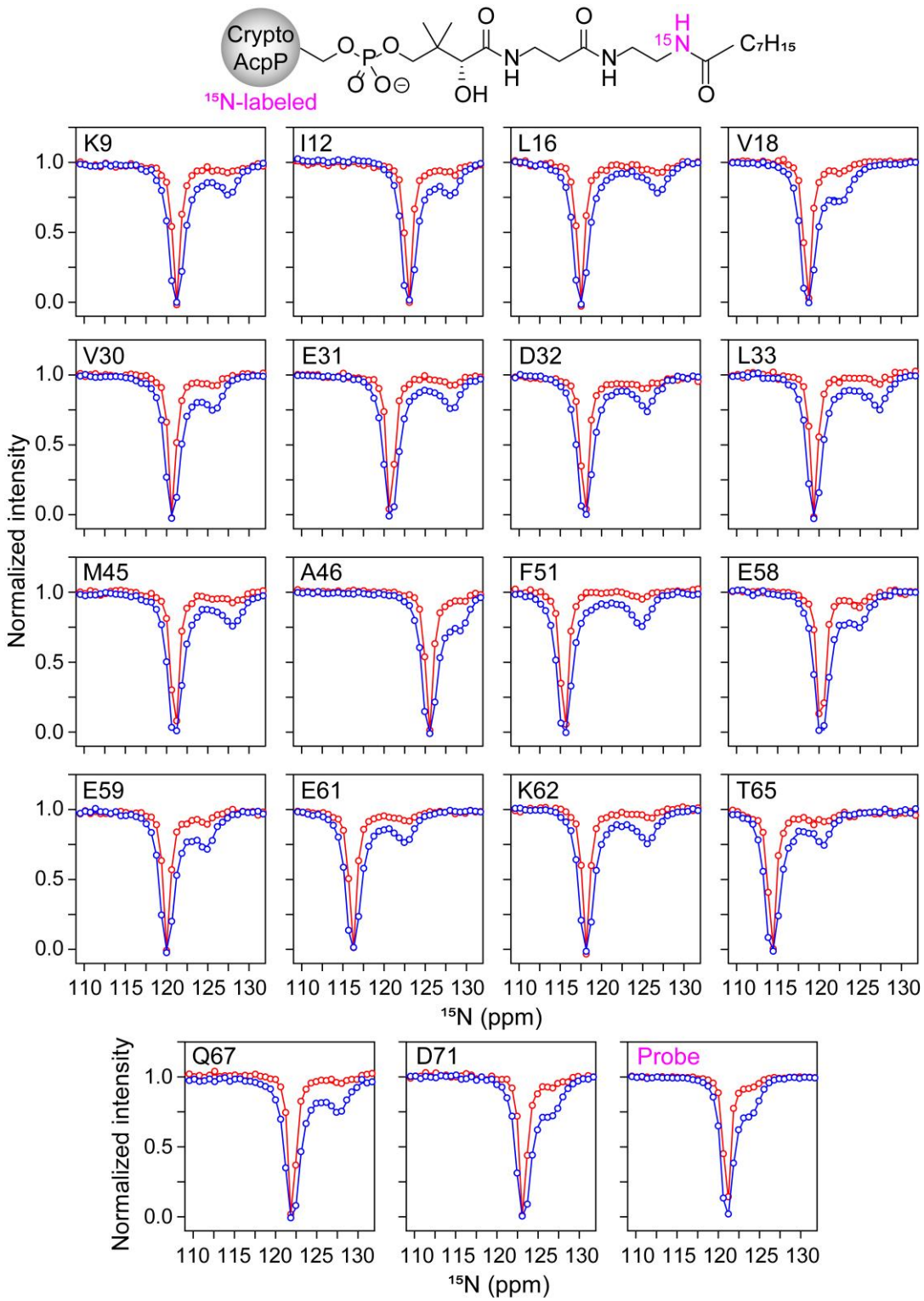


Figure S4.5. NMR analysis of ^{15}N-labeled C_8-*crypto*-AcpP. ^{15}N-CEST profiles at two CW RF fields (14.4 and 28.6 Hz) of 1 mM ^{15}N-labeled C_8-*crypto*-AcpP recorded at a static spectrometer field of 800 MHz. The experimental data (blue, 28.5 Hz; red, 14.4 Hz) are shown as circles and the solid lines represent the global best-fits to a two-state model (see Fig. 4.3E, main text). The data was collected at 37 °C in 50 mM potassium phosphate, pH 7.4, and 0.01% sodium azide.

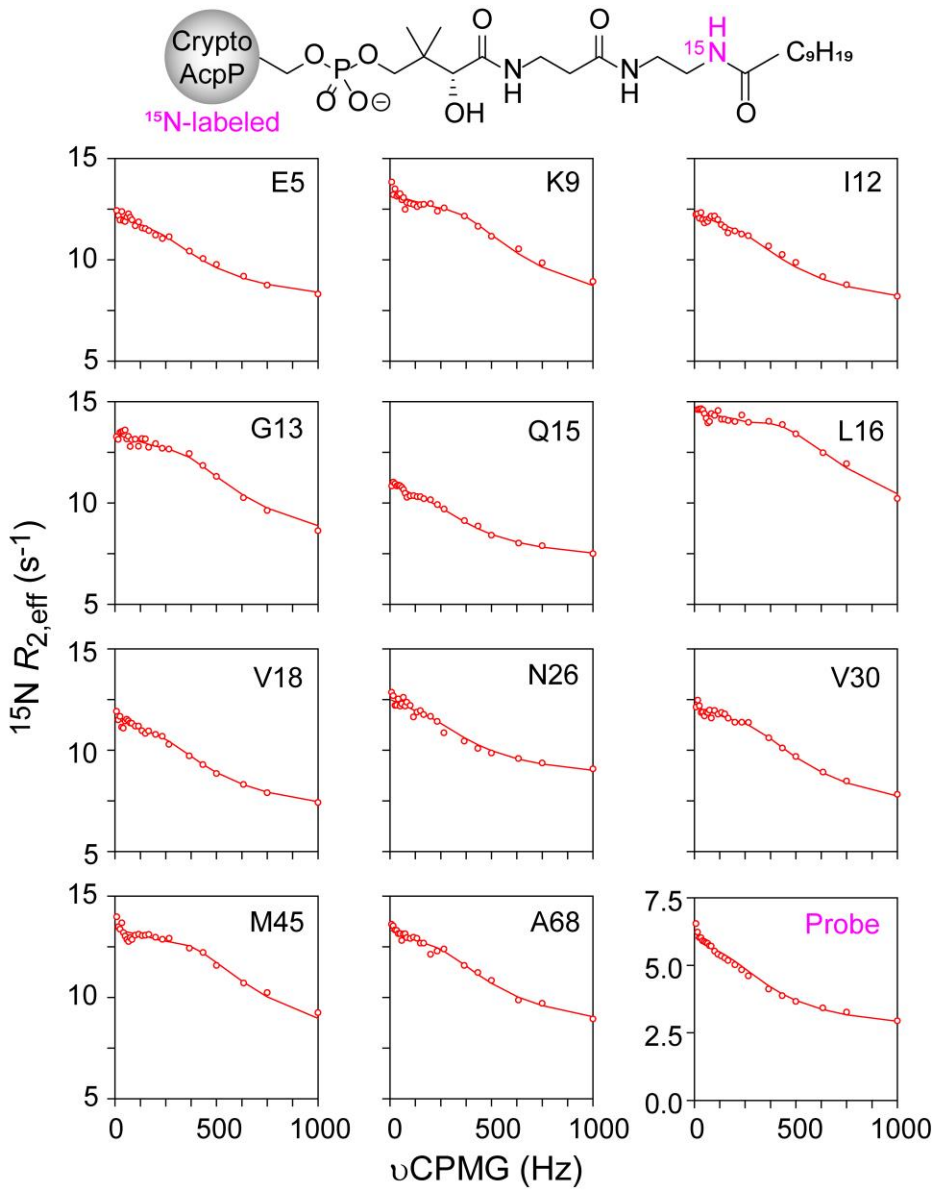


Figure S4.6. NMR analysis of ^{15}N-labeled C_{10}-*crypto*-AcpP. ^{15}N-CPMG profiles of 1 mM ^{15}N-labeled C_{10}-*crypto*-AcpP recorded at a static spectrometer field of 800 MHz. The experimental data are shown as circles and the solid lines represent the global best-fits to a two-state model (see Fig. 4.3E, main text). The data was collected at 37 °C in 50 mM potassium phosphate, pH 7.4, and 0.01% sodium azide.

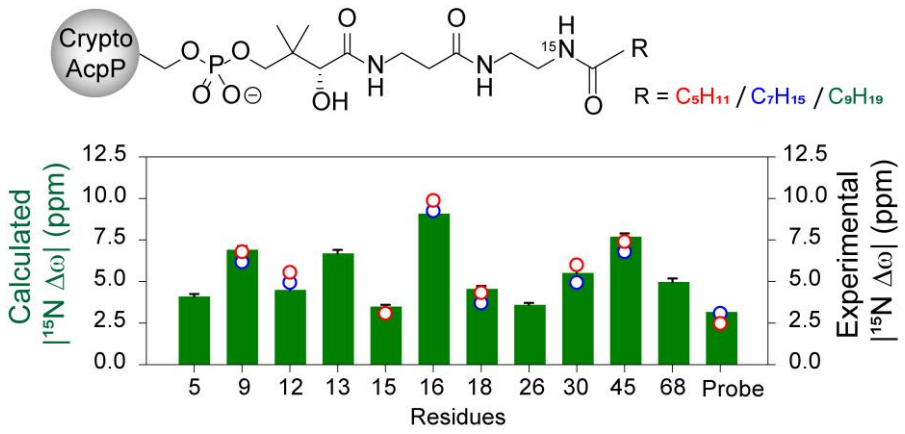


Figure S4.7. Comparison of ^{15}N-CPMG chemical shift predictions against measured chemical shifts from ^{15}N-CEST datasets. Absolute values of ^{15}N-Δω of C10-*crypto*-AcpP (green vertical bars) obtained from a global fit of ^{15}N-CPMG dispersions to a two-state exchange model (A ↔ B), Δω is the difference between chemical shifts of minor state **B** and the (observable) state **A**. Absolute values of ^{15}N-Δω of C_6/C_8-*crypto*-AcpP (red and blue circles, respectively) measured from the corresponding ^{15}N-CEST datasets.

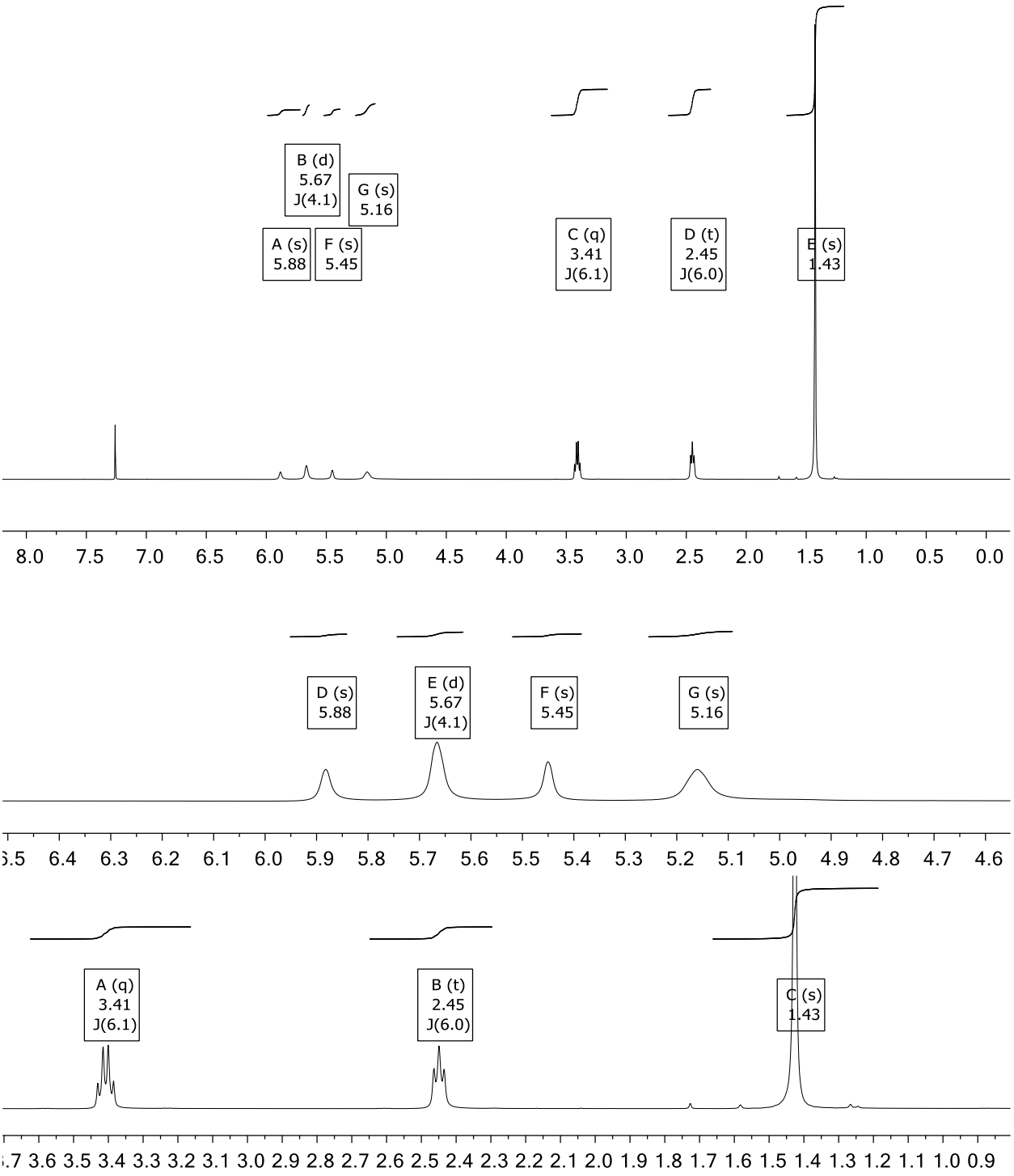
Table S4.1. List of samples used for NMR measurements and the corresponding subsets of residues included in the analysis of NMR data.[a]

Sample	NMR experiment	Residues
^{15}N-labeled *crypto*-AcpP tethered to selectively ^{15}N-labeled C_6-pantetheinamide [b]		
	^{15}N-CEST[c]	9, 11, 12, 15, 16, 18, 19, 28–33, 43–46, 50, 51, 58, 59, 61, 62, 65, 67, 71, probe ^{15}N
^{15}N-labeled *crypto*-AcpP tethered to selectively ^{15}N-labeled C_8-pantetheinamide [b]		
	^{15}N-CEST[c]	9, 12, 16, 18, 30–33, 45, 46, 51, 58, 59, 61, 62, 65, 67, 71, probe ^{15}N
^{15}N-labeled *crypto*-AcpP tethered to selectively ^{15}N-labeled C_{10}-pantetheinamide [b, d]		
	^{15}N-CPMG[e]	5, 9, 12, 13, 15, 16, 18, 26, 30, 45, 68, probe ^{15}N

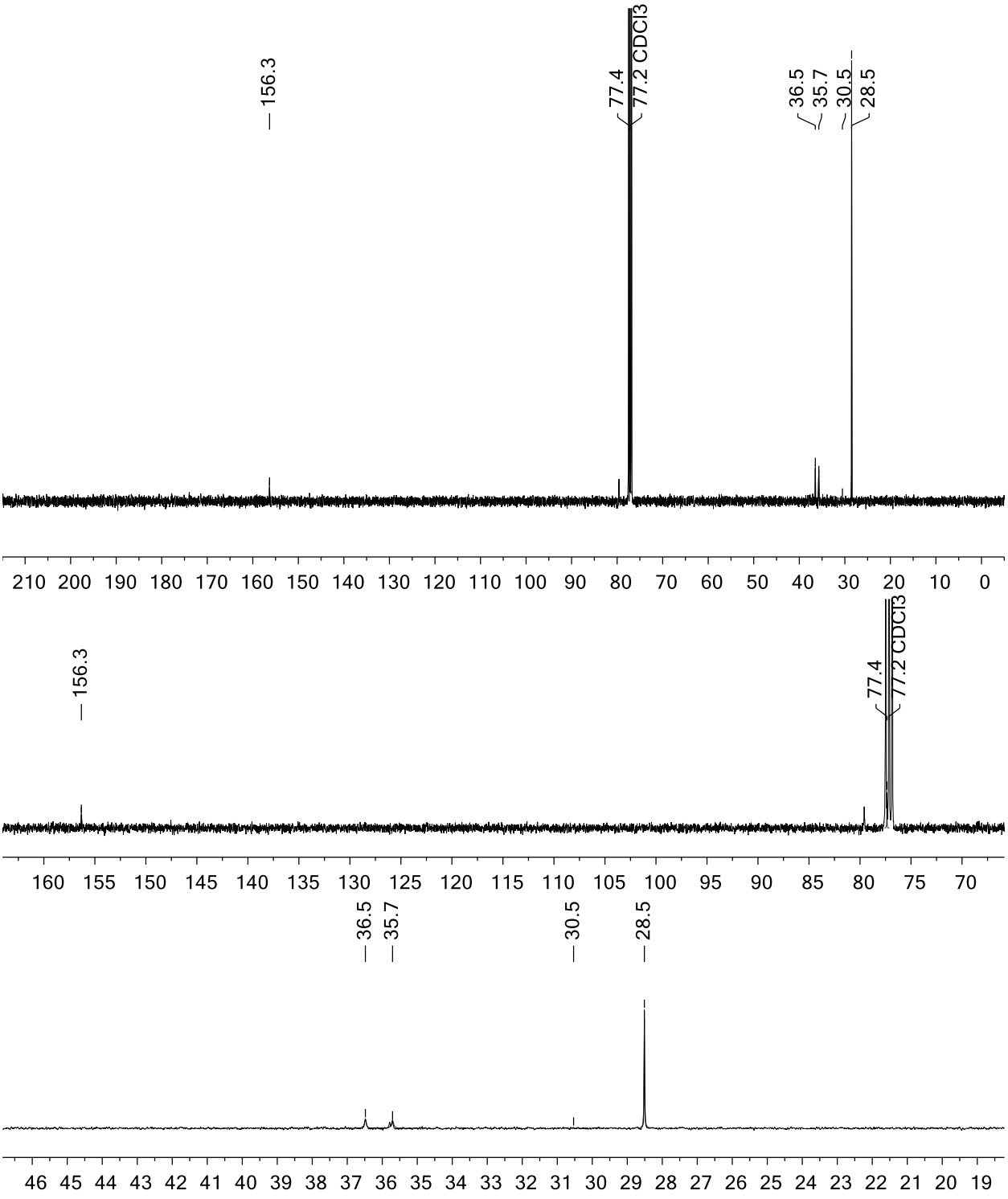
[a] Data were fit globally to a two-state exchange model (Scheme 4.4) with rate constants reported in Fig. 4.3F and Table 4.1 of the main text.
[b]The pantetheinamides with selective ^{15}N-labeling were prepared using Schemes S4.1–S4.3. Individual probes were conjugated to ^{15}N-labeled *apo*-ACP using the one-pot chemoenzymatic method to generate *crypto*-AcpP.
[c] ^{15}N-CEST experiments were recorded with T_{CEST} = 400 ms using CW saturation RF field strengths of 14.4 and 28.6 Hz (^{1}H spectrometer frequency = 800 MHz).
[d] In the case of C_{10}-*crypto*-AcpP, no characteristic second dips in intensity at the resonance frequencies of the minor species were observed (cf. Fig. 4.2C and 4.3B, main text), indicating that the timescale of exchange is not accessible to ^{15}N-CEST measurements.
[e] ^{15}N-CPMG relaxation dispersion experiments were recorded at a ^{1}H spectrometer frequency of 800 MHz.

NMR Spectra

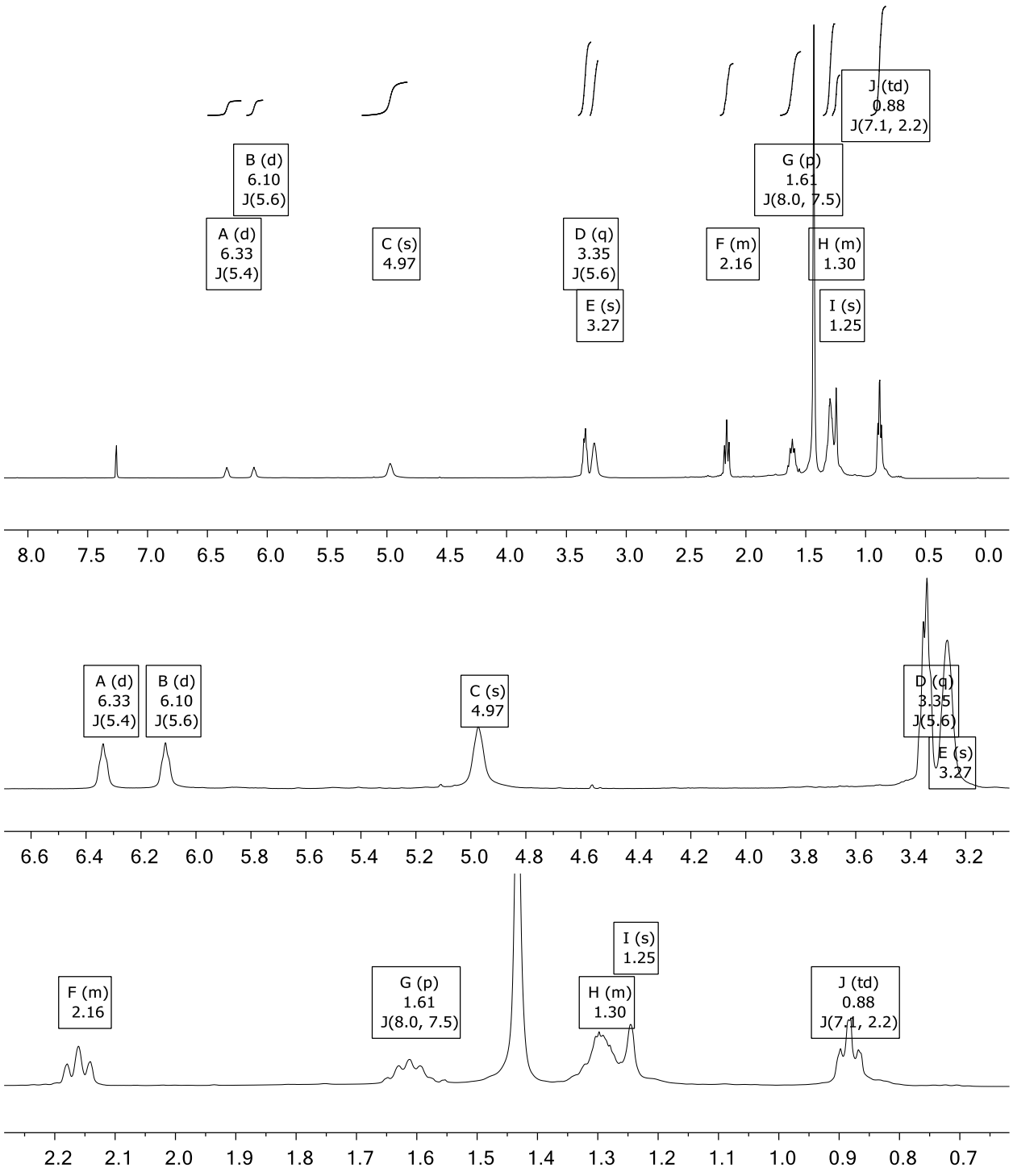
^{1}H-NMR (500 MHz) spectrum of **S1** in $CDCl_3$



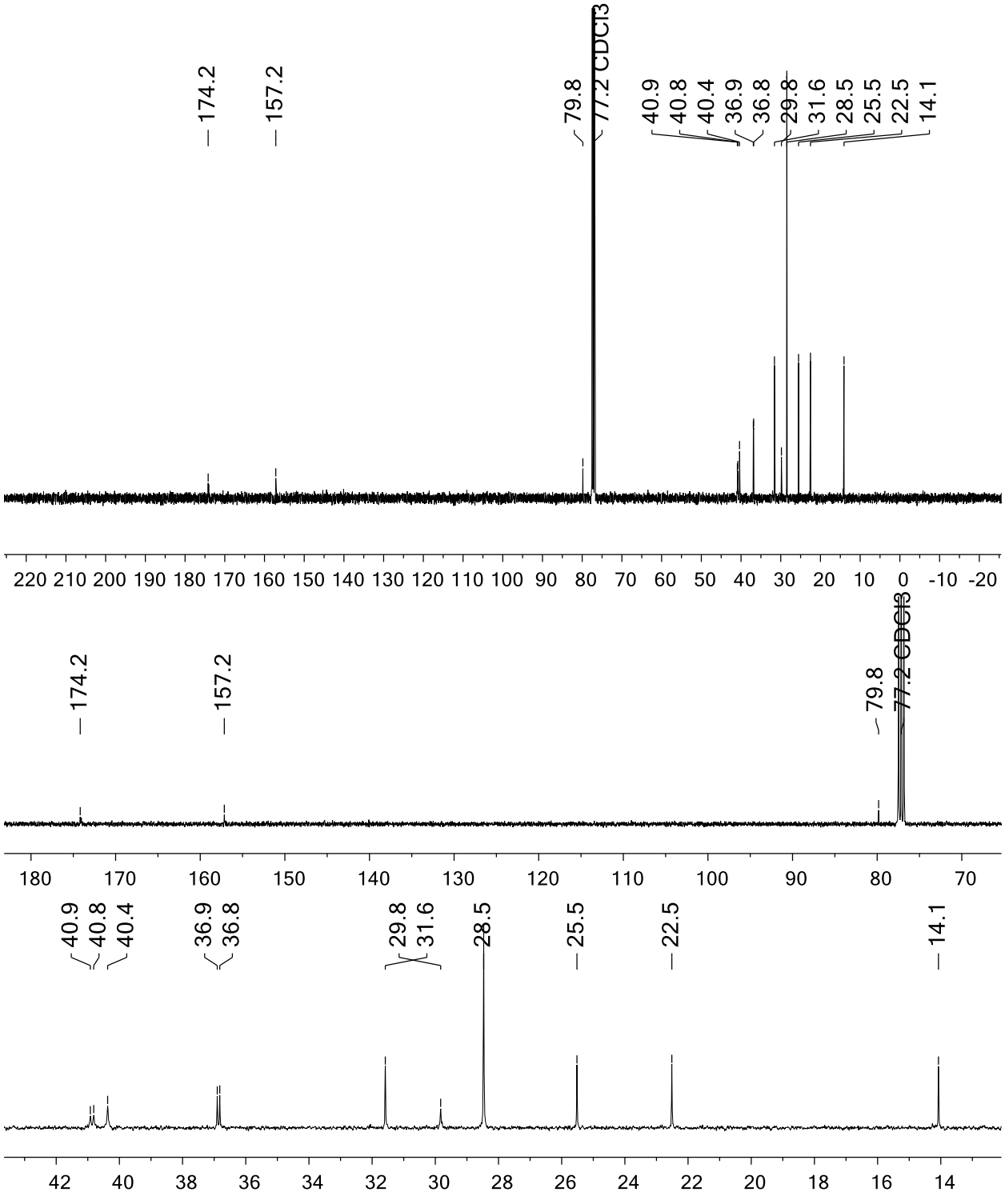
^{13}C-NMR (125 MHz) spectrum of **S1** in $CDCl_3$



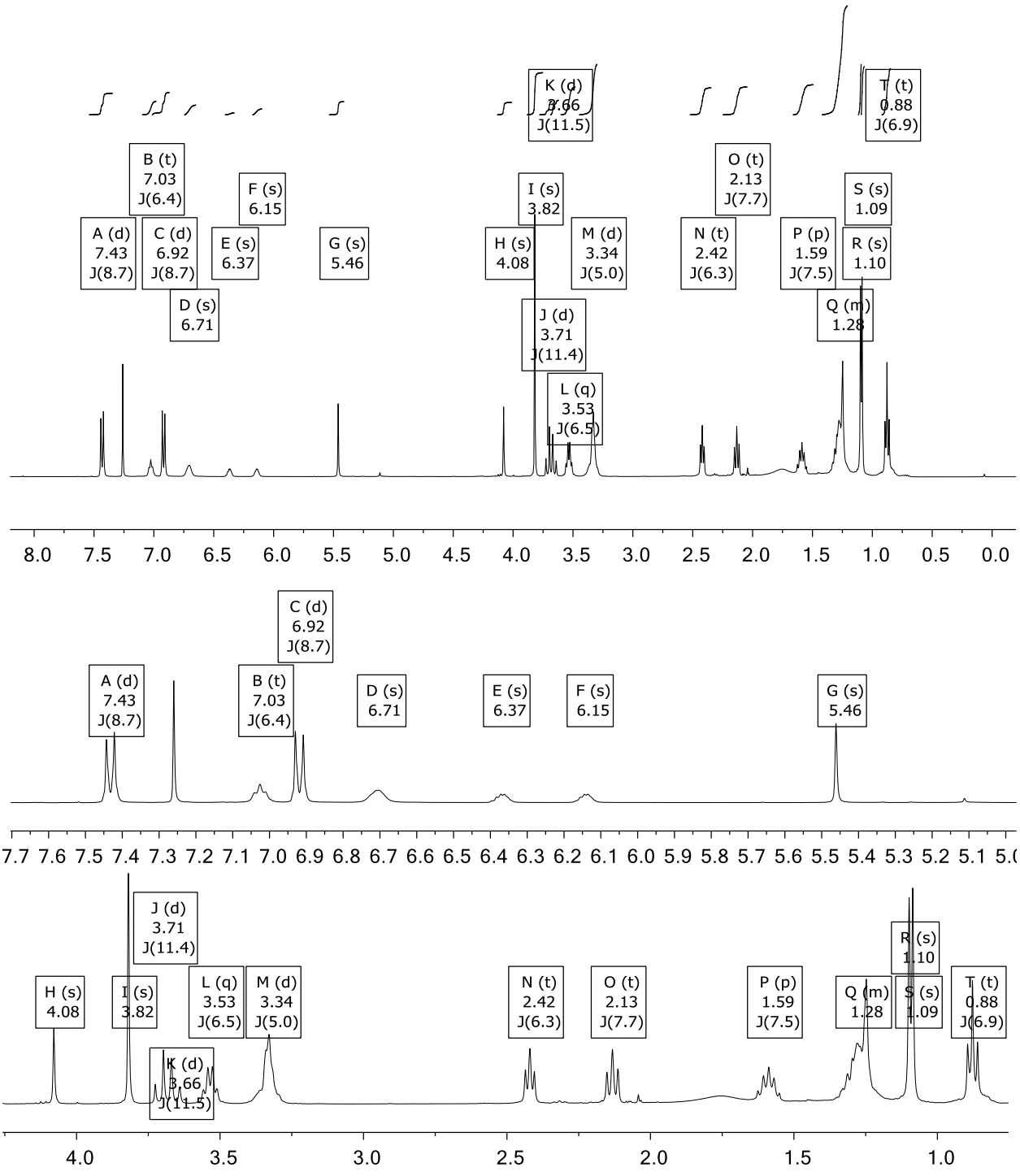
^{1}H-NMR (500 MHz) spectrum of **S2a** in $CDCl_3$



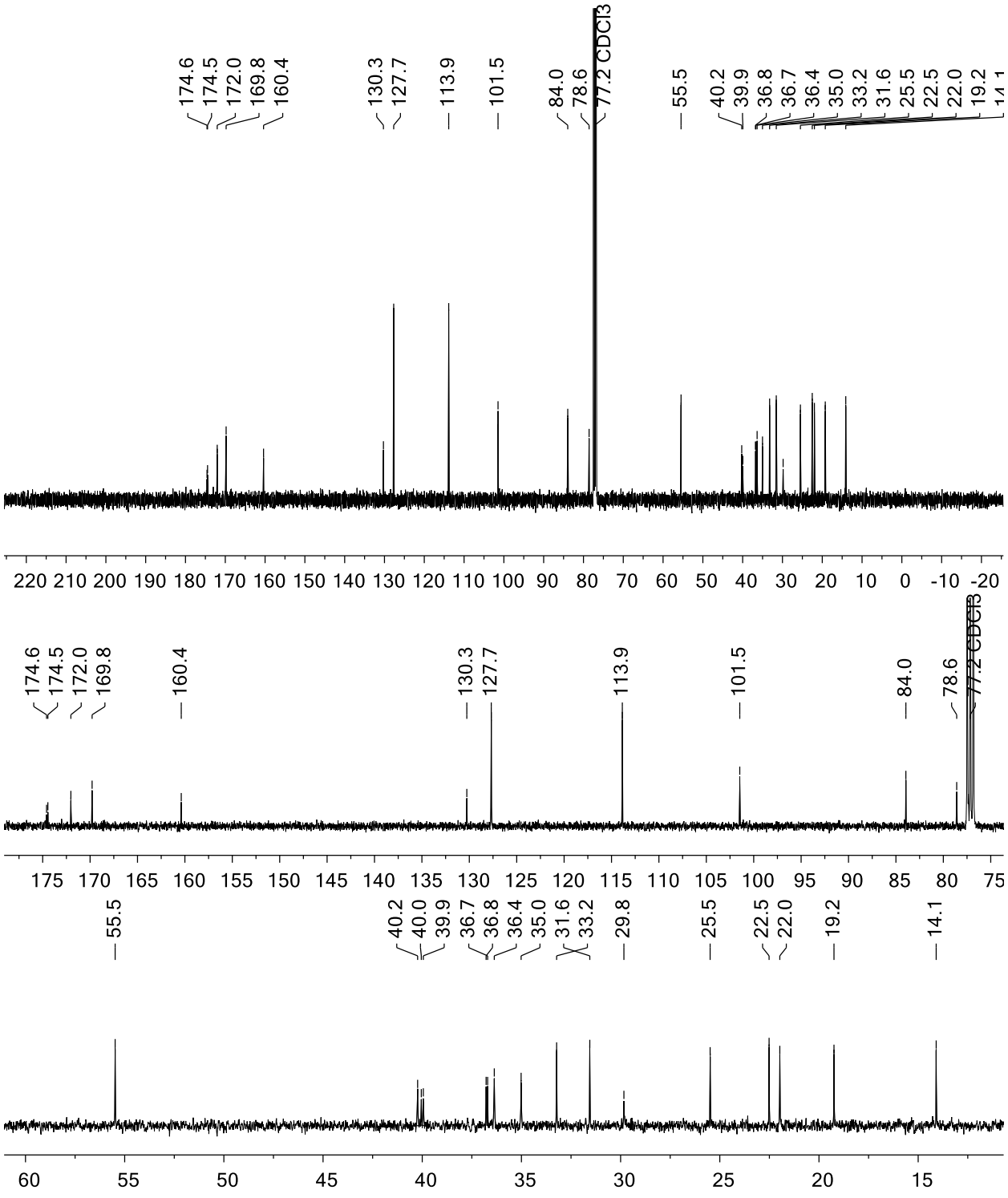
^{13}C-NMR (125 MHz) spectrum of **S2a** in $CDCl_3$



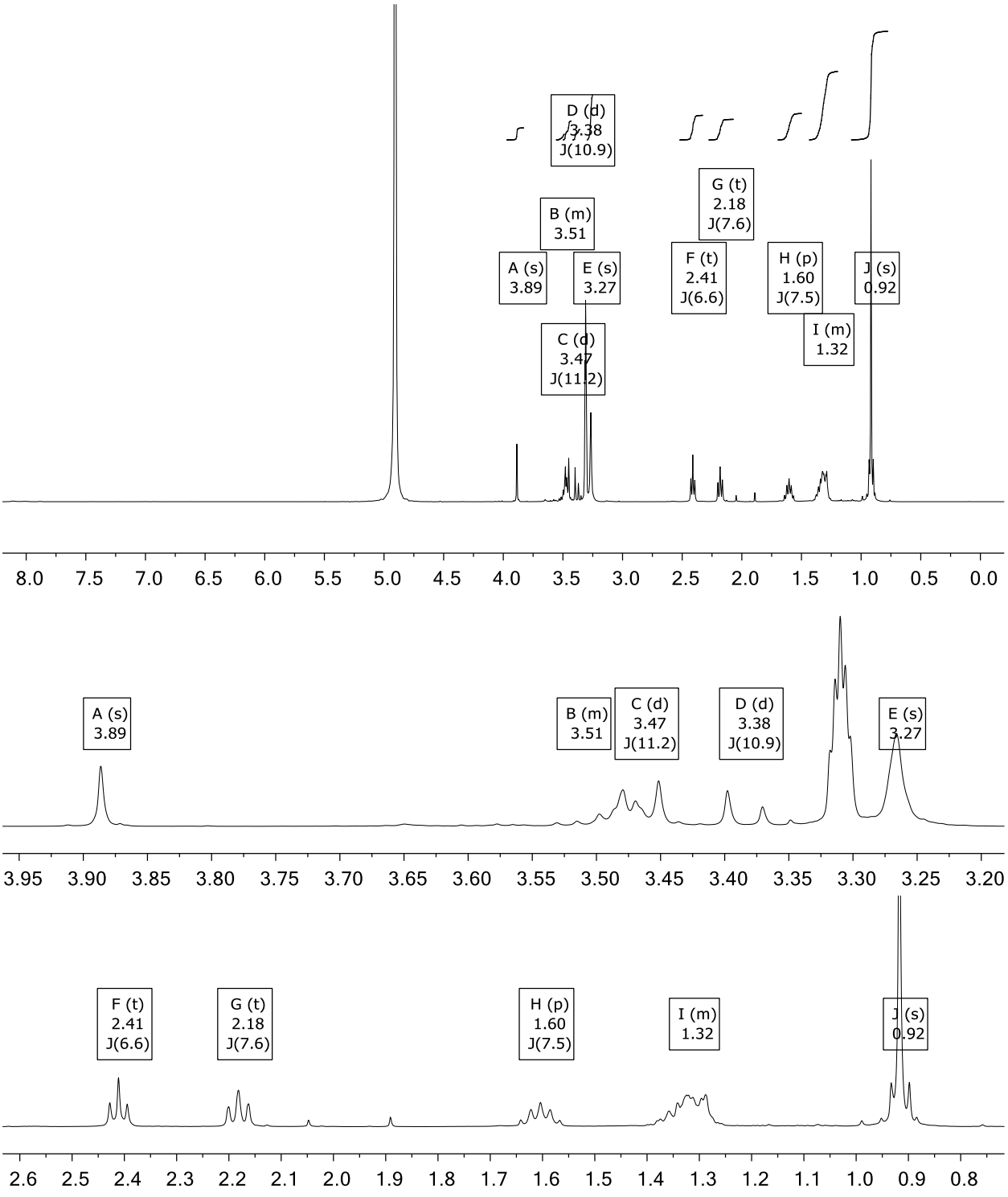
^{1}H-NMR (500 MHz) spectrum of **S5a** in $CDCl_3$



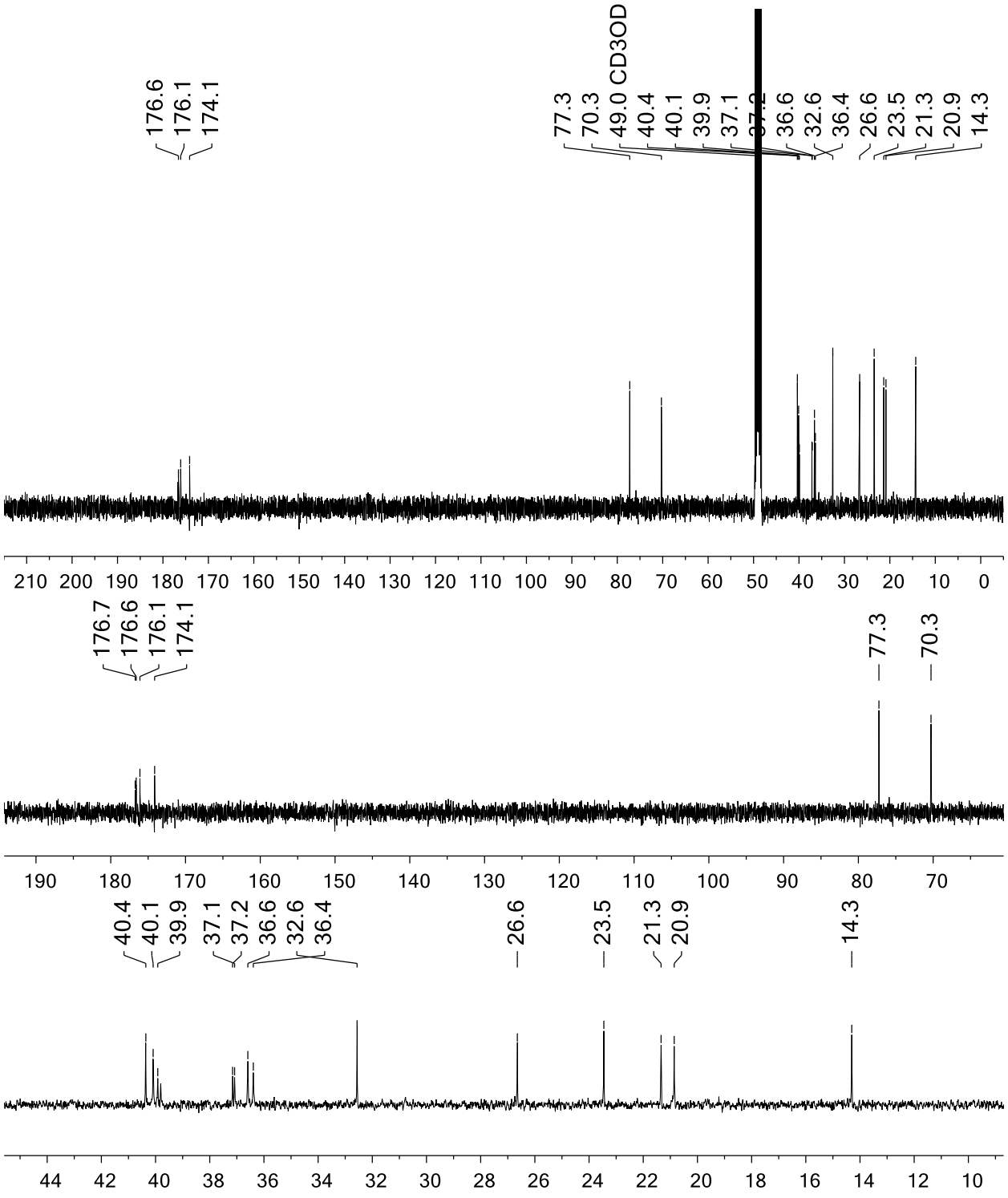
^{13}C-NMR (125 MHz) spectrum of **S5a** in $CDCl_3$



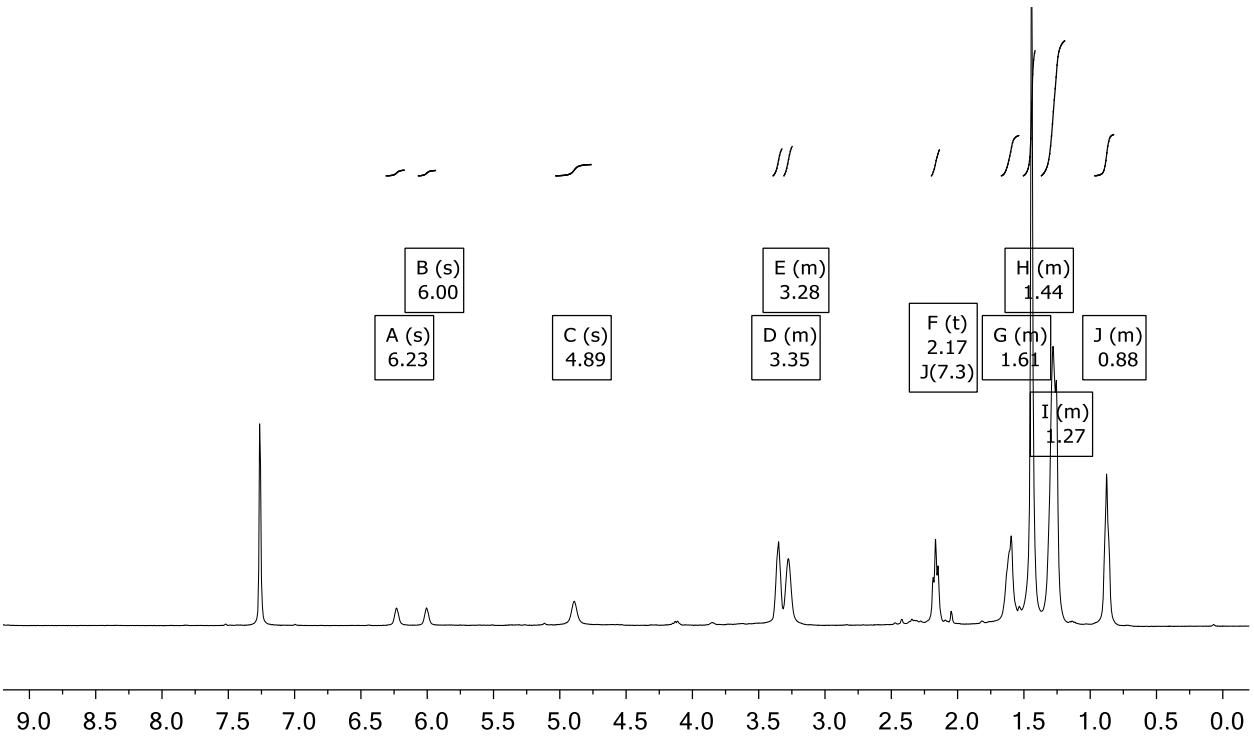
^{1}H-NMR (500 MHz) spectrum of **1a** in CD_3OD

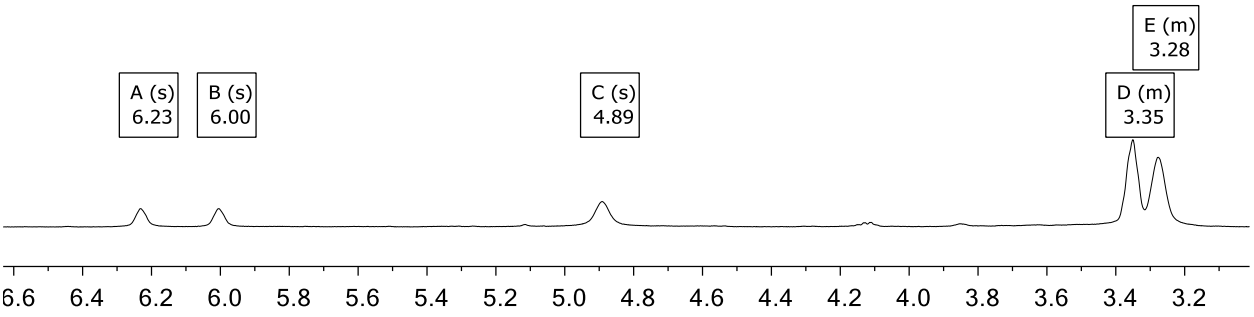


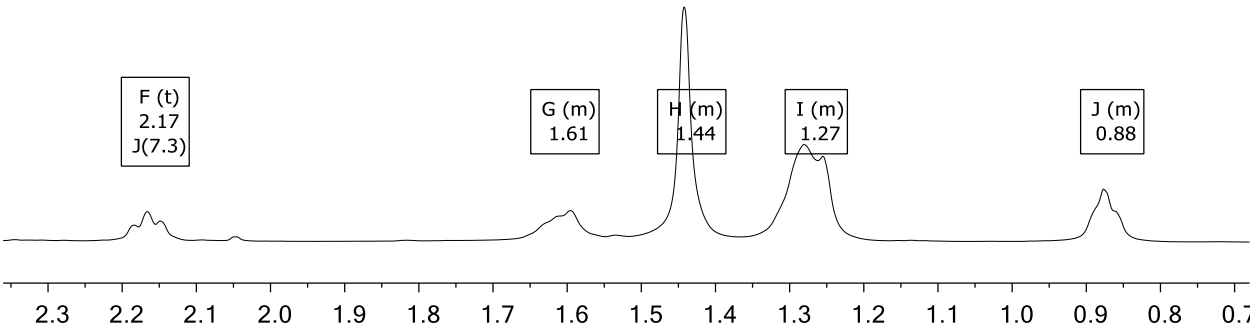
^{13}C-NMR (125 MHz) spectrum of **1a** in CD_3OD



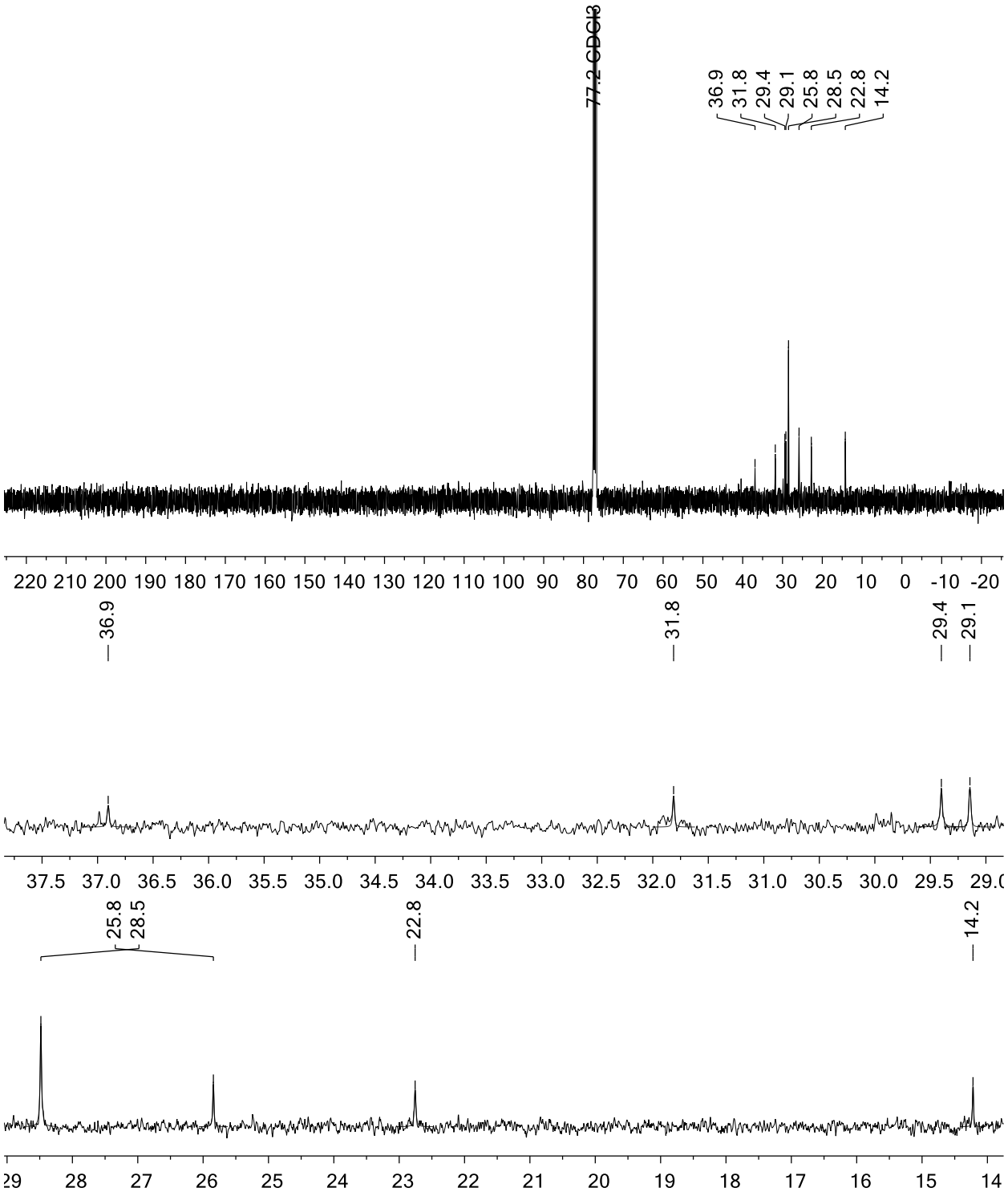
^{1}H-NMR (500 MHz) spectrum of **S2b** in $CDCl_3$

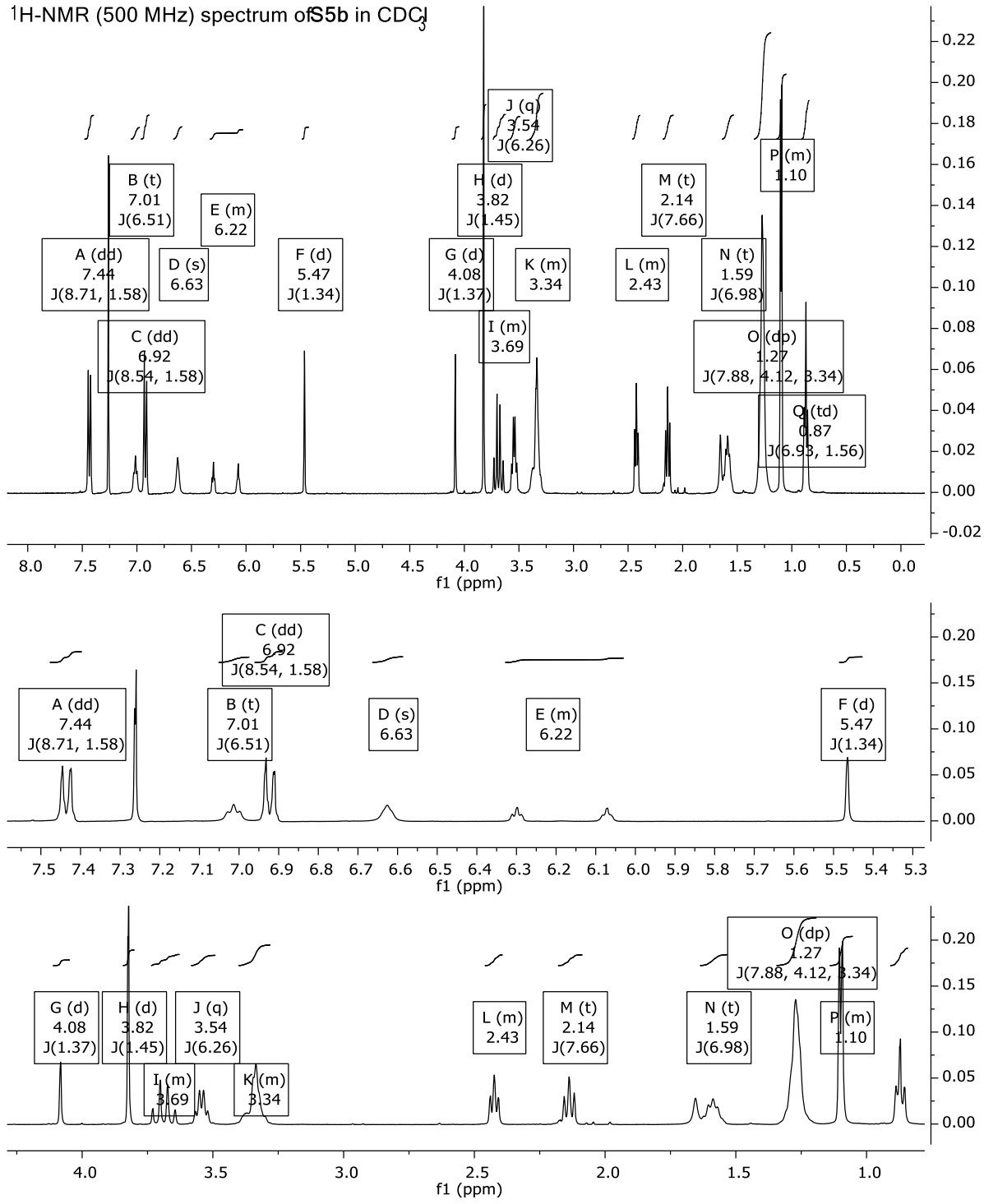




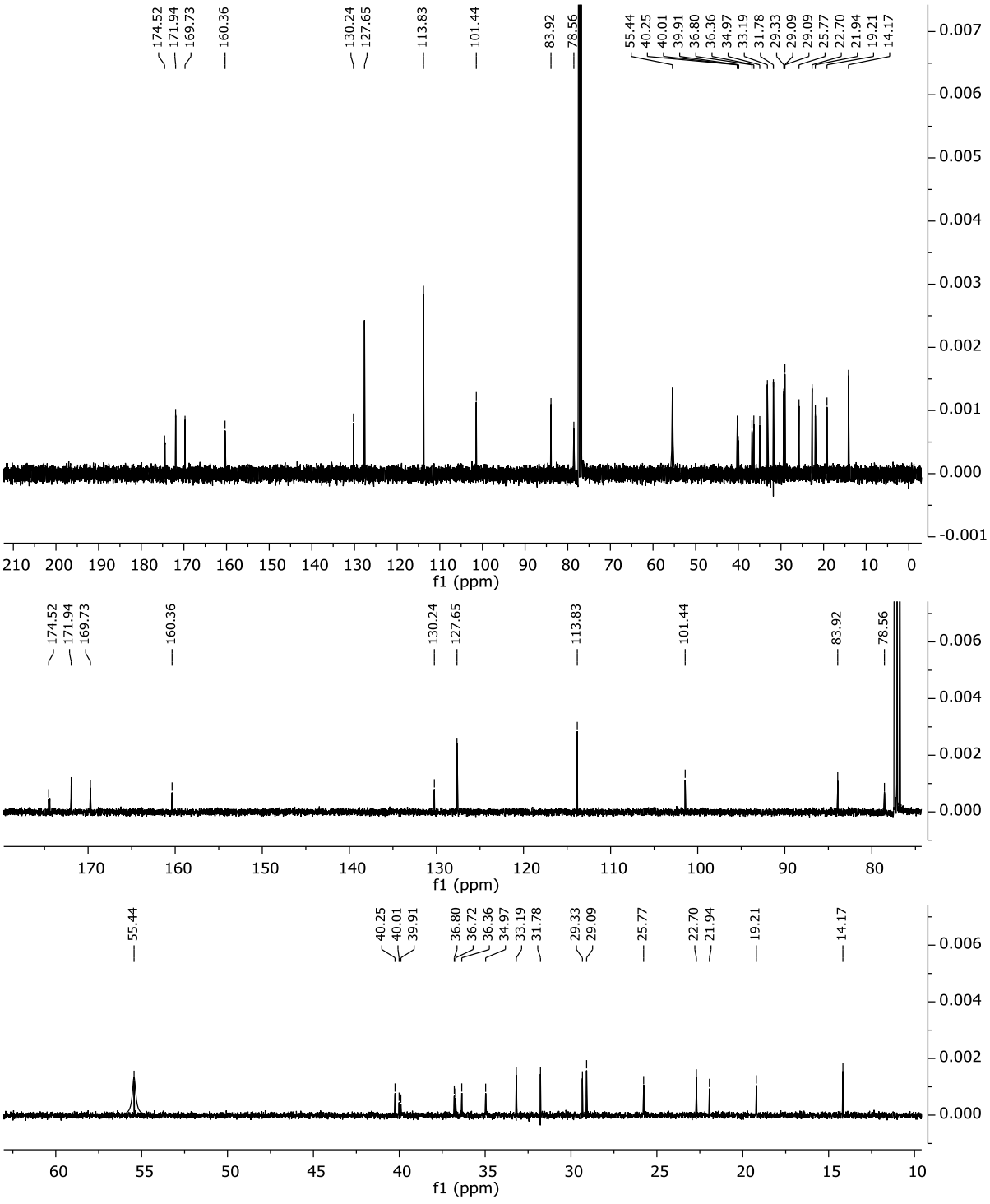


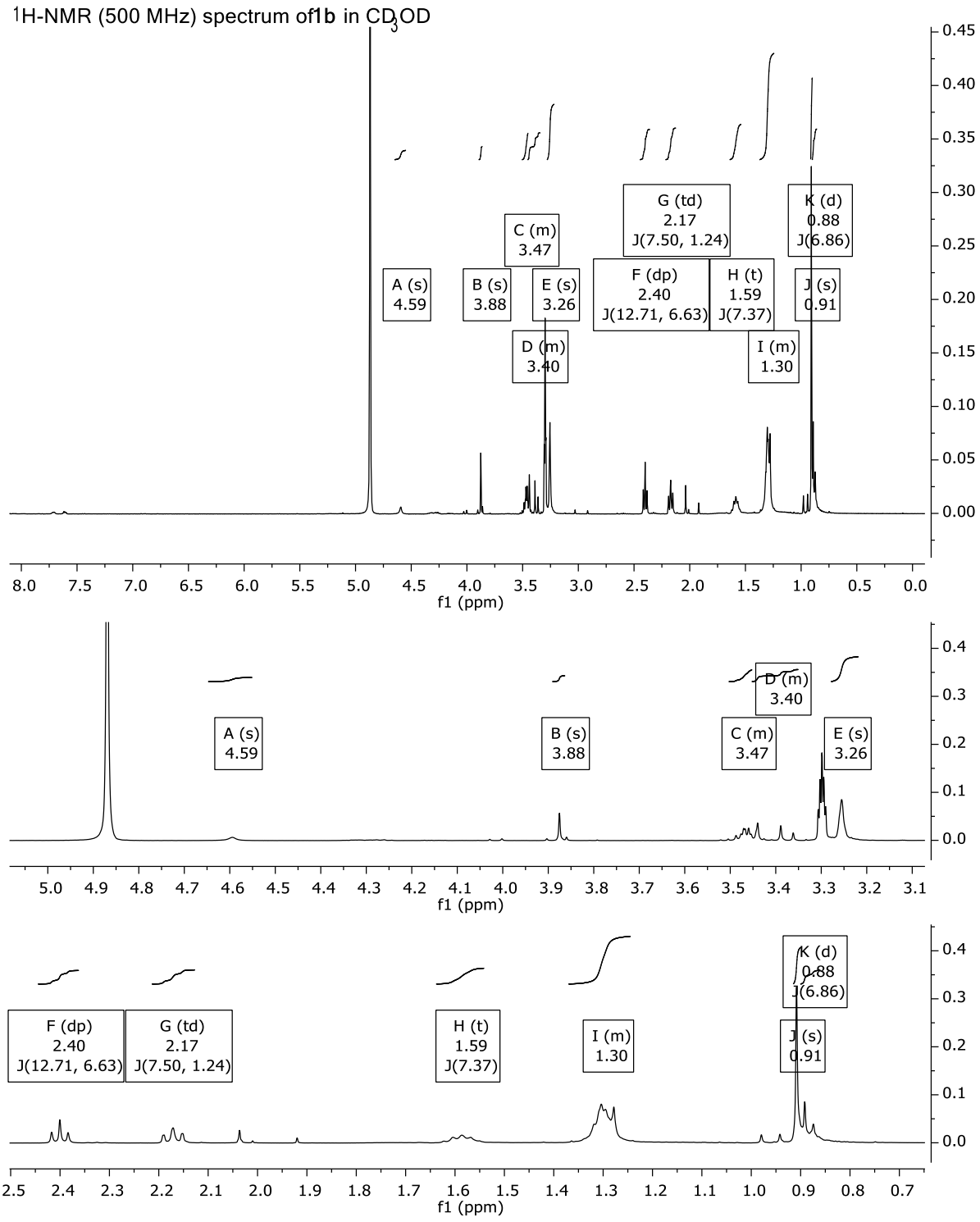
^{13}C-NMR (125 MHz) spectrum of **S2b** in $CDCl_3$



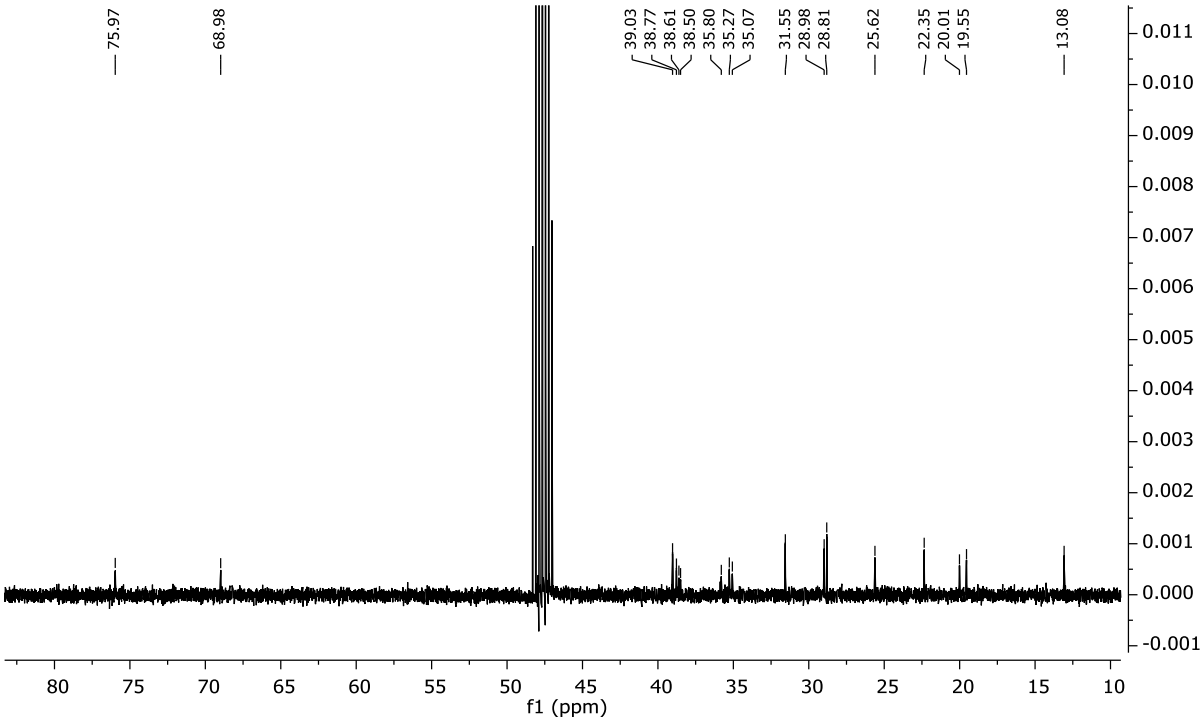

1H-NMR (500 MHz) spectrum of S5b in CDCl3
A (dd) 7.44 J(8.71, 1.58)
B (t) 7.01 J(6.51)
C (dd) 6.92 J(8.54, 1.58)
D (s) 6.63
E (m) 6.22
F (d) 5.47 J(1.34)
G (d) 4.08 J(1.37)
H (d) 3.82 J(1.45)
I (m) 3.69
J (q) 3.54 J(6.26)
K (m) 3.34
L (m) 2.43
M (t) 2.14 J(7.66)
N (t) 1.59 J(6.98)
O (dp) 1.27 J(7.88, 4.12, 3.34)
P (m) 1.10
Q (td) 0.87 J(6.93, 1.56)
f1 (ppm)

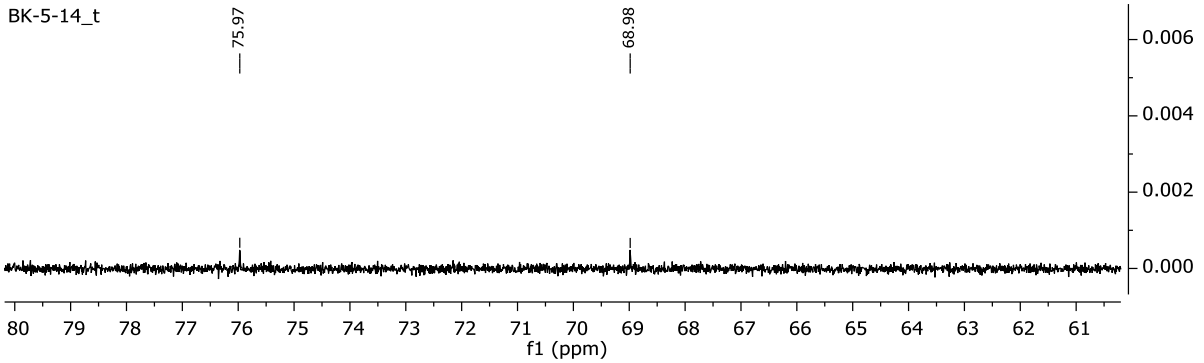
^{13}C-NMR (125 MHz) spectrum of **S5b** in $CDCl_3$

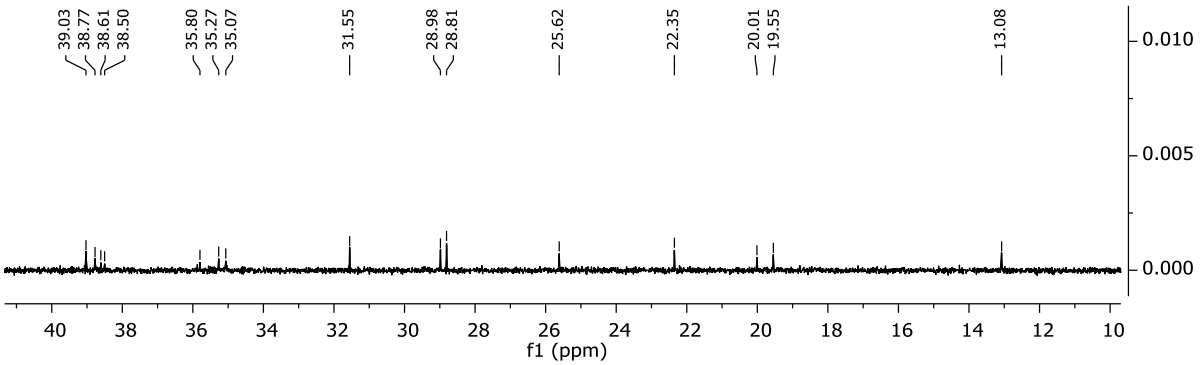



1H-NMR (500 MHz) spectrum of 1b in CD3OD
A (s) 4.59
B (s) 3.88
C (m) 3.47
D (m) 3.40
E (s) 3.26
F (dp) 2.40 J(12.71, 6.63)
G (td) 2.17 J(7.50, 1.24)
H (t) 1.59 J(7.37)
I (m) 1.30
J (s) 0.91
K (d) 0.88 J(6.86)
f1 (ppm)

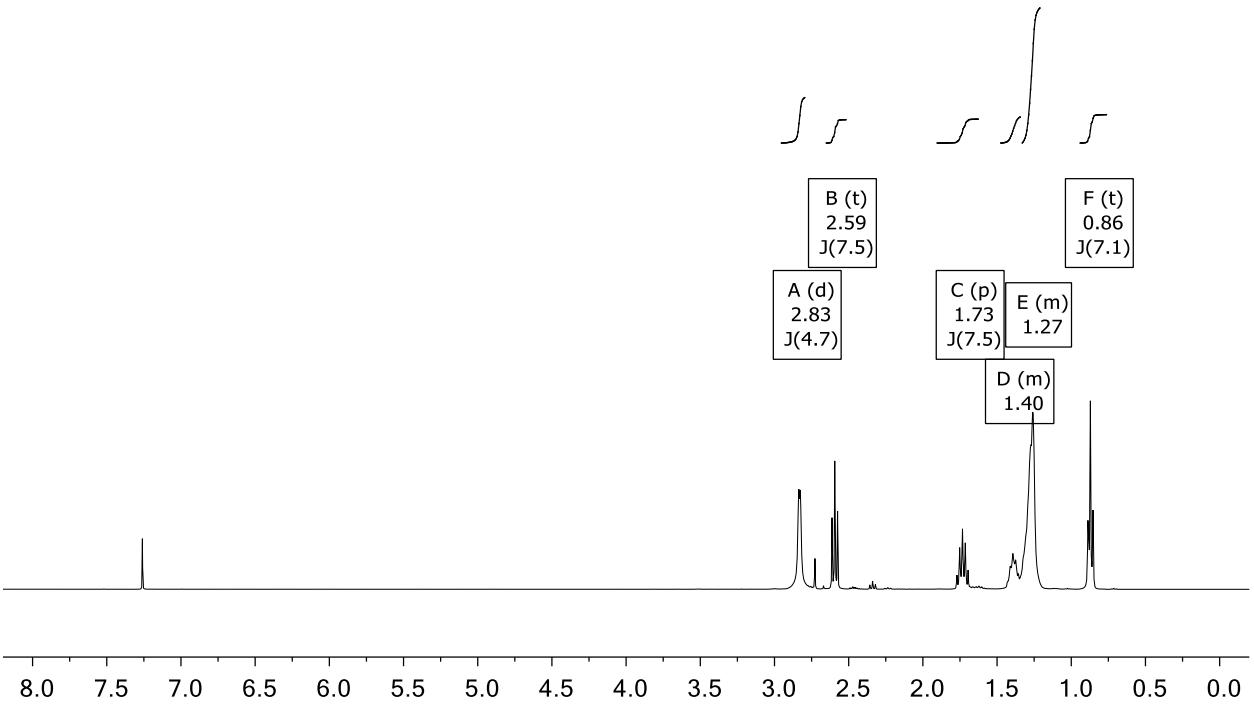
^{13}C-NMR (125 MHz) spectrum of **1b** in CD_3OD

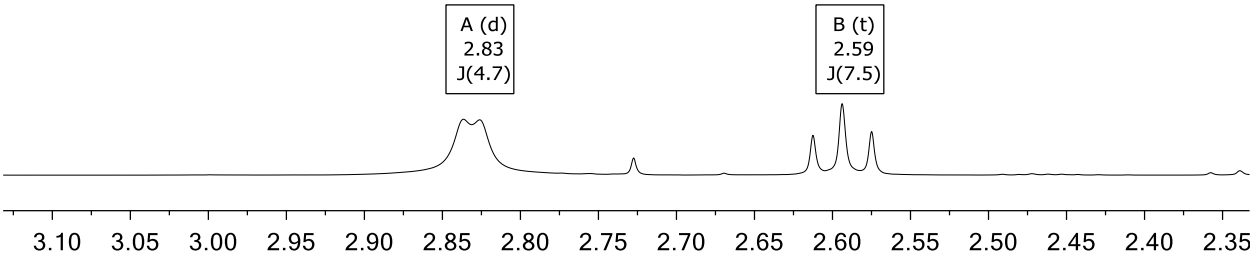


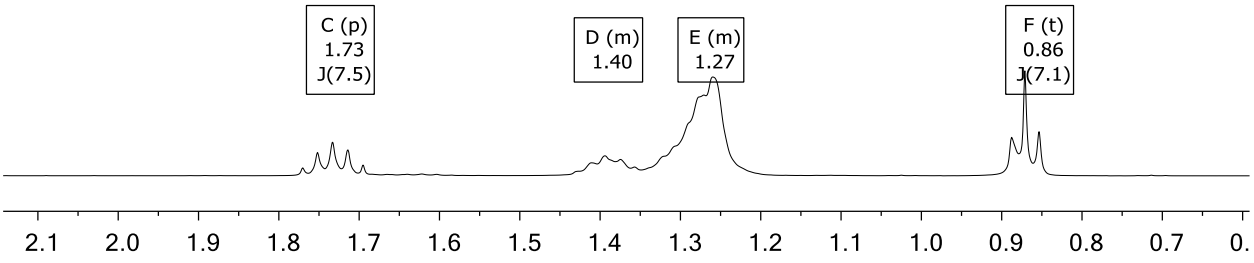




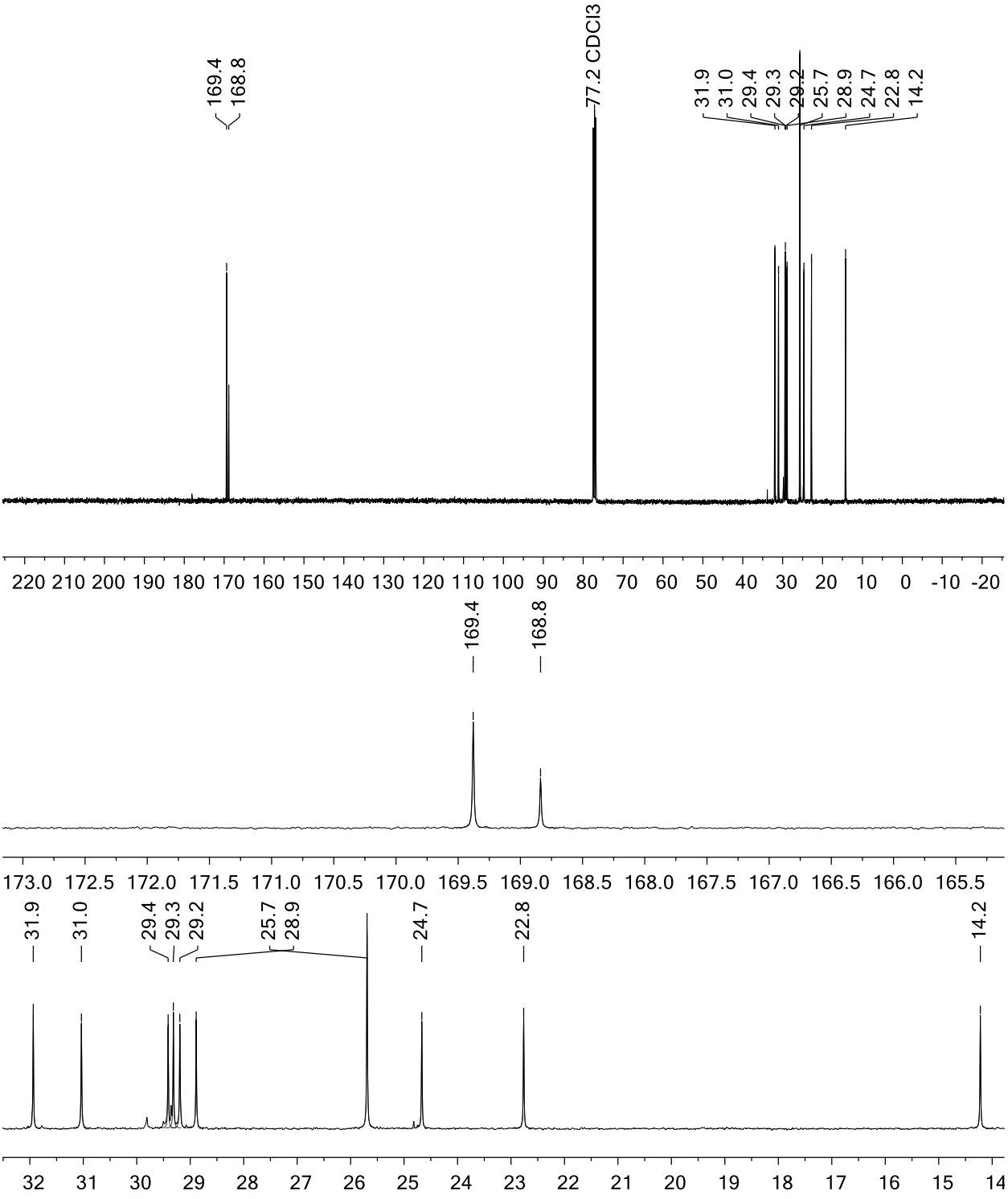
^{1}H-NMR (500 MHz) spectrum of **2,6-dioxopyrrolidin-1-yl decanoate** in $CDCl_3$



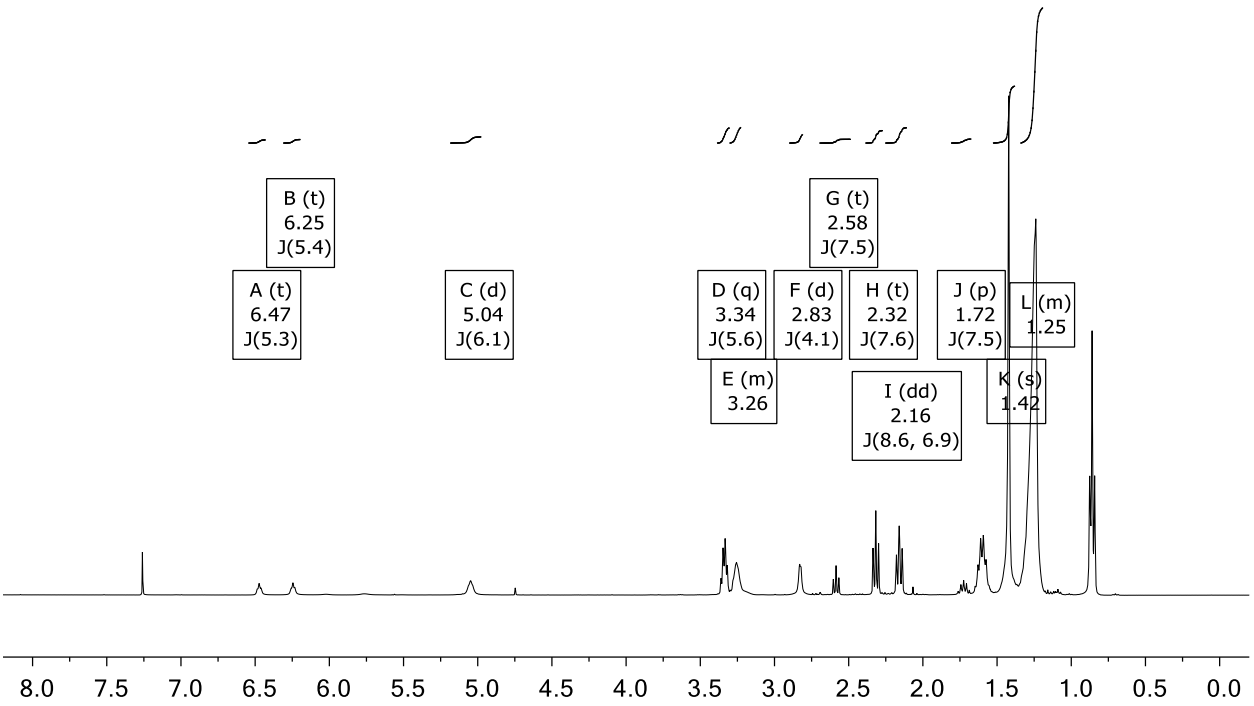


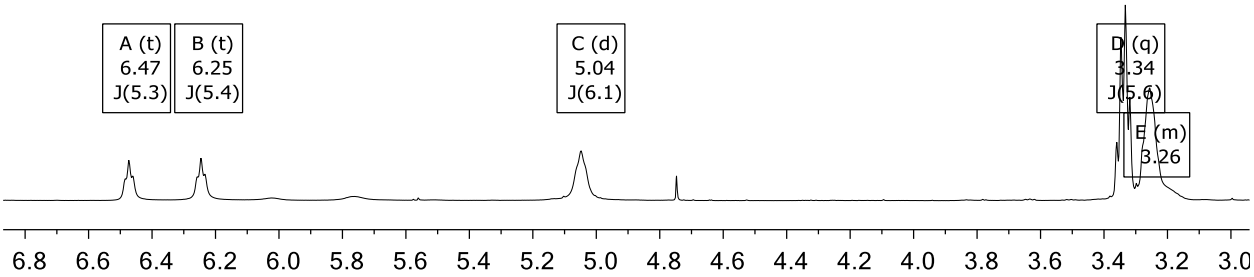


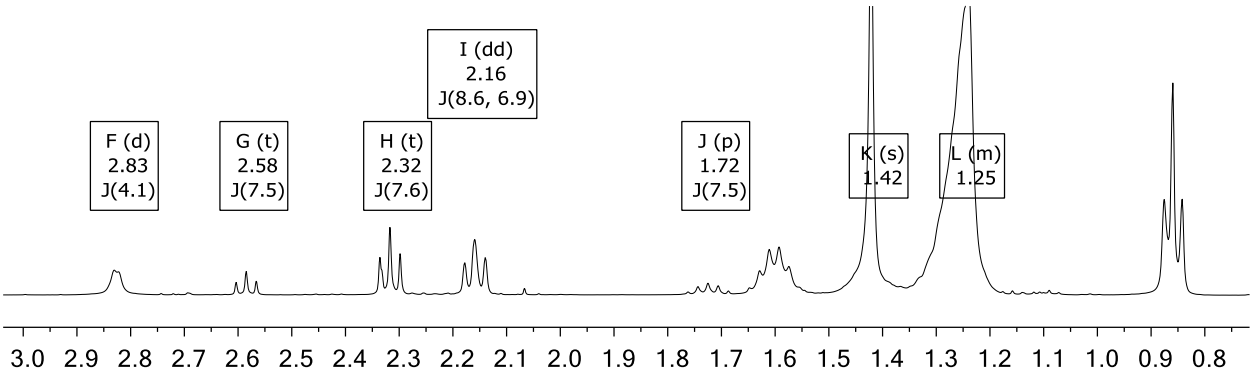
^{13}C-NMR (125 MHz) spectrum of **2,6-dioxopyrrolidin-1-yl decanoate** in $CDCl_3$



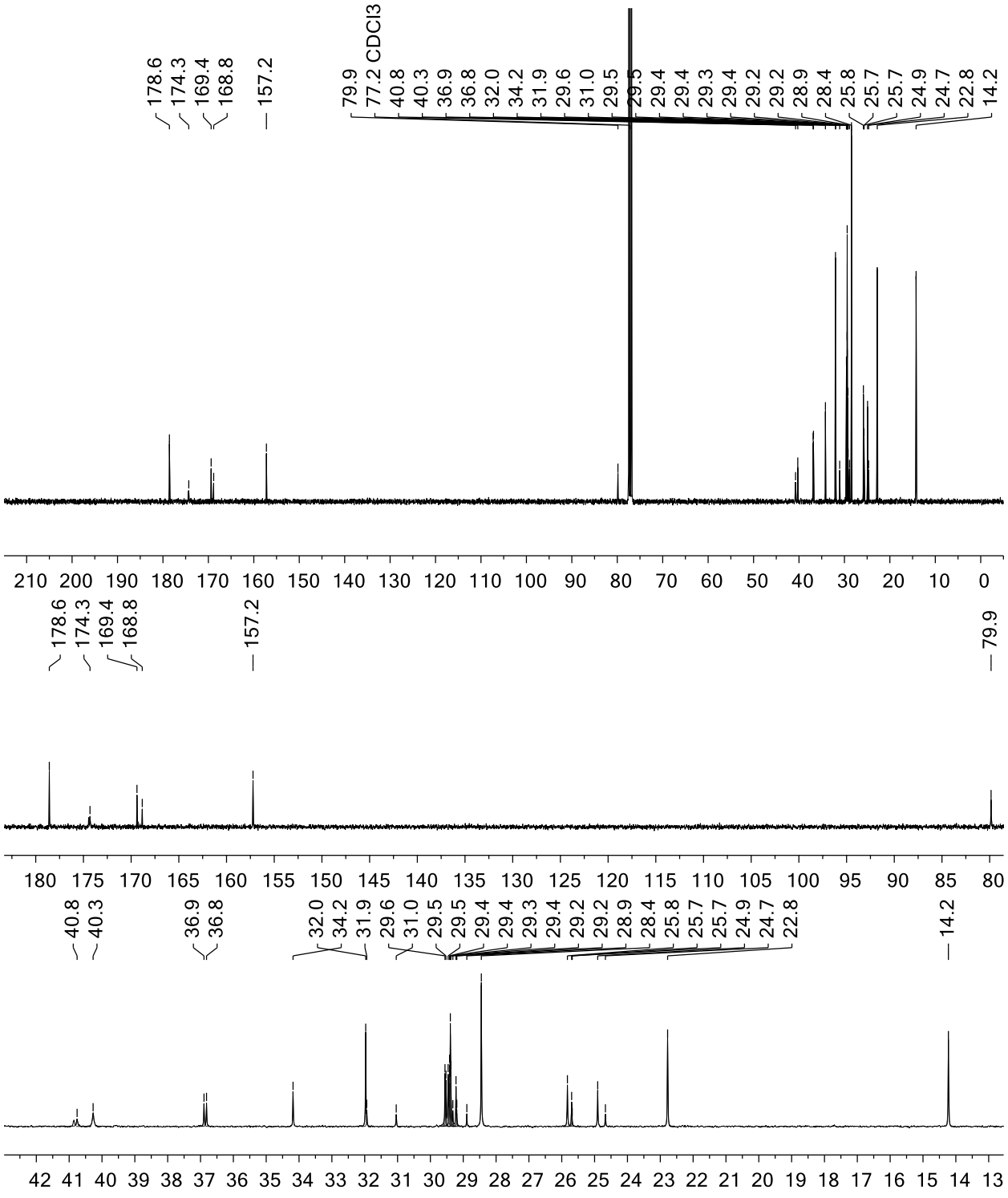
^{1}H-NMR (500 MHz) spectrum of **S2c** in $CDCl_3$

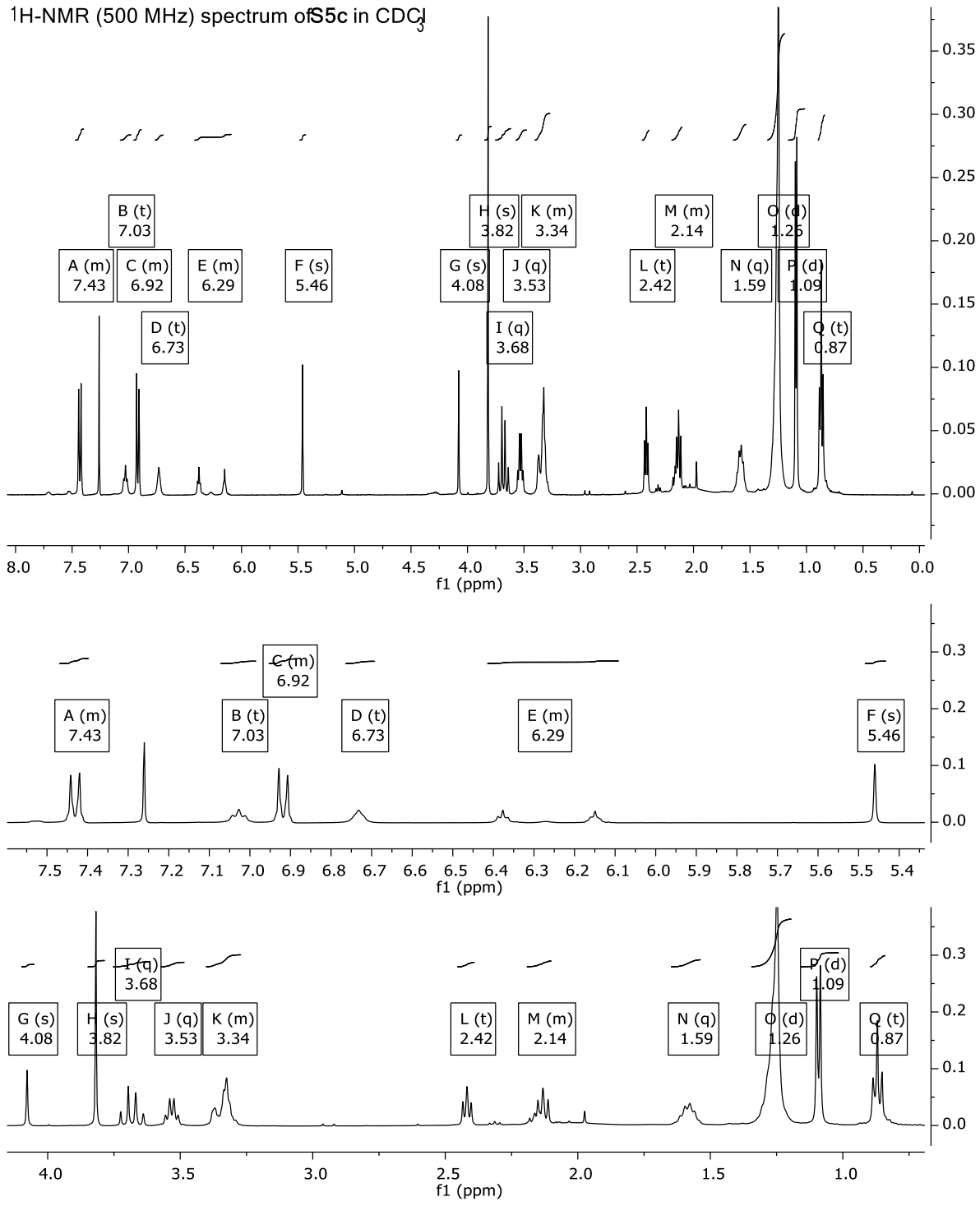




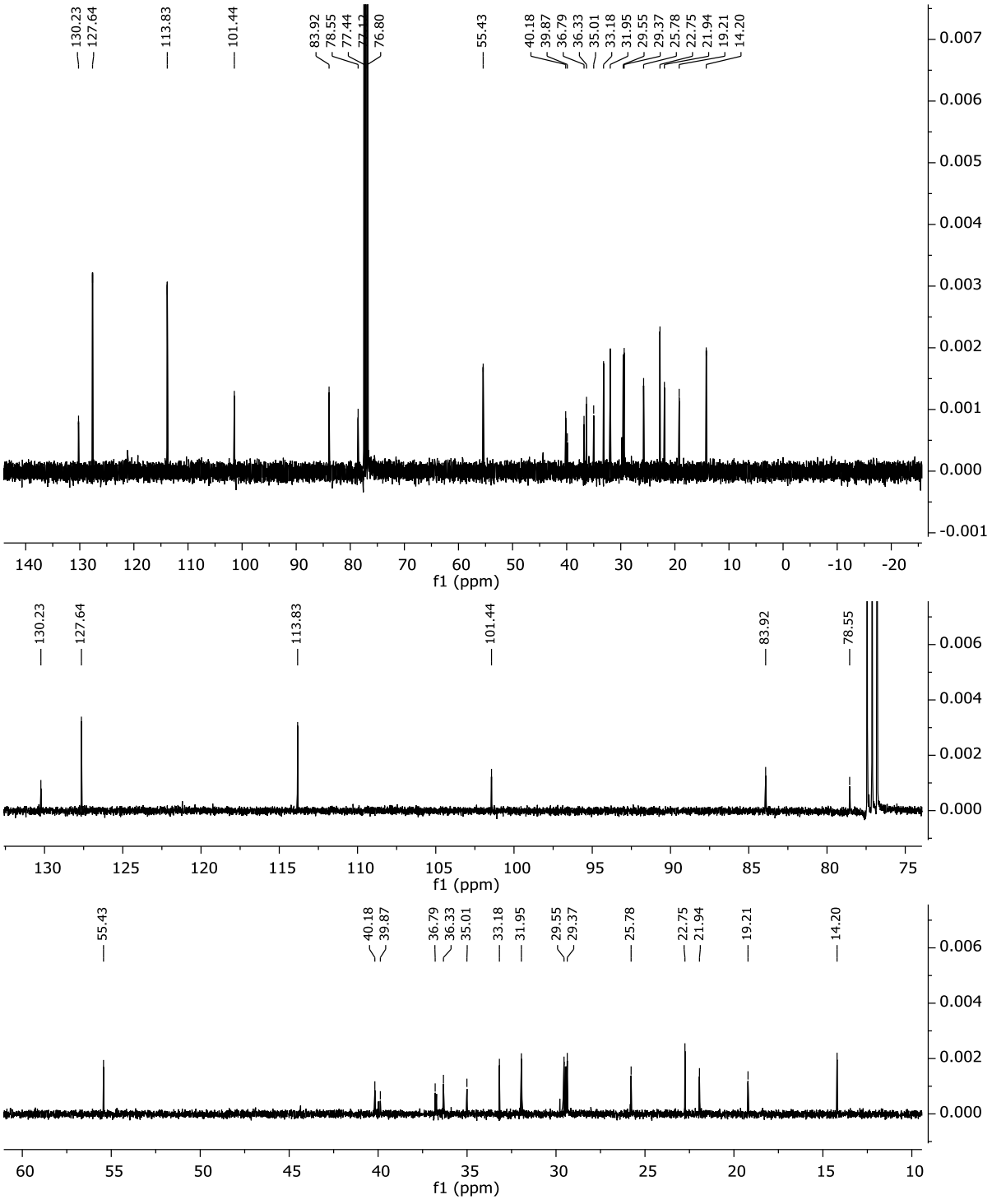


^{13}C-NMR (125 MHz) spectrum of **S2c** in $CDCl_3$

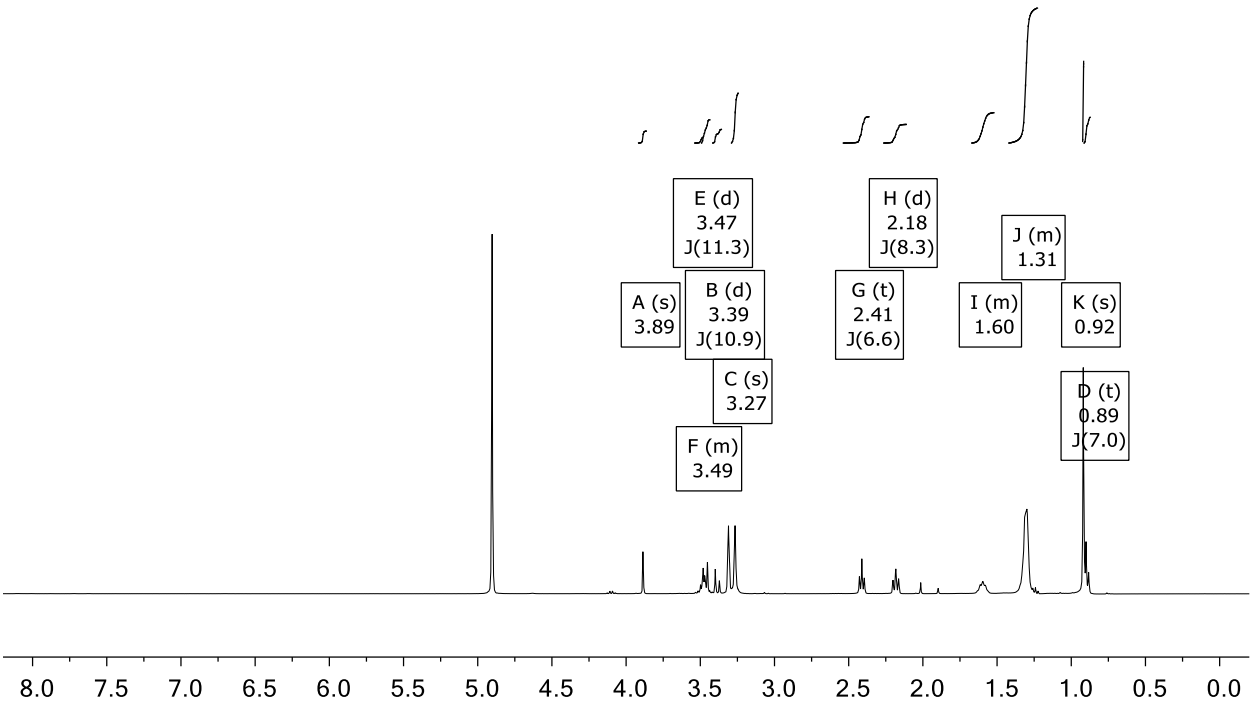


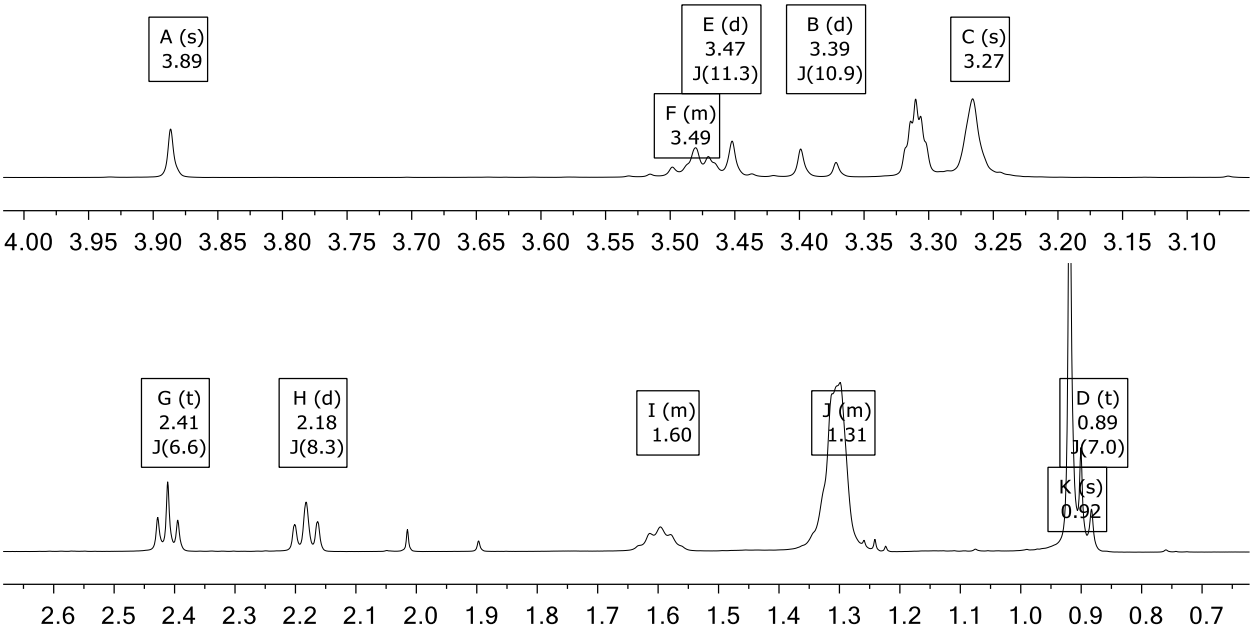

1H-NMR (500 MHz) spectrum of S5c in CDCl3
A (m) 7.43
B (t) 7.03
C (m) 6.92
D (t) 6.73
E (m) 6.29
F (s) 5.46
G (s) 4.08
H (s) 3.82
I (q) 3.68
J (q) 3.53
K (m) 3.34
L (t) 2.42
M (m) 2.14
N (q) 1.59
O (d) 1.26
P (d) 1.09
Q (t) 0.87
f1 (ppm)

^{13}C-NMR (125 MHz) spectrum of **S5c** in $CDCl_3$

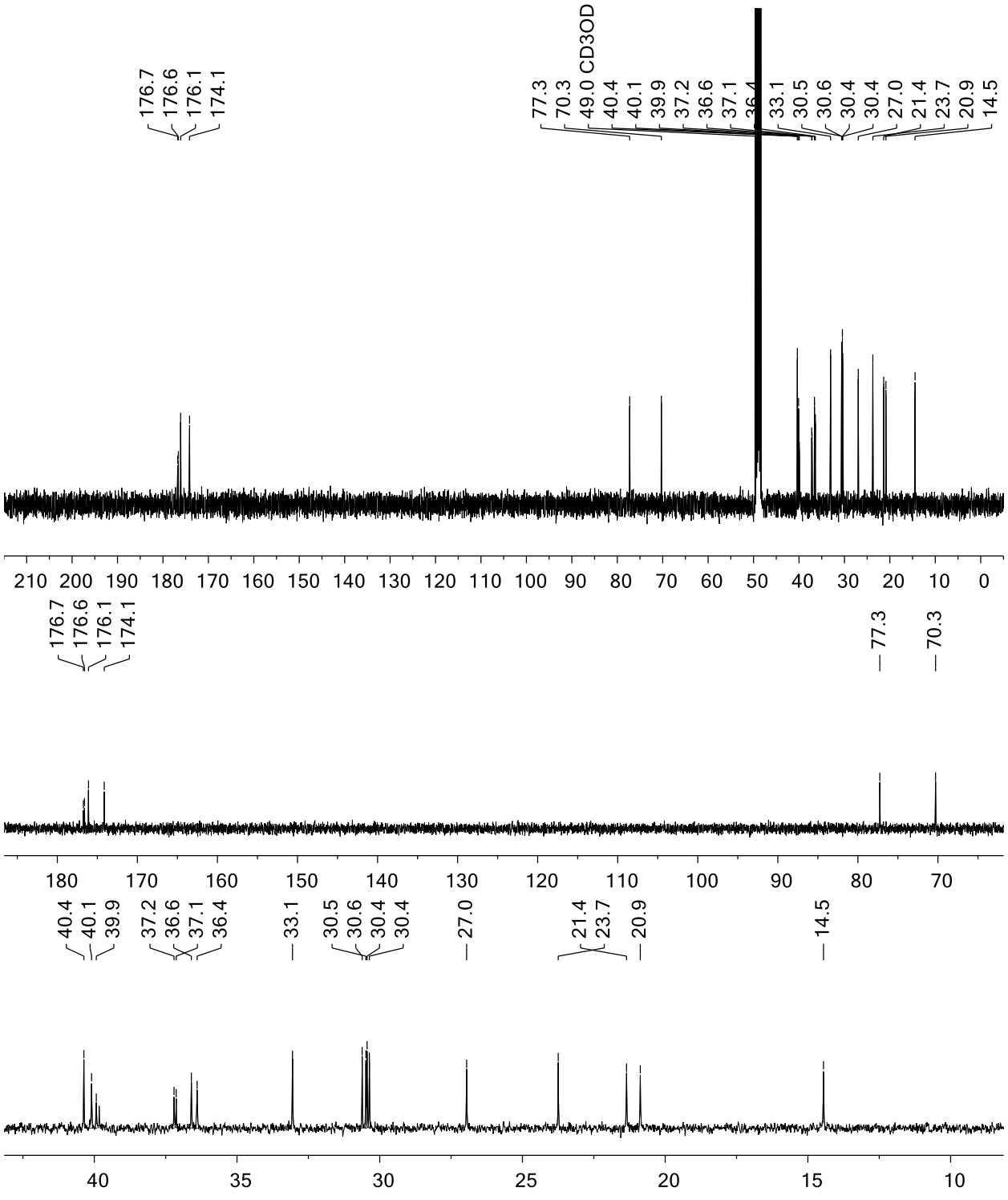


^{1}H-NMR (500 MHz) spectrum of **1c** in CD_3OD





^{13}C-NMR (125 MHz) spectrum of **1c** in CD_3OD



4.4 Acknowledgements

Chapter 4, in full, is currently being prepared for submission for publication of the material as Kalaj, Brianna N.; LaClair, James J.; Shen, Yang.; Schweiters, Charles, D.; Deshmukh, Lalit; Burkart, Michael D. “Quantitative characterization of chain-flipping of acyl carrier protein of *Escherichia coli* using chemical exchange NMR.” The dissertation author was the primary researcher and author of this material.

4.5 References

1. D. M. Byers, H. Gong, Acyl carrier protein: structure–function relationships in a conserved multifunctional protein family. *Biochem. Cell Biol.* **85**, 649-662 (2007).

2. J. Beld, D. J. Lee, M. D. Burkart, Fatty acid biosynthesis revisited: Structure elucidation and metabolic engineering. *Mol. BioSyst.* **11**, 38-59 (2015).

3. C. S. Heil, S. S. Wehrheim, K. S. Paithankar, M. Grininger, Fatty acid biosynthesis: Chain-length regulation and control. *ChemBioChem* **20**, 2298-2321 (2019).

4. J. E. Cronan, The acyl carrier proteins of lipid synthesis are busy having other affairs. *Biochem. J.* **480**, 855-873 (2023).

5. L. Buyachuihan, F. Stegemann, M. Grininger, How acyl carrier proteins (ACPs) direct fatty acid and polyketide biosynthesis. *Angew. Chem., Int. Ed. Engl.* **63**, e202312476 (2024).

6. John E. Cronan, The chain-flipping mechanism of ACP (acyl carrier protein)-dependent enzymes appears universal. *Biochem. J.* **460**, 157-163 (2014).

7. A. Bax, G. M. Clore, Protein NMR: Boundless opportunities. *J. Magn. Reson.* **306**, 187-191 (2019).

8. B. Ramaraju, S. L. Nelson, W. Zheng, R. Ghirlando, L. Deshmukh, Quantitative NMR study of insulin-degrading enzyme using amyloid-β and HIV-1 p6 elucidates its chaperone activity. *Biochemistry* **60**, 2519-2523 (2021).

9. P. Vallurupalli, G. Bouvignies, L. E. Kay, Studying “invisible” excited protein states in slow exchange with a major state conformation. *J. Am. Chem. Soc.* **134**, 8148-8161 (2012).

10. A. G. Palmer, Chemical exchange in biomacromolecules: Past, present, and future. *J. Magn. Reson.* **241**, 3-17 (2014).

11. K. Finzel, D. J. Lee, M. D. Burkart, Using modern tools to probe the structure–function relationship of fatty acid synthases. *ChemBioChem* **16**, 528-547 (2015).

12. T. Sztain, A. Patel, D. J. Lee, T. D. Davis, J. A. McCammon, M. D. Burkart, Modifying the thioester linkage affects the structure of the acyl carrier protein. *Angew. Chem. Int. Ed. Engl.* **58**, 10888-10892 (2019).

13. J. J. Li, "Hofmann rearrangement" in Name reactions: A collection of detailed mechanisms and synthetic applications (Fifth edition). (Springer International Publishing, Cham, 2014), 10.1007/978-3-319-03979-4_136, pp. 319-320.

14. J. Beld, H. Cang, M. D. Burkart, Visualizing the chain-flipping mechanism in fatty-acid biosynthesis. *Angew. Chem., Int. Ed. Engl.* **53**, 14456-14461 (2014).

15. A. Roujeinikova, W. J. Simon, J. Gilroy, D. W. Rice, J. B. Rafferty, A. R. Slabas, Structural studies of fatty acyl-(acyl carrier protein) thioesters reveal a hydrophobic binding cavity that can expand to fit longer substrates. *J. Mol. Biol.* **365**, 135-145 (2007).

16. A. S. Worthington, M. D. Burkart, One-pot chemo-enzymatic synthesis of reporter-modified proteins. *Org. Biomol. Chem.* **4**, 44-46 (2006).

17. Y. Shen, O. Lange, F. Delaglio, P. Rossi, J. M. Armini, G. Liu, A. Eletsky, Y. Wu, K. K. Singarapu, A. Lemak, A. Ignatchenko, C. H. Arrowsmith, T. Szyperski, G. T. Montelione, D. Baker, A. Bax, Consistent blind protein structure generation from NMR chemical shift data. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 4685-4690 (2008).

18. C. D. Schwieters, G. A. Bermejo, G. M. Clore, Xplor-NIH for molecular structure determination from NMR and other data sources. *Protein Sci.* **27**, 26-40 (2018).

19. L. Deshmukh, C. D. Schwieters, A. Grishaev, R. Ghirlando, J. L. Baber, G. M. Clore, Structure and dynamics of full-length HIV-1 capsid protein in solution. *J. Am. Chem. Soc.* **135**, 16133-16147 (2013).

20. L. Deshmukh, J. M. Louis, R. Ghirlando, G. M. Clore, Transient HIV-1 Gag–protease interactions revealed by paramagnetic NMR suggest origins of compensatory drug resistance mutations. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 12456-12461 (2016).

21. L. Deshmukh, C. D. Schwieters, A. Grishaev, G. M. Clore, Quantitative characterization of configurational space sampled by HIV-1 nucleocapsid using solution NMR, X-ray scattering and protein engineering. *ChemPhysChem* **17**, 1548-1552 (2016).

22. T. A. Holak, M. Nilges, J. H. Prestegard, A. M. Gronenborn, G. M. Clore, Three-dimensional structure of acyl carrier protein in solution determined by nuclear magnetic resonance and the combined use of dynamical simulated annealing and distance geometry. *Eur. J. Biochem.* **175**, 9-15 (1988).

23. B. N. Wu, Y. M. Zhang, C. O. Rock, J. J. Zheng, Structural modification of acyl carrier protein by butyryl group. *Protein Sci.* **18**, 240-246 (2009).

24. L. Deshmukh, V. Tugarinov, J. M. Louis, G. M. Clore, Binding kinetics and substrate selectivity in HIV-1 protease−Gag interactions probed at atomic resolution by chemical exchange NMR. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E9855-E9862 (2017).

25. L. Deshmukh, V. Tugarinov, D. H. Appella, G. M. Clore, Targeting a dark excited state of HIV-1 nucleocapsid by antiretroviral thioesters revealed by NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* **57**, 2687-2691 (2018).

26. R. D. Elias, W. Am, R. Ghirlando, C. D. Schweiters, V. S. Reddy, L. Deshmukh, Proline-rich domain of human ALIX contains multiple TSG101-UEV interaction sites and forms phosphorylation-mediated reversible amyloids. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 24274-24284 (2020).

27. Y. Shen, A. Bax, Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology. *J. Biomol. NMR* **38**, 289-302 (2007).

28. L. Yao, A. Grishaev, G. Cornilescu, A. Bax, The impact of hydrogen bonding on amide ^{1}H chemical shift anisotropy studied by cross-correlated relaxation and liquid crystal NMR spectroscopy. *J. Am. Chem. Soc.* **132**, 10866-10875 (2010).

29. S. Paramasivam, A. M. Gronenborn, T. Polenova, Backbone amide ^{15}N chemical shift tensors report on hydrogen bonding interactions in proteins: A magic angle spinning NMR study. *Solid State Nucl. Magn. Reson.* **92**, 1-6 (2018).

30. H. M. McConnell, Reaction rates by nuclear magnetic resonance. *J. Chem. Phys.* **28**, 430-431 (1958).

31. G. A. Zornetzer, B. G. Fox, J. L. Markley, Solution structures of spinach acyl carrier protein with decanoate and stearate. *Biochemistry* **45**, 5217-5227 (2006).

32. G. A. Zornetzer, J. Tanem, B. G. Fox, J. L. Markley, The length of the bound fatty acid influences the dynamics of the acyl carrier protein and the stability of the thioester bond. *Biochemistry* **49**, 470-477 (2010).

33. T. Sztain, T. G. Bartholow, J. A. McCammon, M. D. Burkart, Shifting the hydrolysis equilibrium of substrate loaded acyl carrier proteins. *Biochemistry* **58**, 3557-3560 (2019).

34. A. Roujeinikova, C. Baldock, W. J. Simon, J. Gilroy, P. J. Baker, A. R. Stuitje, D. W. Rice, A. R. Slabas, J. B. Rafferty, X-ray crystallographic studies on butyryl-ACP reveal flexibility of the structure around a putative acyl chain binding site. *Structure* **10**, 825-835 (2002).

35. N. M. Kosa, R. W. Haushalter, A. R. Smith, M. D. Burkart, Reversible labeling of native and fusion-protein motifs. *Nat. Methods* **9**, 981-984 (2012).

36. T. Sztain, T. G. Bartholow, D. J. Lee, L. Casalino, A. Mitchell, M. A. Young, J. Wang, J. A. McCammon, M. D. Burkart, Decoding allosteric regulation by the acyl carrier protein. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2025597118 (2021).

37. J. T. Mindrebo, A. Patel, W. E. Kim, T. D. Davis, A. Chen. T. G. Bartholow, J. J. La Clair, J. A. McCammon, J. P. Noel, M. D. Burkart, Gating mechanism of elongating β-ketoacyl-ACP synthases. *Nat. Commun.* **11**, 1727 (2020).

38. J. T. Mindrebo, A. Chen, W. E. Kim, R. N. Re, T. D. Davis, J. P. Noel, M. D. Burkart, Structure and mechanistic analyses of the gating mechanism of elongating ketosynthases. *ACS Catal.* **11**, 6787-6799 (2021).

39. S. C. Kamerlin, A. Warshel, At the dawn of the 21st century: Is dynamics the missing link for understanding enzyme catalysis? *Proteins* **78**, 1339-1375 (2010).

40. S. D. Schwartz, Protein dynamics and enzymatic catalysis. *J. Phys. Chem. B* **127**, 2649-2660 (2023).

41. G. M. Clore, D. S. Garrett, R-factor, free R, and complete cross-validation for dipolar coupling refinement of NMR structures. *J. Am. Chem. Soc.* **121**, 9008-9012 (1999).

42. S. Salentin, S. Schreiber, V. J. Haupt, M. F. Adasme, M. Schroeder, PLIP: fully automated protein-ligand interaction profiler. *Nucleic Acids Res.* **43**, W443-447 (2015).

S1. J. Beld, H. Cang, M. D. Burkart, Visualizing the chain-flipping mechanism in fatty-acid biosynthesis. *Angew. Chem., Int. Ed. Engl.* **53**, 14456-14461 (2014).

S2. K. Charov, M. D. Burkart, Quantifying protein-protein interactions of the acyl carrier protein with solvatochromic probes. *Methods Enzymol.* **638**, 321-340 (2020).

S3. A. S. Worthington, M. D. Burkart, One-pot chemo-enzymatic synthesis of reporter-modified proteins. *Org. Biomol. Chem.* **4**, 44-46 (2006).

S4. M. M. Nakano, N. Corbell, J. Besson, P. Zuber, Isolation and characterization of sfp: a gene that functions in the production of the lipopeptide biosurfactant, surfactin, in *Bacillus subtilis*. *Mol. Gen. Genet.* **232**, 313-321 (1992).

S5. R. W. Haushalter, A. S. Worthington, G. H. Hur, M. D. Burkart, An orthogonal purification strategy for isolating crosslinked domains of modular synthases. *Bioorg. Med. Chem. Lett.* **18**, 3039-3042 (2008).

S6. R. D. Elias, W. Am, R. Ghirlando, C. D. Schweiters, V. S. Reddy, L. Deshmukh, Proline-rich domain of human ALIX contains multiple TSG101-UEV interaction sites and forms

phosphorylation-mediated reversible amyloids. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 24274-24284 (2020).

S7. R. D. Elias, B. Ramaraju, L. Deshmukh, Mechanistic roles of tyrosine phosphorylation in reversible amyloids, autoinhibition, and endosomal membrane association of ALIX. *J. Biol. Chem.* **297**, 101328 (2021).

S8. B. Ramaraju, S. L. Nelson, W. Zheng, R. Ghirlando, L. Deshmukh, Quantitative NMR study of insulin-degrading enzyme using amyloid-β and HIV-1 p6 elucidates its chaperone activity. *Biochemistry* **60**, 2519-2523 (2021).

S9. F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277-293 (1995).

S10. W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins: Struct. Funct. Genet.* **59**, 687-696 (2005).

S11. G. M. Clore, A. M. Gronenborn, Determining the structures of large proteins and protein complexes by NMR. *Trends Biotechnol.* **16**, 22-34 (1998).

S12. N. C. Fitzkee, A. Bax, Facile measurement of ^{1}H-^{15}N residual dipolar couplings in larger perdeuterated proteins. *J. Biomol. NMR* **48**, 65-70 (2010).

S13. C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65-73 (2003).

S14. C. D. Schwieters, G. A. Bermejo, G. M. Clore, Xplor-NIH for molecular structure determination from NMR and other data sources. *Protein Sci.* **27**, 26-40 (2018).

S15. N. A. Lakomek, J. Ying, A. Bax, Measurement of ^{15}N relaxation rates in perdeuterated proteins by TROSY-based methods. *J. Biomol. NMR* **53**, 209-221 (2012).

S16. P. Vallurupalli, G. Bouvignies, L. E. Kay, Studying "invisible" excited protein states in slow exchange with a major state conformation. *J. Am. Chem. Soc.* **134**, 8148-8161 (2012).

S17. M. Guenneugues, P. Berthault, H. Desvaux, A method for determining B1 field inhomogeneity. Are the biases assumed in heteronuclear relaxation experiments usually underestimated? *J. Magn. Reson.* **136**, 118-126 (1999).

S18. D. F. Hansen, P. Vallurupalli, L. E. Kay, An improved ^{15}N relaxation dispersion experiment for the measurement of millisecond time-scale dynamics in proteins. *J. Phys. Chem. B* **112**, 5898-5904 (2008).

S19. L. Deshmukh, V. Tugarinov, J. M. Louis, G. M. Clore, Binding kinetics and substrate selectivity in HIV-1 protease−Gag interactions probed at atomic resolution by chemical exchange NMR. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E9855-E9862 (2017).

S20. H. M. McConnell, Reaction rates by nuclear magnetic resonance. *J. Chem. Phys.* **28**, 430-431 (1958).

S21. S. Raman, O. Lange, P. Rossi, M. Tyka, X. Wang, J. Aramini, G. Liu, T. A. Ramelot, A. Eletsky, T. Szyperski, M. A. Kennedy, J. Prestegard, G. T. Montelione, D. Baker, NMR structure determination for larger proteins using backbone-only data. *Science* **327**, 1014-1018 (2010).

S22. Y. Shen, O. Lange, F. Delaglio, P. Rossi, J. M. Armini, G. Liu, A. Eletsky, Y. Wu, K. K. Singarapu, A. Lemak, A. Ignatchenko, C. H. Arrowsmith, T. Szyperski, G. T. Montelione, D. Baker, A. Bax, Consistent blind protein structure generation from NMR chemical shift data. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 4685-4690 (2008).

S23. R. Das, I. Andre, Y. Shen, Y. Wu, A. Lemak, S. Bansal, C. H. Arrowsmith, T. Szyperski, D. Baker, Simultaneous prediction of protein folding and docking at high resolution. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 18978-18983 (2009).

S24. J. T. Mindrebo, A. Patel, W. E. Kim, T. D. Davis, A. Chen. T. G. Bartholow, J. J. La Clair, J. A. McCammon, J. P. Noel, M. D. Burkart, Gating mechanism of elongating β-ketoacyl-ACP synthases. *Nat. Commun.* **11**, 1727 (2020).

S25. A. Roujeinikova, W. J. Simon, J. Gilroy, D. W. Rice, J. B. Rafferty, A. R. Slabas, Structural studies of fatty acyl-(acyl carrier protein) thioesters reveal a hydrophobic binding cavity that can expand to fit longer substrates. *J. Mol. Biol.* **365**, 135-145 (2007).